3D Data Creation to Curation:

Community Standards for 3D Data Preservation

edited by
Jennifer Moore,
Adam Rountrey, and
Hannah Scates Kettler

3D Data Creation to Curation:

Community Standards for 3D Data Preservation

edited by
Jennifer Moore,
Adam Rountrey, and
Hannah Scates Kettler

Association of College and Research Libraries
A division of the American Library Association
Chicago, Illinois 2022

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences–Permanence of Paper for Printed Library Materials, ANSI Z39.48-1992. ∞

Library of Congress Control Number: 2022938940

Cover design by Ian Lanius.

Printed in the United States of America.

26 25 24 23 22 5 4 3 2 1

Contents

Acknowledgments

All of this work would not have been possible without the support of the Institute of Museum and Library Services (IMLS), the partner institutions (Washington University in St. Louis, the University of Iowa, Iowa State University, and the University of Michigan), and all of the individuals who willingly devoted their time, whether in person or virtually, over the last few years to the continued development of CS3DP. Contributions included event support, conversations, writing, presentations, research, and reviewing. We thank Judith Lauber for her invaluable and thorough editorial guidance as well as Carol Abraczinskas of the University of Michigan Museum of Paleontology for designing the CS3DP logo and Ian Lanius of Washington University in St. Louis for the book cover design. The IMLS-supported CS3DP Forums have been graciously archived by the University of Iowa Institutional Repository.[1]

As this work was, by and large, created by a group of experts scattered across the globe volunteering their time to the development of this communal effort, it is difficult to put into words the gratitude we have as editors and CS3DP organizers for the enthusiasm and support we have received from all of our collaborators—as everyone who has contributed to the book's development is indeed a true collaborator. We are so grateful for your perspective. You have made this work not only thoroughly enjoyable but invigorating.

community standards
for 3D data preservation

Forum 1 Attendees	Forum 2 Attendees
Andrea d'Andrea	Francesca Albrezzi
Francesca Albrezzi	Jon Blundell
Kari Allen	Justin Bonfiglio
Anne Austin	John Bonnett
Diana Bell	Susan Borda
Stephanie Bemberg	Douglas Boyer
Jon Blundell	Tom Bray
Douglas Boyer	Beverly Chudo
Jon Caris	Perry Collins
Beth Caruso	Michael Conyers
Jasmine Clark	Kyle Courtney
Michael Conyers	Charles Crable
Kyle Courtney	Chelsea Dantuma
Katherine DeVet	Sean Davis
Patti Wood Finkle	Katherine DeVet

Forum 1 Attendees	Forum 2 Attendees
Tassie Gniady	Emily Finch
Kristina Golubiewski-Davis	Jon Frey
Nathan Hall	Tassie Gniady
Juliet L. Hardesty	Kristina Golubiewski-Davis
Christina Harlow	Stephanie Grimes
Cynthia Hudson Vitale	Juliet L. Hardesty
Chuk Ikebude	Sarah Hutton
Lisa Johnston	Martin Jones
Mica Jones	Christin Lampkowski
Deb Katz	Brandon Locke
Monique Lassere	Nichole Lohrke
Melissa Levine	Henry Lovejoy
Shu-Wen Lin	Jennifer Moore
Holly Little	Angela Murillo
Jessica Maisano	Kieron Niven
Narcise Mbunzama	Brandon Patterson
Marcia McIntosh	Samantha Porter
Frank McManamon	Nelson Rios
Jennifer Moore	John Ristevski
Ashley Morhardt	Adam Rountrey
Gary Motz	Albert Rozo
Angel Nieves	Hannah Scates Kettler
Kieron Niven	Gurpreet Singh
Todd Ogle	Eunah Snyder
Heather Richards-Rissetto	Lisa Snyder
Vince Rossi	Rebecca Snyder
Adam Rountrey	Stu Snydman
Will Rourk	Todd Swanson
Hannah Scates Kettler	Cody Thompson
Carla Schroer	Amanda Tickner
Pat Seed	Kate Topham
Paul Segura	Jocelyn Triplett
Lisa Snyder	Julie Winchester
Rebecca Snyder	Scott Witmer
Todd Swanson	Yuanxiao Xu
Andrea Thomer	Ayoung Yoon
Mollie Webb	
Kate Webbink	
Brian Westra	
Ann Whiteside	
Susan Wiesner	
Bill Winston	
Jamie Wittenberg	
Micah Zeller	

Note

1. “Community Standards for 3D Preservation (CS3DP),” University Libraries, University of Iowa, accessed May 4, 2020, https://ir.uiowa.edu/cs3dp/.

Chapter 1

Introduction

Jennifer Moore, Adam Rountrey, and Hannah Scates Kettler

Context for This Work

There has been rapid growth in the production and usage of 3D data over the last decade, yet the **preservation**[*] of these data has lagged behind to the detriment of scholarship and innovation. While the need for digital 3D data preservation is widely recognized, the ongoing development of 3D data creation processes and the evolving usage of content still present many open-ended questions about how to ensure the stability and durability of this data type. Creators, curators, and users of 3D datasets are disadvantaged by the lack of shared guidelines, practices, and standards.[1] This volume, which includes surveys of current practices, recommendations for implementation of standards, and identification of areas in which further development is required, is a result of the efforts of a large practicing community coming together under the Community Standards for 3D Data Preservation (CS3DP) initiative to move toward establishment of standards.[2] The goal of this work is to identify the broad, shared preservation needs of the whole community, and it is viewed as essential to use a collaborative approach for standards development that promotes individual investment and broad adoption. The authorship of the chapters recognizes those who worked to discuss particular aspects of preservation in detail, but throughout the process of development, the entire community has been engaged, shaping the content to meet needs across a diverse base of stakeholders.

---

* All terms in **bold type** are defined in the glossary.

The Democratization of 3D Data Production

While a detailed history of 3D data creation and preservation is beyond the scope of this introductory chapter, it seems appropriate to at least provide a brief summary of how we reached the present situation in which efforts to develop accepted preservation strategies are lagging substantially behind efforts to create and capture new 3D data.

Digital 3D data have been in use nearly as long as digital computers have been available, with foundational concepts in **computer-aided design** and display of digital modes being developed by Ivan Sutherland and others at MIT in the early 1960s.[3] It is important to recognize that the usefulness of digital 3D data in many applications depends critically on having a system that allows users to display and interact with the data. While the original interactive Sketchpad software developed by Sutherland handled only 2D drawing and manipulations, a colleague, Lawrence Roberts, soon extended it to enable display and manipulation of simple digital 3D solids.[4] For some potential applications of this technology, the usefulness of the digital models is dependent on their realistic **rendering** (the process of producing a 2D image of the 3D data for viewing), and it would be several years before more realistic, shaded renderings of digital 3D models would be achieved. In 1967, with a goal of producing renderings that could be intuitively "felt" rather than laboriously interpreted, Wylie and colleagues created a system for displaying halftone shaded renderings of digital 3D objects.[5] They also promoted the now-common use of triangulated **meshes** to represent surfaces, which is computationally efficient. Methods of producing shaded renderings continued to improve, and in the early 1970s it became possible to display realistically more complicated digital models, including those based on real physical objects. Among the earliest "digitized" real world objects rendered in this way by a University of Utah group were the face of Sylvie Gouraud (before June 1971, via orthophotos of polygons drawn on her face), Edwin Catmull's left hand (1972, via polygons drawn on a plaster cast, probably measured with a coordinate measuring machine), and a Volkswagen Beetle (1972, via polygons drawn on the car, measured with rulers, strings, levels, and a volleyball stanchion).[6] Advances in rendering would continue in the following decades, making the use of digital 3D data applicable to a wider array of fields.

In the 1960s and 1970s, creation (and use) of digital 3D data was difficult and time-consuming, and from the present perspective, this was still the case in the late 1990s, when the groundbreaking Digital Michelangelo Project was initiated at Stanford with the goal of creating a digital model of Michelangelo's *David* (and other works) that would be sufficiently accurate and detailed to allow their use in research. Obtaining an accurate digital 3D model of Michelangelo's *David* in 1998–1999 required years of planning, the design and fabrication of a custom laser scanner, a team of twenty-two people, 1,080 person-hours of scanning, and over 1,500 person-hours of post-processing.[7] With so much invested in the creation of early digital 3D models like this, it is not surprising that the issues related to preservation were immediately recognized.[8] Many of the core problems identified by CS3DP participants in recent years are similar to those raised by the Stanford team twenty years ago. In 2000, Levoy and Garcia-Molina highlighted the cost of storage, the complexity of **metadata**, and difficulties related to licensing and distribution of digital 3D data.[9] The Digital Michelangelo Project Archive, which still offers raw data, derivatives, and descriptions for the project outputs, was implemented to preserve and distribute the data.[10]

Where preservation has been a passive process, the results have been mixed. The previously mentioned Volkswagen Beetle data produced forty-eight years ago appear to have been preserved in a slightly modified form, based on comparisons between a 1972 wireframe rendering and a digital 3D data file obtained from a no-longer-operational online service (3DCafe.com) via the Internet Archive.[11] Sutherland suggests that no special considerations were given to preserving the data, but that once the digital data passed to several users, they survived anyway.[12] Indeed, at a time when few digital 3D "test" models were available for study, distribution and reuse of the data probably led to relative safety in a large number of copies. Once the file was made available on the web, it was preserved by the Internet Archive. While the apparent existence of some of the earliest digital 3D data is remarkable, their preservation is notably incomplete. Unlike the Michelangelo Project data, there is no chain of custody or metadata associated with the Volkswagen Beetle data, making it impossible to say with certainty that they are, in fact, genuine.

Levoy and Garcia-Molina noted that online storage of the 500-gigabyte Michelangelo project database was prohibitively expensive for most libraries at the time.[13] While per-gigabyte storage costs have decreased dramatically since 2000, the size and number of digital 3D models being produced have increased in unforeseen ways, leaving the community facing essentially the same storage problems. With the appearance of more affordable, fast, user-friendly methods of data acquisition, such as desktop and handheld laser scanners in the early 2000s and GPU-driven **photogrammetry** in the early 2010s, rates of digital 3D data creation have increased rapidly. High-**resolution** digital 3D models of real objects are now created in minutes at little cost, driving efficiencies and advances in fields like paleontology and archaeology, as well as in the engineering and entertainment industries. Adding to the digital 3D data load is the expansion of use of CT scanners in nonmedical research. All of these advances in technology and accessibility resulted in the democratization of 3D digitization before standards or preservation practices were well established, leaving 3D practitioners in a holding pattern, accumulating data and creating bespoke solutions.

The Audience

The CS3DP project was established to bring together people with diverse backgrounds and experiences with digital 3D data (henceforth, 3D data)* to examine the current practices in 3D data documentation, dissemination, and preservation and to make recommendations for standardization that could bring broad adoption and benefits. For example, efforts that seek to preserve physical objects through digitization are not worthwhile without means of preserving digital products, but this is one of many contexts where 3D preservation is beneficial. The people who make up the community have backgrounds in art, architecture, natural history, information science, medicine, archaeology, and law, and there are representatives from academic institutions, nonprofits, and commercial industries mostly based in the United States; they also have experience with a diversity of approaches. Given the potential scope of the work that this book attempts to cover and complexities arising from different legal frameworks in other countries, it is written for a primarily US audience, although much of what is covered will no doubt be applicable to those outside the US. This work likely has relevance beyond the borders, and the authors did attempt to make reference to related work going on in other areas of the world. It should also be noted that the discussions tended to focus on 3D applications in academic research, cultural heritage, and education, but the resulting material will still be useful to those operating outside of educational contexts.

These pages are intended to be used by people with varying amounts of 3D experience, from novice to seasoned practitioner, and are also intended for people in 3D data preservation support roles, who may or may not be involved in the creation of the data, yet may be tasked with curating, migrating, and sustaining **access** to these data long-term. This last task, sustaining access to these data, is the crux of the problem. Regardless of our intent, whether digitizing a physical space or object, representing imaginative spaces, or creating 3D for entertainment, without some kind of access to these data, they likely will escape preservation. 3D models can be expensive consumables or flashy ephemera, yet in some cases, such as with entities like CyArk that are meant to empower the collective preservation of cultural heritage material in 3D, that ethos can be undermined by a lack of perpetual access and long-term preservation.[14] Access and preservation, and the steps in between, are inseparable from each other. Though this work is indeed focused on preservation, the pages within reflect the entire life cycle of 3D data creation and maintenance, underpinned by concerns of access by various user groups.

* The glossary included in this book was developed collaboratively out of necessity. We quickly realized that in order to understand each other as we spoke about our 3D data work across the various disciplines and modalities, we needed a set of common terms that could be used across these boundaries.

The Creators

In order to create this resource, the CS3DP community had to come together several times over the course of two years to develop a report across disciplinary boundaries as well as 3D creation boundaries. In many cases, common vocabulary needed to be developed to facilitate dialogue and understanding to move the conversation beyond simply acknowledging the issues of **digital preservation**. Only then could the community address the issues of 3D data preservation.

This volume builds upon previous work, such as the London Charter, the Guides to Good Practice from the Archaeological Data Service (ADS), and 3D-ICONS, to address broad concerns related to 3D data preservation. The CS3DP, representing the expertise across the 3D data life cycle from creation to curation, has forged ahead to identify and discuss specific needs, potential outcomes, and recommendations for 3D data preservation, moving the conversation from acknowledgement to action. This volume will act as a guide for implementing real changes in the way we approach 3D data creation and preservation by presenting preservation in the context of representative **workflows** that translate across disciplines and across institutions. By doing so, we set the stage for adaptation and broad applicability to sustain progress in 3D preservation as a network of invested stakeholders.

Before CS3DP, much dialogue was conducted in silos, be they silos of disciplinary expertise (archaeology or architecture, for example) or data creation method (photogrammetry versus laser scanning), rather than looking at 3D data creation and preservation as a whole. Even in its beginning, CS3DP attempted to focus solely on the preservation of the bits and bytes of 3D data, but we quickly learned that it was impossible to detangle the preservation aspects of this work from the creation, use, access, and discoverability of these data. Thus, this multifaceted effort was born and included the simultaneous development and discussion in five areas; overall preservation best practices, metadata creation, access concerns, management and storage, and the rights and ownership related to these data.

The authors and contributors of this text are many and represent voices and expertise from some of the aforementioned projects, as well as from LIB3DVR (IMLS grant project), Building 4 Tomorrow (IMLS grant project), Advanced Challenges in Theory and Practice in 3D Modeling of Cultural Heritage Sites (NEH),[15] and nonprofit organizations such as Cultural Heritage Imaging and CyArk. By bringing all together for collaborative conversations, CS3DP participants were able to communicate concerns and needs across procedural and disciplinary boundaries. This book is a single foundational resource for 3D data preservation, demonstrating from its authorship an approach for 3D data preservation agnostic to platform, discipline, and methods.

Values of CS3DP

CS3DP aims to be open, radically inclusive, and collaborative, and it was developed for and is governed by its members. From the outset, the project aimed to garner as much community input and support as possible. While there was some limited understanding of the community as a result of the data we gathered through our 2017 survey, in the project proposal we outlined a community of practice (CoP) model to build our collective understanding and investment. The advantage of bringing together members from diverse domains and expertise to share knowledge and perspective is that it allows us to build the best possible understanding, recommendations, and ultimately agreement on how to move forward. This means not just hearing from our loudest, most privileged, most cited voices, but also from those who have not yet emerged, have been overlooked, or have been ignored, and this is reflected in our use of some authorship recommendations from McNutt and colleagues.[16]

In our first forum, the official kickoff of the CoP, we made an effort to gather a panel that would lay the foundations for group discussions. Forum 1 participants came from diverse geographic, economic, and disciplinary backgrounds, and the forum was meant to cultivate relationships between participants and develop a sense of shared ownership. To facilitate in- depth discussion and writing, smaller working groups were established during the first forum, and they remained open to new participants throughout the discussion and writing process. The working groups are responsible for the chapters in this volume, which address specific areas around foundational questions that were identified in a whole-group town hall. The five working groups established were best practices, management, rights and ownership, metadata, and access and **discovery**. The groups were given a charge designed to provide scope, roles, and expectations about working together and communicating. From there on out, working groups took the reins and reported back to the group on a monthly basis. The CS3DP participants have endeavored to do research, have discussions, and make decisions as a group. At Forum 2 we heard from groups about progress to date and together scoped our next steps. This second forum was focused on refining what working groups were learning and shaping that into the recommendations.

The writers and editors recognize the importance of ethical considerations around the ownership, dissemination, attribution, and general treatment of 3D objects and data in our collections and virtual environments. There were indeed many discussions about ethics at the CS3DP forum events in 2018, but we did not form an independent ethics work group during our first forum. While there was interest in the topic, we came to understand that without interdisciplinary expertise grounded in the study of ethics, we could not address it with appropriate scope. While many chapters herein may touch on various ethical concerns, we recognize we have not and cannot cover all ethical considerations for 3D data preservation in this work. Our hope is to see focused development of 3D Data Ethics following this publication.

Given that the recommendations are intended for a diverse audience, from the outset we considered it essential that the CS3DP recommendations be developed by a diverse group. The intention is that it be a valuable resource regardless of discipline and experience—for those who are new to 3D data work, but also for those who've been working with it for decades. Chapter 2, "Best Practices for 3D Data Preservation," introduces the **Good/Better/Best** approach to the recommendations, which intends to provide scalability and meet practitioners where they are. Our collective experience demonstrated to the group that all-or-nothing recommendations are often out of reach for many. Instead, we are advocating for attainable, extensible recommendations rather than one-size-fits-all solutions.

We also recognize that even more diverse perspectives will further enliven this work. Having voices that represent cultures affected by colonization is very important to the growth of the community and creating truly ethical recommendations. In Forum 1, Angel Nieves asserted that "preservation is an act of social justice." He described how digital reconstructions of destroyed cultural landscapes provide an opportunity for reconciliation. This is true only if reconstructions are preserved for long-term access.

From Creation to Preservation

The process of preserving 3D data increases FAIRness (findability, accessibility, interoperability, reuse). As mentioned above, working groups approached the problem through their respective lenses: best practices planning, metadata and documentation, long-term repository storage and management, articulation of rights and ownership, and restrictions and access. These topics are distinct and interconnected; they make up the framework for this volume and will be presented as independent chapters, which reference and leverage their companion chapters.

To begin to contextualize the problem, there are two main branches of 3D model generation: reality capture and manual modeling. Reality capture means the creator has a physical thing that they would like to replicate using a camera or scanner. Within the reality capture branch are two more major distinctions: whether the model is volumetric or a surface capture. **Volumetric data** are data that include measurements or other values in a 3D array or grid. A common method that produces volumetric data is computed tomography (CT) scanning. Surface-based captures aim to digitally reproduce the shape of the object in 3D but do not collect density information. Manual modeling also often creates a 3D surface object, but not via a digital capture. Common manual creations are results of creative modeling and reference-based modeling. Models can also be created by combining multiple methods. Specific methods or modes of creation will be described in the next section.

In the best case scenario, the life cycle of 3D data begins, as much data creation does, with planning. Among the questions that need addressing in the planning stage are what is the purpose of generating a 3D object, what is the origin information that will contribute to its creation, and what mode of creation best fits the purpose and origin. Often the purpose and originating information guide the creation tools and methods. Chapter 2, "Best Practices for 3D Data Preservation," will articulate how careful planning can impact the preservation of data for various modes of creation. That chapter will plant many seeds that will flourish throughout the rest of the volume.

Modalities Represented in the Chapters

There are two broad digital data types covered in this book: (1) data that may include 3D points, edges, and faces, such as a polygonal mesh representing the surface of, for example, a scanned statue, and (2) volumetric, or **voxel**, data, which are a 3D array or grid with values assigned to cells in the grid (e.g., CT scan data). There are a variety of ways in which data in these two types are produced. Some of the most common methods are listed below.

- 3D point/mesh data
  - Photogrammetry: The extraction of three-dimensional measurements from two-dimensional data (i.e., images). Developments in GPU-based processing allow rapid reconstruction of 3D surface meshes from sets of conventional photographs of a physical object or environment. The mesh output from this technique may be enhanced by color information at vertices (i.e., vertex color) or an associated 2D image representing surface color, which is mapped to the mesh (**texture map**).
  - Laser scanning: The process of recording precise three-dimensional information about a real-world object or environment by rapidly sampling or scanning an object's surface with lasers. The information is often returned to the user as a dense collection of precisely located *x,y,z* coordinates referred to as a **point cloud**. Laser scanning devices may use a time-of-flight method, a phase method, or a triangulation method. The point cloud or mesh output from this technique may be enhanced by color information at points/vertices (i.e., vertex color) or an associated 2D image representing surface color, which is mapped to the mesh (texture map).
  - **Structured light**: Method of 3D capture that relies on the distortion of projected light to calculate surface form. A known pattern (often a grid or horizontal lines) of light projected onto a surface appears distorted from perspectives other than that of the projector. This distortion can be used

for geometric reconstruction of the surface shape. The mesh output from this technique may be enhanced by color information at points/vertices (i.e., vertex color) or an associated 2D image representing surface color, which is mapped to the mesh (texture map).
    - Bibliography/sources-based modeling: A method of model production based on documents, reference photographs, or other sources of information about a real-world object or place. Models are often created in a CAD (computer-aided design) software system.
    - Creative modeling: A method of model production in which the user designs a 3D object or environment based on creative vision.
- Volumetric data
    - CT scanning: Also known as a computed tomography scan, and formerly known as a computerized axial tomography scan or CAT scan. This method makes use of computer-processed combinations of many X-ray measurements taken from different angles to produce cross-sectional (tomographic) images (virtual "slices") of a scanned object, allowing the user to see inside the object without cutting. Other medical imaging methods (e.g., MRI) produce similar volumetric data based on different properties of the object.
    - Voxel art: A method of modeling in which objects are represented by many 3D cubes that may vary in color (e.g., "Minecraft style").
- **Multimodal modeling**
    - 3D model resulting from a combination of methods (modes).

While dividing these data types by representation (points and meshes versus volumetric data) is useful for many parts of 3D data preservation planning, there are times when it is more useful to separate data by whether or not they are a faithful attempt at recording the geometry and characteristics of a measurable (at the time of capture) real-world object or environment, a faithful attempt at recreating geometry and characteristics of a real but nonmeasurable object or environment, or a creative output in which expression is more important than representation of a real-world object or environment. In this book, we will refer to reality-capture models, sources-based models, and creative models respectively to differentiate these when appropriate.

Many of the methods mentioned above will be covered in more detail in chapter 2, "Best Practices for 3D Data Preservation," and referenced throughout the book. As much as the authors could, those different modalities are acknowledged and considered, and the discussions and recommendations are made with the myriad pathways of 3D data creation in mind. We chose to focus mostly on commonly used methods that do not have well-established preservation standards, so some types of data, such as medical CT and 3D GIS data, are not discussed in detail.

What to Expect

Best Practices for 3D Data Preservation

Chapter 2, "Best Practices for 3D Data Preservation," begins by describing the triad of digital preservation (management, technology, content) in its relation to 3D data preservation and introduces three tiers of preservation to which institutions can aspire. These tiers will be used in other chapters throughout the book: good, better, and best practices for 3D preservation. The concept of **preservation intervention points** (PIPs) is introduced as a method by which practitioners can address preservation actions throughout the data life cycle, and those points considered most relevant (planning, collection or creation, curation, and long-term access) to 3D data preservation are unpacked.

PIPs are described as project-dependent and related to the purpose of data generation, target audiences, preservation needs, and imaging modality. However, the chapter outlines general Good/Better/Best guidelines that may be useful to both new and established practitioners to consider adopting at the level that suits that institution's resource level.

Management and Storage of 3D Data

Chapter 3, "Management and Storage of 3D Data," aims to discuss the unique features of 3D data management, how **repository managers** and creators are currently wrestling with 3D, and what recommendations—based on current data management standards—should be upheld in 3D data preservation. The chapter opens by acknowledging that all general digital preservation principles apply. It references the "triad of preservation" and explains how it applies to the 3D context. The group conducted two surveys of 3D-specific and other repositories that hold 3D data. Authors walk the reader through the methodology and briefly describe results of the survey, which are articulated more fully later. Respondents provided details on the systems and platforms used to store 3D data, how preservation packages are composed, what cost models are used to finance creation and storage, and other areas of interest, including data types, citations, and retention methods. The chapter describes how the management of 3D data differs from other data types, how existing managers are managing their 3D data, and what standards serve as best practice in the work of preserving these assets.

Metadata Requirements for 3D Data

As discussed in chapter 4, "Metadata Requirements for 3D Data," preservation and accessibility of 3D data depend on the use of metadata to document characteristics of the data related to their creation, management, distribution, retrieval, and archiving, the components of the **digital asset life cycle**. The broad types of 3D data considered in this chapter are reality-capture models (representations of real-world objects), sources-based models (models based on sources such as documents and photographs), and creative models (models based on the creative process of the artist). It should be recognized that, while 3D data have some unique metadata requirements that are not currently covered by existing standards, many parts of the documentation system can be borrowed from more generalized standards such as Dublin Core, Darwin Core, or MODS. To identify specific needs for 3D data, the authors of this chapter surveyed stakeholders currently involved in the creation and curation of 3D data regarding current practices and workflows and assembled recommended generalized metadata fields at the good, better, and best levels. This flexibility in recommendation level was chosen in recognition of the spectrum of resources available to entities with interests in preserving 3D data. Recommendations for 3D data are in the form of metadata types needed in a given context and can be used as guides in identifying particular metadata standards and controlled vocabularies for use with 3D data.

The standards, recommendations, and further needed inquiry identified in this volume are shared to provide a foundation from which to further build truly shared, adaptable, and flexible standards for the 3D data community of users and creators. With a nod to the contexts from which we all arose, we've included an appendix reminding us of a selection of disciplines' and societies' (albeit US-focused) ethical grounding that may further influence any application of said standards in our various contexts.

Copyright and Legal Issues Surrounding 3D Data

Legal concepts related to the ownership and licensing of 3D can be complex, and chapter 5, "Copyright and Legal Issues Surrounding 3D Data," clarifies the legal context by giving an accessible overview of US copyright law and highlighting relevant case law. What do phone books have to do with ownership of 3D data? Quite a bit, as it turns out! To make the material more relatable to readers who might work in the production, analysis, or curation of 3D data, four case studies are presented that cover representative situations. These include photogrammetry of a natural history object within an academic institution, photogrammetry of cultural object by a nonprofit in collaboration with a tribal authority, submission of CT scan data to an institutional repository by an individual outside of the institution, and publication of a virtual architectural model created from many sources. These analyzed examples should allow readers to recognize legal issues within their own work more easily. The chapter also covers the basics of contracts and licensing, as well as fair use.

Accessing 3D Data

The issue of access and discoverability is not simply a matter of permissions and availability. Chapter 6, "Accessing 3D Data," addresses concerns widely identified by the community regarding identification, location, retrieval, and use of 3D data. Different audiences have different needs related to effectively finding and using these data, yet the character of the data (capture methodology, pipelines) also influences data discovery and user interactions. Thus, the topics of access, audience, and 3D methodology are all presented together due to their interdependence. Only by developing standards in the context of these interacting considerations can we move forward with plans for long-term preservation and reuse of 3D data. This chapter elaborates on access-related issues conveyed in chapter 3, "Management and Storage of 3D Data," and chapter 2, "Best Practices for 3D Data Preservation," and it illustrates the importance of appropriate rights and ownership consideration and documentation, which impact access.

Glossary

In Forum 1, we recognized immediately the need for the disambiguation of terms. The glossary committee brought together members from all of the working groups to pull together key terms and acronyms that emerged in the forum events, virtual CS3DP meetings, and working group discussions. The glossary committee added to and refined this list of terms as the book developed. The intention is to define terms used in the book and to be useful to those from all backgrounds.

Notes

1. Community Standards for 3D Data Preservation (CS3DP), "CS3DP 2017 Community Survey," accessed January 10, 2020, https://osf.io/tcn6h/.
2. University of Iowa, "Community Standards for 3D Preservation (CS3DP)," accessed May 4, 2020, https://ir.uiowa.edu/cs3dp/.
3. E.g., Ivan Sutherland, "Sketchpad: A Man-Machine Graphical Communication System," in *AFIPS '63 (Spring): Proceedings of the May 21–23, 1963, Spring Joint Computer Conference* (New York: Association for Computing Machinery, 1963), 329–46, https://doi.org/10.1145/1461551.1461591.
4. Lawrence Roberts, "Machine Perception of Three-Dimensional Solids" (PhD diss., Massachusetts Institute of Technology, 1963).
5. Chris Wylie et al., "Half-Tone Perspective Drawings by Computer," in *AFIPS '67 (Fall): Proceedings of the November 14–16, 1967, Fall Joint Computer Conference* (New York: Association for Computing Machinery, 1967), 49–58, https://doi.org/10.1145/1465611.1465619.
6. Henri Gouraud, "Computer Display of Curved Surfaces," technical report, Department of Computer Science, University of Utah, 1971, https://collections.lib.utah.edu/ark:/87278/s69k4vts; Andrew Utterson, "A Computer Animated Hand," Library of Congress, 2011, https://www.loc.gov/static/programs/national-film-preservation-board/documents/computer_hand2.pdf; Robert McDermott, "Robert Remembers: The VW Bug," *Utah Teapot*, Fall 2003.
7. Marc Levoy et al., "The Digital Michelangelo Project: 3D Scanning of Large Statues," in *SIGGRAPH '00: Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques* (New York: ACM Press, 2000), 131–44, https://doi.org/10.1145/344779.344849.
8. Marc Levoy and Hector Garcia-Molina, *Creating Digital Archives of 3D Artworks*, white paper (Stanford, CA: Stanford University, 1999), https://graphics.stanford.edu/projects/dli/white-paper/dli.html.
9. Levoy and Garcia-Molina, *Creating Digital Archives*.

10. Stanford University, "The Digital Michelangelo Project Archive," accessed April 15, 2020, https://accademia.stanford.edu/dmich-archive/.
11. University of Utah School of Computing, "Wireframe Model of Sutherland's Volkswagen," 1972, Computer History Museum, https://www.computerhistory.org/revolution/computer-graphics-music-and-art/15/206/558; "Vehicles," 3D Café, http://www.3dcafe.com/vehicles.htm, archived at the Internet Archive, https://web.archive.org/web/19971210195553/http://www.3dcafe.com/models/yellobug.exe.
12. Ivan Sutherland, e-mail message to author, October 16, 2019.
13. Levoy and Garcia-Molina, *Creating Digital Archives*.
14. CyArk home page, accessed January 15, 2020, https://www.cyark.org/.
15. National Endowment for the Humanities, "Advanced Challenges in Theory and Practice in 3D Modeling of Cultural Heritage Sites," accessed January 15, 2020, https://www.neh.gov/divisions/odh/institutes/advanced-challenges-in-theory-and-practice-in-3d-modeling-cultural-heritage.
16. Marcia K. McNutt et al., "Transparency in Authors' Contributions and Responsibilities to Promote Integrity in Scientific Publication," *Proceedings of the National Academy of Sciences* 115, no. 11 (March 2018): 2557–60, https://doi.org/10.1073/pnas.1715374115.

Bibliography

Community Standards for 3D Data Preservation (CS3DP). "CS3DP 2017 Community Survey." Accessed January 10, 2020. https://osf.io/tcn6h/.

CyArk home page. Accessed January 15, 2020. https://www.cyark.org/.

Gouraud, Henri. "Computer Display of Curved Surfaces." Technical report. Department of Computer Science, University of Utah, 1971. https://collections.lib.utah.edu/ark:/87278/s69k4vts.

Levoy, Marc, and Hector Garcia-Molina. *Creating Digital Archives of 3D Artworks*. White paper. Stanford, CA: Stanford University, 1999. https://graphics.stanford.edu/projects/dli/white-paper/dli.html.

Levoy, Marc, Kari Pulli, Brian Curless, Szymon Rusinkiewicz, David Koller, Lucas Pereira, Matt Ginzton, et al. "The Digital Michelangelo Project: 3D Scanning of Large Statues." In *SIGGRAPH '00: Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*, 131–44. New York: ACM Press, 2000. https://doi.org/10.1145/344779.344849.

McDermott, Robert. "Robert Remembers: The VW Bug," *Utah Teapot*, Fall 2003.

McNutt, Marcia K., Monica Bradford, Jeffrey M. Drazen, Brooks Hanson, Bob Howard, Kathleen Hall Jamieson, Veronique Kiermer, et al. "Transparency in Authors' Contributions and Responsibilities to Promote Integrity in Scientific Publication." *Proceedings of the National Academy of Sciences* 115, no. 11 (March 2018): 2557–60. https:/doi.org/10.1073/pnas.1715374115.

National Endowment for the Humanities. "Advanced Challenges in Theory and Practice in 3D Modeling of Cultural Heritage Sites." Accessed January 15, 2020. https://www.neh.gov/divisions/odh/institutes/advanced-challenges-in-theory-and-practice-in-3d-modeling-cultural-heritage.

Roberts, Lawrence. "Machine Perception of Three-Dimensional Solids." PhD diss., Massachusetts Institute of Technology, 1963.

Stanford University. "The Digital Michelangelo Project Archive." Accessed April 15, 2020. https://accademia.stanford.edu/dmich-archive/.

Sutherland, Ivan. "Sketchpad: A Man-Machine Graphical Communication System." In *AFIPS '63 (Spring): Proceedings of the May 21–23, 1963, Spring Joint Computer Conference*, 329–46. New York: Association for Computing Machinery, 1963. https://doi.org/10.1145/1461551.1461591.

University of Iowa. "Community Standards for 3D Preservation (CS3DP)." Accessed May 4, 2020. https://ir.uiowa.edu/cs3dp/.

University of Utah School of Computing. "Wireframe Model of Sutherland's Volkswagen." 1972. Computer History Museum, https://www.computerhistory.org/revolution/computer-graphics-music-and-art/15/206/558.

Utterson, Andrew. "A Computer Animated Hand." Library of Congress, 2011. https://www.loc.gov/programs/static/national-film-preservation-board/documents/computer_hand2.pdf.

Wylie, Chris, Gordon Romney, David Evans, and Alan Erdahl. "Half-Tone Perspective Drawings by Computer." In *AFIPS '67 (Fall): Proceedings of the November 14–16, 1967, Fall Joint Computer Conference*, 49–58. New York: Association for Computing Machinery, 1967. https://doi.org/10.1145/1465611.1465619.

Chapter 2

Best Practices for 3D Data Preservation

Kristina Golubiewski-Davis, Jessica Maisano, Marcia McIntosh, Jennifer Moore, Kieron Niven, Will Rourk, and Rebecca Snyder

ABSTRACT

The CS3DP Preservation Best Practices working group has focused on creating a framework that can be applied across multiple 3D-data-generating modalities and tiers of implementation, from national to local governmental, educational, and private entities. The primary focus in this chapter is on identifying ***preservation intervention points*** *(PIPs) within project* ***workflows****, that is, moments to stop and ask, "What files should I save?" and "What information should be recorded at this point?" The answers to these questions are project-dependent and related to the purpose of generating the data, target audiences,* ***preservation*** *needs, and creation modality. However, some general guidelines are provided for new and established practitioners to consider. This chapter breaks down the PIPs into four main areas: planning; collection and creation; processing; and curation and long-term access. At each stage, we recommend levels of implementation to accommodate different levels of infrastructure, funding, and audience requirements. The chapter concludes with a discussion of* ***Good/Better/Best*** *recommendations and is supported by an extensive appendix comprising case*

studies illustrating PIPs and the Good/Better/Best recommendations for a range of 3D-data-generating modalities.

Introduction

As a practice, **digital preservation** has deep connections with disciplines such as the library and archival sciences. Libraries and **archives**, with their centuries-long history of experience, have become so ingrained within the consciousness of the people they serve that their inherent value is implied in their very names. Say "library," especially within a museum or academic context, and no one would question why it is important or what purpose it serves. The challenge for digital preservation, as a relatively new field, is to continually demonstrate the value it provides to the communities it serves. It can be difficult to make a case to an institution's administration as to why additional funds, staffing, and other resources must be given to ensure long-term viability, discoverability, and **access** to digital resources. Often it is assumed that, because paper documentation is "so easy" to maintain over time, digital data should not be any different. Paper preservation seems effortless to a non-conservator simply because of the decades of work and research that have provided a strong foundation for it to become common practice. Digital resources are essential to the mission of contemporary organizations, society, and knowledge creators. Over time—*and many lost files later*—the institutional value of digital preservation will likewise become an assumption. Best practices for digital preservation provide an important foundation toward supporting this goal.

Existing Standards

Digital preservation typically is divided into three major foci—management, technology, and content—often with an accompanying diagram of a three-legged stool or a Venn diagram "triad" to illustrate how all preservation activities are interlinked and interdependent.[1] Indeed, the intention to provide access to and preserve digital materials long into the future is impossible without institutional will, funding, clear policies, and review cycles (management); a well-designed and robust repository that is tailored to the targeted types of data and designated communities (technology); and well-defined selection criteria, **metadata** standards, and a deep understanding of formats (content). Because digital preservation touches nearly every aspect of an organization, the task of converting that intent into a tangible plan of action can be daunting.

A group of archivists responsible for astronomical data recognized the need for preservation content management standards and formed the Consultative Committee for Space Data Systems. Their reference model for the Open Archival Information

System (OAIS) has become an International Organization for Standardization (ISO) standard and the basis for thinking about the management and preservation of digital materials among digital preservation practitioners.[2] The companion, *Audit and Certification of Trustworthy Digital Repositories* (often referred to as TRAC), was designed to help institutions implement the particularly dense OAIS reference model and assess compliance through a checklist.[3] The standard divides the OAIS model into three sections: Organizational Infrastructure; Digital Object Management; and Technical Infrastructure and Security Risk Management. The checklists require an institution to consider multiple factors when designing a repository or considering adding new functionality or support for new data types. More than simply focusing on software and storage, both OAIS and TRAC place great emphasis on clearly established policies, documentation of user needs, and long-term maintenance requirements. It is a misconception that digital preservation is primarily about technology and formats. While these are clearly important, sustainable preservation cannot occur without a robust organizational investment and long-term commitment. See chapter 3, "Management and Storage of 3D Data," for additional discussion.

PREservation Metadata: Implementation Strategies (PREMIS) is a working group focused on creating common metadata standards for preservation activities. The *PREMIS Data Dictionary for Preservation Metadata*, currently in version 3.0, is designed to assist digital repositories in capturing vital information relating to digital preservation concerns.[4] The best practices detailed in this chapter are compatible with the concepts detailed in the **data model** and provide a means for implementing PREMIS for 3D data. See chapter 3, "Management and Storage of 3D Data," for a more in-depth discussion. Ideally, a **digital repository** would combine the OAIS model with PREMIS-compatible metadata standards.

Since 2016 the concept of FAIR data (data that are findable, accessible, interoperable, and reusable) has had a growing presence within digital preservation with the aim of highlighting the value and reuse potential of research data.[5] As with the three foci highlighted above, the FAIR principles are built around the three similar core areas (entities) of data, metadata, and management infrastructure.[6] While the FAIR principles are not a rigid, precise standard, they provide a backdrop against which projects, workflows, tools, and systems can be assessed in order to ensure that data and metadata are accurately and consistently created and in ways that are discoverable, accessible, and persistent. The FAIR principles relate directly to the preservation intervention points (PIPs) framework discussed below, the latter framing these considerations within a data-creation-to-deposition workflow. To some extent, implementation of many of the FAIR principles is the responsibility of the archive or repository to which the data may ultimately be submitted. However, an awareness of these key considerations by data creators early within a project can only help to ensure that datasets—and the institutions in which they are deposited—are able to comply with these principles.

The goal of digital preservation is challenging enough for standard archival and library materials, which are well studied and understood as content types and formats. It becomes much more nebulous when trying to manage and preserve emerging digital forms, covering numerous approaches and producing a variety of potentially unstandardized open and proprietary data formats. This is particularly true for 3D data, from their raw stage to their fully processed final form. As the name implies, the CS3DP group is dedicated to creating community standards for the preservation of 3D data. This chapter in particular considers 3D as a new content type that must be brought into existing OAIS best practices. These criteria and considerations are well established for other content types; the challenge is to modify and customize them for the unique properties and workflows associated with the major 3D data collection and processing methodologies, while remaining flexible enough to accommodate new 3D data collection modalities. Best practice considerations include assessment concerns by collection method, data types, retention schedule guidelines, evaluation criteria, discipline-specific standards, and the varied needs of designated communities. There can be no one solution that accommodates all institutions, all budgets, all data types, and all community needs. Questions like "How often should I review file format viability?" "Should I maintain all the raw data long-term?" "How would my designated community expect to discover and receive content?" "What derived data and versions should be preserved?" will be answered by each institution according to its capabilities and requirements.

This chapter seeks to help institutions and individuals ask the vital questions and provides tiered solutions based on need and resources, using a PIPs approach that details key assessment points throughout the life cycle of a 3D project, and poses key questions to consider. An institution's priorities, capabilities, selection criteria, designated community requirements, and available funding will all inform how these questions are answered and what recommended actions are taken. Baseline recommendations for preservation are offered in a Good/Better/Best format at the end of this chapter (tables 2.4a, 2.4b, and 2.4c), with the intent that the PIPs framework will guide the implementation of preservation work. It is recommended that, at a minimum, data be saved in open-standard, nonproprietary, and human-readable file formats (e.g., ASCII text or XML-based formats), as this removes a dependency on specific software packages and allows greater flexibility for future access and migration. Throughout this chapter, the project lead is referred to as the decision maker. This choice was intentional to reflect the decisions that project leads often make early in the project, although this decision-making process could also be undertaken by a team or an institution.

Preservation Intervention Points

The basis for these best practices recommendations is the framework of preservation intervention points (PIPs). This framework comes from the Archaeology Data Service

VENUS Preservation Handbook, expanded upon in the Archaeology Data Service and Digital Antiquity Guides to Good Practice.[7] The basic premise of the PIPs framework is to identify and assess critical decision points in the data creation process and record those points within the context of their long-term implications for data preservation and reuse (see figure 2.1). Given the variety of 3D data acquisition methods discussed in this chapter (and in this volume as a whole), the following sections identify the stages where PIPs are most likely to occur in any 3D data project, with specific examples of PIPs in the 3D data acquisition workflow of several case studies (see appendix).

The following stages are considered most relevant for 3D data collection and/or creation workflows:

- planning
- collection or creation
- processing (post-acquisition data manipulation)
- curation and long-term access

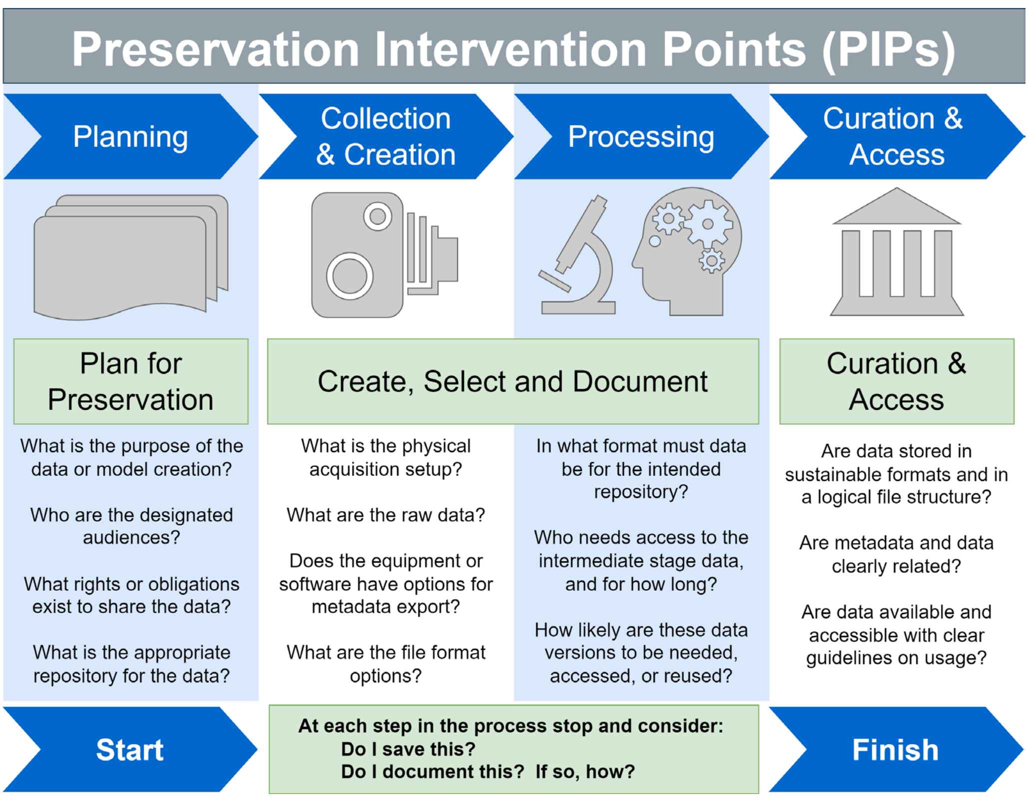


Figure 2.1
Basic preservation intervention points (PIPs) within a digital project. Image created in diagrams.net.

Ideally, when considering preservation of specific 3D data, the data collection and/or creation workflow will be mapped prior to beginning the process. Greater documentation of a specific project will help to identify which aspects of the workflow are worth preserving. 3D data collection and/or creation include a wide variety of technologies and processes. The framework outlined in this chapter is intended to be flexible, allowing the project lead to implement the stages that are relevant to their project. Not all stages will be necessary for all projects. For example, a sources-based modeling project may not include 3D data collection, but may include a phase of collecting other types of **source material**. The sections below describe the generalities of the PIPs framework and some specific examples.

To illustrate the PIPs framework in the context of an entire project, case studies are provided in the appendix that document specific workflows for projects that use sources-based, **photogrammetry**, X-ray CT, large-scale **laser scanning**, and **structured light** scanning acquisition techniques. Each case study identifies the PIPs stages within that workflow and recommended actions for preservation. These examples also serve to illustrate how the PIPs framework can be adapted to fit specific institutional priorities and needs, as each project deals with these differently.

Each of the 3D project stages below indicates a point at which the project lead identifies and records information for long-term preservation. For each preservation stage, we have identified a series of questions for project leads to consider and the preservation implications of the answers to those questions. Within the context of a project, PIPs occur when the project has reached a predetermined milestone or when the data or data format has been altered in an irreversible way. These are the points at which the process and any appropriate metadata should be recorded, together with an assessment of data format in light of long-term preservation and access (i.e., avoiding pathways that ultimately lock data into proprietary or inaccessible formats). The goal of the PIPs framework is to create a plan that takes preservation into account while being flexible enough to allow pivoting within the project based on unforeseen variables. An outline of known PIPs considerations and documentation methods created at the beginning of the project will benefit the process and should be available to all members of the project group.

As project leads consider the PIPs and questions detailed below, they will need to make context-specific decisions that meet the needs of their project, their institution, and their designated audiences. As a guideline, at the end of this chapter there is a set of Good/Better/Best recommendations for implementation; see tables 2.4a, 2.4b, and 2.4c.

Planning

The planning stage is the most important stage at which to consider preservation because this is when the scope of the project and requirements for long-term preservation are decided. The purpose of the project, expected reuse, and needs for dissemination and preservation will affect how information is created and preserved during the workflow. It is strongly recommended that the 3D data creator or collector consider all project stages early in the process, including determination of an appropriate long-term storage solution for the project data and consultation with repository representatives, if relevant. While some of the information can be recreated at a later point in time, some information can be captured only during the data creation stage (see the "Create" section in chapter 4, "Metadata Requirements for 3D Data"). As the preservation and reuse of the data can hinge on specific data format choices and whether or not metadata and **paradata** (see chapter 6, "Accessing 3D Data") were recorded at the appropriate time, the 3D data collector or creator should carefully consider how the questions in table 2.1 affect a specific project and workflow. Educational institutions often offer support for project planning and 3D data preservation, usually via institutional repositories, libraries, humanities centers, or technical support in various departments.

The planning stage also includes time spent preparing for 3D data collection and/or creation. Preparing the location, creating collection procedures, gathering equipment, and testing the conditions to determine the best collection protocol are all done prior to acquiring the first 3D data collection point. The case study "Large-Scale Laser Scanning" in the appendix provides an example that includes complex planning factors, multiple experts, and limited access to the site prior to its destruction. Any decisions or factors that affect the collection process should be documented in a consistent manner. Careful attention should be paid to file organization techniques, and unique identifiers should be used for projects that include 3D data creation for multiple objects. Documentation of both the process and the reasons for collection or data optimization methods can inform the choices and priorities determined during those stages if access to the material is restricted by time or other unknown factors.

Purpose of the Data Collection and/or Creation

The primary focus of 3D data collection is to provide enhanced dimensional information for the documentation of an object beyond that which can be acquired through 2D capture. This information may include both external and internal 3D properties of the object. For cultural heritage subjects, the geometry and texture are vital for the historical record of physical places and things. Data associated with specific buildings or sites may have specific points associated with real-world coordinates. The original purpose of the data collection or creation will impact the preparation of derivatives of the 3D data. For example, projects intended to provide accurate surface measurements will require point clouds with a high density of points, or surface models with a higher polygon count and recorded unit size for reference, whereas those intended for 3D printable models will comprise non-manifold closed surfaces and will not require as high a polygon count as detailed research data. Likewise, it is more important to consider accurate color representation for models used to increase public engagement through online interaction than for data that will be used to take measurements. These early considerations can help guide a practitioner in determining which data types and formats to preserve to meet the needs of the data or model use, and also the potential reuse of the data. For example, if the data were originally created for a 3D printing project, but the project was created through 3D scanning that yielded high-quality original scans that could be useful for measurement projects, then the preservation of detailed metadata, descriptions of the scanning process, and files from intermediate steps would increase the likelihood of reuse of the project data and the ability to check for scientific **accuracy** (see chapter 6, "Accessing 3D Data"). The case studies in the appendix include data whose purposes include educational outreach ("Photogrammetry"), cultural heritage preservation by use of 3D documentation ("Large-Scale Laser Scanning" and "Multimodal"), visualization of data for academic narratives ("Sources-based"), scientific analysis ("High-**Resolution** X-ray CT"), and creation of a 3D assemblage of objects for academic study ("Structured Light Scanning"). Each case study offers an explanation for how and why each institution chose to record information based on the needs of the project and the capabilities of the institution.

Detailed metadata and paradata should be collected for any project (see chapter 4, "Metadata Requirements for 3D Data"). The decision to collect and preserve higher quality data than what are required for a specific project should be made based on the ability of the research team to collect those data and the ability of the institution to properly preserve those data. Tables 2.4a, 2.4b, and 2.4c offer baseline recommendations for which files to preserve based on the purpose of the project. Ultimately, the 3D data collected and preserved for a project will be a balance between the best practice recommendations and individual project needs.

Intended Audiences

When creating 3D data, the acts of documenting the process and preserving the data are recommended but rarely done unless the data will be critically evaluated (e.g., by scientific peers) or reused. At a minimum, the audience consists of the individual doing the work and the client for whom the work is done. For these reasons, careful documentation and proper preservation will benefit continued work involving the 3D data. In chapter 6, "Accessing 3D Data," the authors define six categories of audiences: scholars and researchers; educators; students; museums, public outreach, and nongovernmental organizations (NGOs); professionals; and general user/personal interest. Each of these groups will have different needs regarding the level of documentation and access to raw data or intermediary files. For example, scholars and researchers typically require a high level of documentation, in-depth metadata, and access to the raw files in order to verify the accuracy of the data or model and incorporate relevant data into their own projects. Public outreach groups will require access to the final project products in open file formats that are easy to reuse across a variety of platforms. Considering the needs of the intended audience during the planning stage will identify important PIPs stages for recording information and determine the level of preservation and access appropriate for the project.

Repository Selection

Long-term preservation of 3D data is best achieved by depositing the project into a repository, preferably a trusted digital repository.[8] In some cases, this type of preservation may not be possible due to limited resources or restrictions on the data, in which case an alternative may include making the 3D data publicly accessible via an open-access repository or online database and finding medium-term solutions for storing the data (repository and storage examples are provided in tables 2.4a, 2.4b, and 2.4c). The purpose of the project and funding requirements will help to determine the appropriateness of long-term preservation and the breadth and depth of the data preserved. Frequently, grant-funded projects have an obligation to publish and preserve data in a publicly accessible manner. During the planning stage, it is important to identify what preservation strategy will be applied to the data.

Long-term preservation often requires the identification of a repository in which to deposit the data. As of this publication, there is no universally accepted file format, agreed-upon set of metadata, or standardized input format for 3D metadata. Identifying the repository for a project early on allows the project lead to understand the restrictions and implications of using different repositories. For example, some repositories allow files to be uploaded in a nested folder structure while others allow only a single folder for all of the files. Repositories might have a file type preference for the final ingest or specify a format preference and structure for metadata collection (see chapter 3, "Management and Storage of 3D Data," for more information). Understanding the requirements and limitations of the intended repository early in the process can expedite the collection of metadata and organization and formatting of files by matching the collection strategy to the needs of the preservation strategy. In addition to ensuring that data to be preserved meet the requirements of the selected repository, it is important to verify that the repository itself is capable of storing and preserving data in the long term. Should the chosen repository be unable or unwilling to preserve the volume of 3D data, metadata, and paradata necessary to fully encapsulate the project, either another repository should be considered or a secondary preservation solution should be identified for the additional data. The OAIS model describes a standardized framework in which repositories can operate, and, building from this model, a number of standards, assessment, and certification processes have been created in order to allow institutions to highlight themselves as trustworthy data repositories (e.g., the CoreTrustSeal).[9] Repository certification is also discussed in the "Certification" section of chapter 3, "Management and Storage of 3D Data." 3D data creators can prepare to discuss depositing their data with a repository staff person by providing answers to several questions, including the following:

- What are the origin and the context of the data?
- What are the initial and target forms and formats of the data?
- What is the expected lifespan of the data?
- How might the data be reused or repurposed?
- How large is the dataset, and what is its rate of growth?
- Who are the potential audiences for the data?
- Who owns the data?
- Does the dataset include any sensitive information?
- What publications or discoveries have resulted from the data?[10]

Table 2.1 provides some examples and potential answers to fundamental questions that should be addressed at the initial stage of planning for 3D data preservation within the data workflow.

TABLE 2.1

Planning stage considerations

Question/Answer	Implications
What is the purpose of the data or model creation?	The purpose of the project will determine the ways in which the data should be preserved and made accessible. See chapter 6, "Accessing 3D Data," for an in-depth discussion.
Visualization	The final version of the project (model, animation, etc.) will be the most important set of files to preserve. Self-contained formats are more likely to maintain integrity. When multiple files are required to reproduce the visualization, file structure and relationships between files should be clearly organized.
Data analysis	The analytical procedure is crucial to preserve. Original 3D data and information about how those data were edited/analyzed to obtain the final results will enable others to understand and verify those results.
3D printing	Files intended for 3D printing should be relatively small and in .stl or .obj format. While raw data may be useful in the long term, it is important to provide them in a format that is easily migrated to current technology. Accompanying metadata should include print settings appropriate for the model or settings and equipment used for specific printouts when the model is used in scholarly publications. Any editing done to make the model printable (e.g., hole filling, decimation, smoothing) should be clearly documented.
Historical or cultural record	3D data documentation inherently provides dimension and may provide texture or color information of the subject. For cultural heritage and historical subjects, these data provide crucial documentation for the historical record.

Question/Answer	**Implications**
Who are the designated audiences?	Different designated audiences will have different needs regarding documentation and metadata that will also affect which stages and formats of the data should be preserved. See chapter 4, "Metadata Requirements for 3D Data," for an in-depth discussion of metadata documentation and chapter 6, "Accessing 3D Data," for definitions and discussion of audience types. The following list provides examples of academic audience user groups for whom 3D data can help support research and pedagogy.
Scholars and researchers	Documentation should include as much information as needed to replicate the study, including both technical and field-specific metadata. The 3D data should be made available minimally to peer reviewers. Persistent identifiers (e.g., DOI—digital object identifier [https://www.doi.org/] and ORCID—Open Researcher and Contributor ID [https://orcid.org/]) will increase data reuse and citation after publication.
Educators	The 3D data should be easily accessible, provided in a usable format, and documented to the level of intended use. Instructional supplements should also be included.
Students	The 3D data should be easily accessible and provided in a usable format.
Museums, public outreach, and nongovernmental organizations (NGOs)	3D data ownership and reuse policies must be addressed and proper credit given for data or model generation and funding source.
Professionals	Depending on the professional use, considerations for other groups may also apply here. When applicable, the level of documentation should follow the recommendations of any relevant government regulations for that field as well as any policies internal to the institution; these may include limitations on long-term preservation and data sharing and use.
General user/personal interest	Guidelines should be provided for attribution of the 3D data and restrictions on their use (see chapter 5, "Copyright and Legal Issues Surrounding 3D Data").
What rights or obligations exist to share the data?	Obligations to share data may come from grants or institutional policies. The rights of the data collector or creator to share 3D data will depend on the purpose and nature of the project. See chapter 5, "Copyright and Legal Issues Surrounding 3D Data," for an in-depth discussion of rights.
Grant-funded obligations	Publicly funded grants increasingly include obligations for openly sharing data where appropriate. These obligations should be planned for when determining the long-term preservation and access needs of the project.

Question/Answer	Implications
Cultural sensitivity access	In some cases, the objects being transformed into 3D data may be culturally sensitive. This must be taken into account when determining the appropriate way to share their data.
Restrictions by object owners	Objects that are not owned by the project lead or their institution may have restrictions on the use of their 3D digital representations. See chapter 5, "Copyright and Legal Issues Surrounding 3D Data," for more information.
Georeferenced data	References to the geolocation of object-based data may be culturally sensitive even if the data themselves are not. Sites that are thought to contain artifacts of significant number or cultural value or sites that are underwater may be susceptible to looting. Sites with endangered species may be subject to over-collection. These concerns should be taken into account before sharing geolocation of object-based data.
What is the appropriate repository for the data?	At the time of this writing, there is no standard for how 3D data are accepted into a repository, with most repositories having their own workflow preferences. Identifying a repository early in the project allows the project lead to incorporate repository requirements (e.g., data formats and metadata) into the workflow. See chapter 3, "Management and Storage of 3D Data," for an in-depth discussion.
Institutional repository	Institutional repositories may not be familiar with 3D data. Be sure to consult the repository contact to understand possible limitations of the repository and discuss preservation strategies.
Noninstitutional repository	Each repository will have a different cost structure and ingest mechanism. Identify these early for incorporation into the project workflow and funding requests.
No repository	Understand the ramifications of not putting the data into a repository. Identify other options for disseminating the data if long-term preservation is not possible due to funding limitations. As with repositories, identify the file formats and associated information needed to use those resources early in the project.

Collection and/or Creation

The collection and/or creation stage refers to the point at which the raw data are generated. The duration of this stage and the number and volume of files generated depend on the method used to create the 3D data. It is important to document instrument settings, environmental settings, and acquisition protocols that affect this process. When possible, preserve any file outputs that the equipment software might generate (e.g., **README files** that record equipment settings). Also, when possible, collect the highest resolution data with the most complete metadata (see chapter 4, "Metadata

Requirements for 3D Data," for specific guidelines) and paradata available for the project—extra data are better than missing data. This is particularly true if the real-world object being captured by the 3D data will not be easily accessible for additional analysis. In addition to metadata, the collection and creation stage should include documentation of both the standard methods followed and the ability to note when and for what reason deviations occurred. If intentions behind particular data collection methods were developed during the planning stage, these can be used to help justify deviations from the standard method. Early discussion of the purpose of collecting the data, the level of detail in the metadata and paradata for the project, and the capabilities of the repository to preserve those data long-term are crucial to ensuring the project meets the minimum best practices recommended in this volume. Table 2.2 provides some examples of fundamental questions that should be considered at the beginning of the collection and/or creation stage for 3D data preservation within the data workflow.

TABLE 2.2

Collection and/or creation stage considerations

Question	Implications
What are the raw data?	The raw data should be recorded and saved. Refer back to the intended audience and purpose of collecting or creating the 3D data to determine if they should also be preserved long-term. Projects that are of an academic nature or result in the destruction of the original object, such as archaeological digs, should always preserve the raw data.
Does the equipment or software have options for metadata export (embedded or sidecar)?	Some software and imaging equipment have options for exporting a sidecar document (such as an associated README text file) that describes the settings of the equipment and/or records a series of settings that the software used. In some cases, this information can also be embedded into the file itself. This information should be retained and kept with the 3D data.
What are the file format options?	During 3D data creation and collection stages, it is normal that raw data are saved in their native file format. In cases where that format is proprietary, a nonproprietary format should be identified and preserved as well, particularly if continued access to software cannot be guaranteed. Where possible, both proprietary and nonproprietary formats should be preserved long-term for research audiences. Some proprietary raw data files are not just a single file but a main project file and associated dependencies; this file structure must be retained for preservation and archiving. Refer to the section "Good/Better/Best Recommendations for Implementation" for preservation strategies for proprietary and nonproprietary file assets depending on archival institution resources.

Question	Implications
What is the physical acquisition setup?	When using a 3D data acquisition method, the physical environment, setup, and equipment affect the quality of the resulting data. Document this information.
At what steps in the process should one stop and consider: • Do I save this? • Do I document this? If so, how?	Each 3D data acquisition method will have a unique process. Determining the steps at which files are saved or processing is documented will create a series of PIPs for that particular process. This information can then be used to create a protocol that includes the appropriate documentation points and methods. All documentation should be preserved.

Processing (Post-acquisition Data Manipulation)

After the data acquisition phase, many 3D imaging modalities include a stage of data processing, which involves any modifications to the data between their acquisition and the final product. For example, raw data for photogrammetry comprise the original photographic images. These images then go through a processing stage where geometry and texture are created, cleaned, and exported as a final object. Structured light and laser scanners create one or more sets of data that are then aligned, optimized, and merged. Data intended for 3D printing may require editing (e.g., filling of holes, decimation, etc.). All of these modifications after the raw data acquisition should be carefully documented (see the case study "Structured Light Scanning" in the appendix for an example). When possible, protocols and policies should be created that can be shared to assist the audience in understanding what manipulation was done to the raw data to create the final 3D data or model. In the case of bespoke data optimization, the changes that the data have undergone should be described and documented to the best of the ability of the individual making those changes. It is critical to record data processing or correction parameters (as exemplified in the case study "High-Resolution X-ray CT" in the appendix) so that the final result may be replicated or understood for research and scholarly applications. As with the collection and creation stage, a standard method of processing should be documented along with any deviations from that method. Table 2.3 provides some examples of fundamental questions that should be addressed at the initial processing stage of planning for 3D data preservation within the data workflow.

TABLE 2.3
Processing stage considerations

Question	Implications
At what steps in the process should one stop and consider: • Do I save this? • Do I document this? If so, how?	All modifications to the raw data should be documented and that documentation preserved, but some PIPs project files need be kept only as working files and may not be included in the final project. Refer back to the intended audience and purpose of the 3D data creation to determine which working files should be preserved long-term (see the section "File States and Submission Packages" in chapter 3, "Management and Storage of 3D Data").
In what format must the data be for the intended repository?	Understanding the structure and requirements of the repository and how these relate to the format and documentation of the project will expedite the translation of 3D data and documentation for inclusion in the repository. To determine what these requirements might be, contact the intended repository (see chapter 3, "Management and Storage of 3D Data," for guidance on choosing a repository).
Who needs access to the intermediate-stage data, and for how long?	Not all intermediate-stage data need be preserved, but there may be value in retaining these data for future work. Determining the audience and need for intermediate-stage data will help define an appropriate storage location and who should have access.
How likely are these data versions to be needed, accessed, or reused?	If intermediate data are likely to be needed for reuse of the final 3D data product, they should be preserved in the final data package.

Curation and Long-Term Access (SIP, AIP, and DIP)

SIP (**Submission Information Package**), AIP (**Archival Information Package**), and DIP (**Dissemination Information Package**) are conceptual terms used by the OAIS reference model (ISO 14721) to refer to packages of information or data that are submitted (SIP), archived (AIP), and disseminated (DIP) by an archival body or repository.[11] While the requirements and internal workings of repositories vary dramatically depending on location and scope, the basic concept that data may be deposited in one form, stored and preserved by the repository in another, and then disseminated to users in yet another range of formats highlights the fact that there is neither a one-size-fits-all nor a permanent solution to digital preservation and dissemination. The inner workings of digital repositories are beyond the scope of this chapter (see chapter 6, "Accessing 3D Data," and chapter 3, "Management and Storage of 3D Data"). However, it is worth bearing in mind early on in any project involving

3D data destined to be preserved and accessed in the long term that the data created may be better preserved in a variety of formats, and in order for others to understand and use these data, sufficient documentation of their collection, creation, and processing should also be stored alongside them. As discussed above, the FAIR principles should inform many aspects of this stage, ensuring that data are made available, and in suitable formats for reuse, alongside consistent and well-structured metadata.

The SIP—the package prepared for submission to the intended repository—is the point at which the creation of the project archive ends and the long-term preservation process begins. *Prepared* is the key term here, as almost all archives and repositories have specific guidelines and requirements for the submission of data and may refuse to accept a dataset if it fails to meet these requirements. Generally speaking, these requirements fall into two main areas: the type of data being submitted (covering both the content and the format) and the documentation and metadata that accompany the 3D data. Some archives may accept only specific types of data, others only certain data formats and file types, typically with restrictions on both. A key to successfully preserving 3D data is to identify the repository and its requirements early on so that compliance can easily be built into the project workflow, ensuring data are archive-ready at the project's completion. A worst-case scenario is becoming aware of these requirements only as the project is drawing to a close, resulting in a lack of time and money—or physical or technical ability—to undertake the tasks required to make the 3D data suitable for deposit (e.g., format migration, data documentation, data cleaning).

Even if data are not to be deposited into a repository (not recommended), a final preparatory stage formalizing a "project archive" and taking into account some of the considerations highlighted in this chapter is a worthwhile activity and will help ensure that a dataset is coherent and complete. For many projects this will include a stage of data selection based on previously highlighted PIPs but should also include general tasks such as the removal of duplicate data or multiple versions of files, ensuring filenames and directory structures are consistent, and making sure that there are no barriers to long-term preservation and access (e.g., technical, legal, or ethical restrictions or considerations).

A detailed early consideration of the software used in a project and the resulting file formats employed for data storage may bring to light future issues regarding long-term access and sustainability. In general, most approaches to digital preservation favor open-standard, nonproprietary, and human-readable file formats (e.g., ASCII text or XML-based formats), as these remove a dependency on specific software packages and allow greater flexibility for future access and migration. The sustainability of digital formats is discussed in detail on the Library of Congress Digital Preservation web pages,[12] and general good practice guides such as the Digital Curation Centre's "Five Steps to Decide What Data to Keep" provide useful checklists.[13]

Once the 3D data have been successfully packaged and ingested by the repository, an AIP can be generated. The AIP is the full package of data and metadata that forms the basis of the archival dataset. In OAIS technical terms, the AIP should consist "of the Content Information and the associated Preservation Description Information (PDI), which is preserved within an OAIS"[14] In practice this consists of the original 3D data submitted for archiving, in a format suitable for long-term preservation, alongside the necessary information and metadata to aid preservation and to provide key details on **provenance**, context, **fixity**, and so on. While the types of documentation that can be included in the PDI vary depending on the data and project type (see the Digital Preservation Coalition's [DPC] wiki page on preservation description information for examples[15]), detailed descriptions of provenance and context (e.g., recording methodology, instrument setup, software processing, and intention) at various stages or PIPs can be included here to allow a greater understanding of the 3D data and their purpose and limitations. Data relationships, both between data within the package and to external elements, are key here, as the repository itself may need to repackage data for preservation purposes. A logical and structured directory system can make understanding a 3D dataset easier and make explicit the relationships between raw and processed elements. It is recommended, however, that such relationships also be documented elsewhere, and documentation is discussed in more detail in the following section.

Within a repository, additional metadata documenting any processing that the repository has undertaken on the 3D dataset will be created during ingest or migration. These metadata might include format migrations, corrections to data, renaming, or restructuring, and aims to provide a chain of custody for the dataset (see the sections "Distribute and Publish," "Access and Reuse," and "Archive" in chapter 4, "Metadata Requirements for 3D Data"). Again, even for 3D data not formally deposited within a repository, the recording of edits made to those data beyond the scope of the original project is important to allow derivatives and later versions of data to be traced back to their original forms.

In addition to the preservation data stored as the AIP, a DIP is created by the repository to allow access (at whatever level is appropriate) to the 3D data, usually as a download. In practice this may include all the submitted data, or it may be a discrete subset (e.g., processed data) or data derived from the AIP by the archive itself to provide different formats or different resolutions from those originally deposited. Such packages may also take into consideration other issues such as copyright restrictions or time-limited embargoes (see the section "Embargoes" in chapter 3, "Management and Storage of 3D Data," and chapter 5, "Copyright and Legal Issues Surrounding 3D Data"). While the DIP is largely specific to the repository, it is a worthwhile concept to consider when creating the SIP. If data contained within the SIP are simply different versions of the same data, then these relationships should be made clear to aid both the archive and any end user's access and reuse of the data in the correct context. Likewise, any restrictions should be clearly documented so that data can be withheld from or disseminated to the correct users.

Documentation

A major goal of the PIPs framework is to create points of documentation for the 3D data collection or creation workflow. This documentation should be complete enough that someone can understand the work done to the data at any point in the process. This includes both individuals looking at the archived data and individuals who might join in the middle of a project. The documentation structure should be compatible with institutional, project, and preservation needs. Certain repositories or institutions will request documentation in particular file formats or following specific guidelines. Additionally, different 3D imaging modalities may require different types of accompanying documentation. Without proper documentation, 3D data are useful as visualizations but lack the necessary information for reuse by scholars or researchers (see the section "Audience Categories" in chapter 6, "Accessing 3D Data"). The form that documentation takes varies based on project needs, but it may include

- embedded or sidecar metadata exported from 3D acquisition equipment software
- incorporation of metadata into file naming conventions
- information pertaining to file relationships needed to properly open the data
- associated spreadsheet data with documented information
- published written protocols for replication of method
- associated README text files that describe any of the above information

The methods of in-process and preservation-level documentation should be determined during the planning stage. The format of the data collection can be either structured or unstructured. Structured data, such as a spreadsheet, are appropriate for information that can be searched in a database, while unstructured data, such as a text file, are appropriate for recording protocols or other narrative information about the project. A common document found in US standards is a README.txt file, also known as a codebook. This is an associated (sidecar) file that describes the collection and/or creation of the 3D data, the meanings of any column headings in spreadsheet data, rights management information, preferred citation attribution, and explanations of file structure. It may also include protocols for processing the data, including the software used (including version). Examples of how each stage mentioned above can be expressed in a README file can be found in the case studies "High-Resolution X-ray CT" and "Structured Light Scanning" in the appendix. The project lead should identify the appropriate standards for their field and region of the world in regard to how to format and name the README document. An example of good documentation in practice for photogrammetry is the Cultural Heritage Imaging Digital Lab Notebook.[16]

Documentation should also take into consideration the appropriate file formats for the project, including 3D data formats, metadata storage, and associated process documentation. Proprietary formats are tied to specific software suites, are the closest version to the raw data, and contain metadata specific to the capture process. Proprietary formatted data usually reside in the project file for the **registration** of datasets and usually contain metadata from the instruments that collected the data. For reproducibility of the data, whether in the long or short term, the information bound up within a proprietary formatted project file is vital as it can contain instrument settings and readings from various sensors in the data acquisition instrument, such as altimeter, GPS, and inclinometer sensors. Also important to the project file are the settings that were used to acquire 3D data, such as the resolution of data at the time of capture: for example, the number of points captured per scan session or the **voxel** element dimensions in a CT scan. Many datasets comprise multiple datasets that have undergone some process of registration to combine them to create one comprehensive dataset. Proprietary formatted project files can contain the data regarding the **precision** with which the data were registered as well as the method of registration, thus ensuring fidelity and accuracy of the 3D data in the documentation of a subject. The ability to export project-related metadata cannot always be assured, but when it is possible to do so, they should be exported to a text-based format such as a .txt or XML document, especially if it is not permissible to archive the proprietary formatted project files. The case study "Large-Scale Laser Scanning" in the appendix is an example of a 3D data acquisition type that produces proprietary formatted project files such as those described here.

Nonproprietary formats offer a significant benefit in that they can be read without specific software suites, and therefore they are more accessible and it is easier to maintain file fidelity in an archive (see Good/Better/Best recommendations for implementation in tables 2.4a, 2.4b, and 2.4c for some examples). The 2020–2021 version of the Library of Congress recommended format standards now includes 3D data types.[17] These file types are more interoperable and reusable and thus in line with FAIR principles. Human-readable files formatted as ASCII text are easily parsed by both humans and computers for information regarding the documentation of a subject. Binary formatted data are an encoded form of the data requiring software to decode those data so they are human-usable. Binary formatted files are generally smaller in size and run much faster when processed or executed by a program. However, decoding binary data is not assured deep in the future, whereas human-readable text is more likely to be parsed by humans many years from now.

As noted earlier, long-term preservation of 3D data is achieved by depositing open-format, nonproprietary files in a repository. However, proprietary formats of raw data offer indispensable metadata embedded in the file that are often lost in the process of converting the data to nonproprietary formats. In addition to interoperable open-source formats, there is a compelling argument to keep the data in their raw format as part of the local documentation of the project, if not the final archive. Work done by the Software Preservation Network (SPN) to ensure continued access to previous software versions and advances in **emulation** software increase the probability that researchers will be able to retain access to these important data structures. As appropriate, 3D data should be retained in their original file structure and format.

Good/Better/Best Recommendations for Implementation

Guidelines

Good/Better/Best (GBB) recommendations are offered as guidelines to address

- the level of documentation to target based on audience and use
- recommendations for file formats in consideration of access and/or preservation
- PIPs at which to save and preserve the 3D data
- databases or repositories that target the needs of the intended audience

This format of GBB can be found throughout the volume. Any level of 3D data preservation and documentation is a benefit to the community, and the Good level presented in this volume should be considered a target minimum for best practices. The tiered nature of preservation strategies is also seen in other frameworks, such as the Federal Agencies Digital Guidelines Initiative (FADGI) Guidelines[18] and is not intended to place a value judgment on each level, but rather to create a structure that is easy to understand and remember. Not all 3D projects will require the highest level (Best) of preservation. Additionally, the level of preservation achievable will depend on the funding and resources allocated to the project. This is particularly true in regard to where the 3D data are preserved. At the time of this publication, many institutions do not have access to repositories with an infrastructure intended for 3D data. Project leads are encouraged to determine which tier (Good/Better/Best) best fits their needs for preservation and access based on their intended audience, with a goal to get as close to that tier as planning and resources allow.

- *Good:* Preservation necessary for general access and basic use of the data. This tier targets general users, personal interest, and pre-packaged educational use of the final data package.
- *Better:* Preservation necessary for customizable educational use of the data and outreach. This tier targets reuse of the data for a variety of purposes, including both the final data package and the original raw data.

- *Best:* Preservation necessary for cultural heritage preservation and documentation of scientific and scholarly work. This tier targets the ability to reproduce and verify studies using the available information about the 3D data. Should include any proprietary data registration project files.

TABLE 2.4A
Good: General user, personal interest, educators, and students

	Recommendation	**Examples**
Level of documentation	Context information about the data + Identification of creation method/creator + Basic metadata information	Associated README text document with the recommended information. Refer to chapter 4, "Metadata Requirements for 3D Data," for examples of tiered implementation in metadata.
Preserved file formats	A program-neutral format	.obj, .stl, .ply, .fbx, .tif, etc.
Data preservation points	Final project files	
Long-term access	Publicly accessible database	Thingiverse (https://www.thingiverse.com/), Sketchfab (https://sketchfab.com), etc.

TABLE 2.4B
Better: Museums, public outreach, nongovernmental organizations (NGOs), and professionals

	Recommendation	**Examples**
Level of documentation	Context information about the data + Identification of creation method/creator + More robust metadata	Associated README text document with the recommended information. + .csv files with specific metadata information. Refer to chapter 4, "Metadata Requirements for 3D Data," for examples of tiered implementation in metadata.
Preserved file formats	An open-source, program-neutral format that is indexed by PRONOM	.obj, .ply, .stl, .tif, etc.

	Recommendation	Examples
Data preservation points	Original raw data + Final project files	
Long-term access	Publicly accessible, institutional, or government repository	Dataverse (https://dataverse.org/), tDAR (the Digital Archaeological Record; https://www.tdar.org), etc.

TABLE 2.4C

Best: Scholars, researchers, and cultural heritage preservation

	Recommendation	Examples
Level of documentation	Context information about the data + Identification of creation method/creator + Robust metadata—including technical metadata outputs from software where available + Written explanation of the methodology (paradata)	Associated README text document with the recommended information. + .csv files with specific metadata information. + Technical reports associated with the software used (if available). Refer to chapter 4, “Metadata Requirements for 3D Data,” for examples of tiered implementation in metadata.
Preserved file formats	The original, proprietary format + An open-source, program-neutral format, indexed by PRONOM, that allows for structured and customized embedded metadata	Proprietary: .fls, .skp, .vue, .psx, .txrm, etc. Open-source: .dae, .x3d, etc.
Preservation points	Original raw data + Relevant intermediary steps that preserve the decision-making process + Final project files	
Long-term access	Institutional, government, or commercial repository with an infrastructure specifically for 3D data	MorphoSource (https://www.morphosource.org/), MorphoBank (https://morphobank.org/), Figshare (https://figshare.com/), Nature Scientific Data (https://www.nature.com/sdata/), etc.

Conclusion

As the name implies, the CS3DP group is dedicated to creating community standards for the preservation of 3D data. This chapter considers 3D as a new content type that must be brought into existing preservation best practices such as OAIS and TRAC. These criteria and considerations are well established for other content types; the challenge is to modify and customize them for the unique properties and workflows associated with the major 3D data collection and processing methodologies while remaining flexible enough to accommodate new 3D data collection modalities.

This chapter introduces the Preservation Intervention Points (PIPs) framework that details key assessment points throughout the life cycle of a 3D project and poses key questions to consider. An institution's priorities, capabilities, selection criteria, designated community requirements, and available funding will all inform how these questions are answered and what recommended actions are taken. The proposed framework can be applied across multiple 3D-data-generating modalities and tiers of implementation, from national to local governmental, educational, and private entities. As there can be no one solution that accommodates all institutions, all budgets, all data types, and all community needs, this chapter provides Good/Better/Best guidelines for baseline preservation recommendations that take into consideration these varying needs and available resources.

For those interested in seeing how the PIP framework and the Good/Better/Best guidelines work in a real-world context, please see the appendix for six case studies that provide in-depth descriptions of several 3D data acquisition methods to demonstrate how to apply the recommendations from this chapter in existing workflows.

APPENDIX 2A
Best Practices for Preservation

The case studies below provide in-depth descriptions of several different 3D data acquisition methods to demonstrate how the PIP (preservation intervention point) framework can be applied to existing workflows. These case studies are written from the perspective of a member of the project team. The purpose of each case study is to provide the necessary information to understand how the PIPs framework can be employed in different 3D imaging modalities. Because of the nature of the technology, in some instances terminology specific to the digital software or hardware for the project may be present.

Each case study examines a particular method of 3D data acquisition, the type of data collected, and the audience for those data to identify preservation and documentation needs for the project. Where it exists, accompanying text from a README file is provided as an example of best practices. At the end of each case study, a table identifying the PIPs for that project and recommended actions is provided for reference. The case studies are presented in alphabetical order—Sources-based, High-Resolution X-ray CT, Large-Scale Laser Scanning, Multimodal data collection, Photogrammetry, and Structured Light object scanning.

Each of the case studies presented here originates a different institution, for different purposes, and illustrates diverse methods and technologies for curating 3D data. Recommendations for implementation following the Good/Better/Best (GBB) guidelines given in the main text are purely contextual based on an individual institution's capabilities and resources for preservation. A single GBB assessment is given within each case study based on the method being illustrated and the institution to provide a practical explanation of how GBB is applied in the preservation workflow. These case studies are written from the perspective of the data creator and do not address SIPs and AIPs as expressed by the OAIS model for data preservation. For more information on repository operation, review the "Curation and Long-Term Access (SIP, AIP, and DIP)" section in this chapter or chapter 3, "Management and Storage of 3D Data," for more in-depth information.

Sources-based

Case Study: Lhasa VR 3D GIS and virtual environment
Project Dates: 2009–2016
Author: Will Rourk, University of Virginia

Acquisition Method

Sources-based/manually created 3D

Esri ArcGIS Desktop 10.7
Esri CityEngine 2019.0
Esri Unity 3D v5.6
Autodesk 3D Studio Max 2017
Agisoft PhotoScan 1.3

3D content was generated from the 2D GIS mapping of historical Lhasa and converted to an interactive virtual 3D environment using the Unity 3D game engine. 2D GIS is a method of creating maps with features linked directly to data either locally or from an online database. 3D GIS is a method of converting 2D maps in an *x,y* coordinate system to 3D content in an *x,y,z* coordinate system, allowing for spatial dimensions of geographical features to be emulated. Specific 3D models of historical buildings were manually created in 3D Studio Max, and 3D data were collected from the photogrammetry of historical monuments using photographs from fieldwork on site. These 3D assets were incorporated into the 3D GIS as detailed in supporting information. An online 3D virtual environment was created for interaction with 3D content through a web browser using the WebGL JavaScript API.

Good/Better/Best Assessment

The University of Virginia Library supports general methods of preserving data without specifically focusing on 3D data. A Best method of GBB is generally applied to data generated by 3D data collection methods or manual modeling. 3D assets associated with this project can be easily archived in the UVA Library's open scholarship platform, Libra (based on Harvard's Dataverse platform),[19] such as historically referenced 3D models created in 3D Studio Max or generated using photogrammetry. Archiving 3D GIS data is a procedure that is currently being researched due to the complexity of content involved. Geographic information system (GIS) projects involve many dependency files and file formats that are unified by a central project file. The UVA Library is moving in the direction of examining the use of data containers and emulation environments to archive the complexity of GIS projects and the myriad assets involved. Due to the current research conditions of archiving the complexity of GIS projects, a GBB assessment is difficult to administer. A Best method of archiving all project assets is achievable given the flexibility of the Dataverse open-access repository to ingest any type of uploaded file format. This is also true for the UVA Library's long-term archive, the Academic Preservation Trust.[20]

This project was part of a larger Mellon-funded effort under the umbrella of the Humanities Virtual Worlds Consortium (HVWC), in which the University of Virginia was a contributing member.[21] The main goal of research of this group was to create a platform for creating and exploring digital narratives with 3D interactive technologies. The main software platform adopted for this task was the Unity 3D game engine augmented by Drupal for custom interaction within the 3D model or world.[22]

The UVA team contributed to the HVWC by focusing upon Lhasa, the historical spiritual and political capital of Tibet. The objective was to create a comprehensive GIS of historical Lhasa and then convert the map data to a 3D GIS that would then be converted into interactive 3D for the HVWC platform. At the core of this effort was the generation of a 3D model of Lhasa and the Kyichu River valley in which it resides. Two methods were used to create 3D content: manual 3D modeling in 3D Studio Max and procedural modeling based on GIS data in CityEngine. All phases of the 3D data curation centered around the creation of individual models that were incorporated into the larger city model.

Acquisition Process

Planning

Coordination happened at two levels: within the HVWC and inside the UVA team. The HVWC planning included choosing tools to build the interactive 3D platform and selection of third-party programmers to help build custom code. Specific functionality of the platform was a major planning issue because interaction with 3D content needed to match the criteria of the four contributing institutional partners.

At UVA a team was assembled that included Tibetan historians, GIS experts, and a 3D content specialist. A historical narrative was created with a temporal focus on Lhasa as it grew from the mid-seventeenth century CE to early twentieth century CE. The team planned to first create a comprehensive GIS model of historical Lhasa, prior to Sino occupation in 1959. Many historical buildings were demolished and removed after 1959, so the intent of the GIS was to identify as many buildings as possible that once composed the city. Much of the planning centered around identifying maps, aerial images, text, and other documentation that would reveal original position, use, and appearance of these buildings—information that would be crucial to building the GIS of historical Lhasa. It was decided that the data would be held in the University of Virginia's Tibetan and Himalayan Library (THLib) Places database, an existing repository of buildings, sites, and place feature within the extents of Tibetan cultural areas.[23]

Planning Recommendation

This project involved many forms of data brought into a single platform for access and exploration. Organizing these data may require different techniques such as databases, GIS, and digital asset management tools.

Collection

Data collection began with obtaining documentation needed to build an accurate representation of pre-Sino-occupied Lhasa. This included the location of several historical maps, high-resolution USGS declassified aerial photos from the 1960s, SRTM digital elevation data, and the selection of historical images of Lhasa from the early twentieth century CE purchased from the Pitt Rivers Tibetan image collection at Oxford University.

Architectural information was gathered from quintessential books of Tibetan architecture, including *The Lhasa House: Typology of an Endangered Species* and *The Temples of Lhasa*, both by André Alexander, and *The Lhasa Atlas*, by Knud Larsen and Amund Sinding-Larsen.[24] These books provided detailed plans, sections, elevations, and textures for manually building key buildings in Lhasa by hand in 3D.

Collection Recommendation

Data collection is synonymous with research in this stage. Research asset collection methods can be best aided by consulting a librarian for strategies.

Processing

The processing stage was the most intensive for this project, as it was where the 3D content was generated using several methods centered around the construction of an interactive 3D GIS model:

1. GIS: A 2D GIS model was first constructed using Esri ArcGIS. The first task in creating the 2D map of Lhasa was to import the high-resolution aerial image from 1966 and digitize building footprint polygons on top of it. Most of Lhasa's historical buildings were still intact at this point in time, and their outlines were clearly discernible from the aerial image. Peter Aufschnaiter's map of Lhasa from 1948 was also initially used because it defined over 860 buildings with place names that existed in the city core of Lhasa at that time. After defining a building footprint, the polygon feature was then connected to an entry in the Tibetan and Himalayan Library's Places database. Over 3,500 features were defined and entered into the database. SRTM digital elevation data were also imported and georeferenced with the 2D map data. Images from the Pitt Rivers collection were connected to position markers in the map that indicated where the original pictures were taken. Other historical maps were overlaid and georeferenced with the map data. Roads, bodies of water, and vegetation were also mapped according to information from the historical maps and other documentation.
2. CityEngine 3D GIS: The 2D GIS data were imported via shape files into Esri CityEngine, a tool for creating 3D models from 2D GIS data. 3D digital terrain was generated from the SRTM digital elevation data to create the landscape of the Kyichu River valley. Buildings were auto-generated based on the data from the THLib Places database. A height field

existed in the database that indicated how many stories the building has in elevation. Code was written to take that field parameter and generate a placeholder 3D model by extruding the 2D footprint to an elevation in the building height field. Buildings were textured with a generic image based on traditional Tibetan architecture. Thus, 3,500 buildings could be generated quickly based on the back-end data. Roads, rivers, and pathways were also auto-generated based on GIS features.

3. 3D models: Custom 3D models of key buildings in Lhasa were hand-modeled in 3D Studio Max. These were based on the architectural information from the books listed above. The models were not extremely detailed but showed more information than the auto-generated models. The completed model was then exported out of 3D Studio Max and imported into CityEngine and connected to the THLib Places database via a unique identifier.
4. Unity 3D: Once the model was fully composed within CityEngine from procedurally created and manually constructed 3D models, the whole model needed to be transferred to the Unity 3D game engine platform. Elements were grouped together and exported simultaneously, such as buildings from different zones in the city, river features, road features, terrain, and so on. These groups were then imported into Unity 3D.
5. HVWC platform integration: The Unity 3D interactive model was then integrated with the custom code from the HVWC, which was downloaded from the third-party developers via GitHub.

Processing Recommendation

In this case, processing is complex and highly project-specific. However, thorough documentation of this processing is essential.

Curation and Long-Term Access

It was decided by HVWC that projects would be stored by the respective institutions.

The UVA project site is stored within the Tibetan and Himalayan Library, as is the database back end. Project and construction files are backed up in cloud storage. Plans to make these assets available to the scholarly community are still being decided. The original code for the consortium is available on GitHub.[25]

Curation and Long-Term Access Recommendation

Archival platforms must be able to handle several files and their dependency structures.

Output

The main output for this project is the Unity 3D online model and Drupal code dependencies. Many files were generated in the construction of this model:

- GIS Data
  - ArcGIS and CityEngine: .shp shape files, .tif geoTIFF (SRTM), .gdb geodatabase
  - WebGL ArcGIS webviewer
- Database
  - Tibetan and Himalayan Library Places Database—PostGres database
- 3D Content
  - 3D Studio Max: .obj, .fbx, .dae
  - CityEngine: .fbx
  - Unity 3D: .unityproj, export to .fbx, WebGL
- Web Content
  - Drupal dependencies

Output Recommendation

Complex projects generate a complexity of file types. Flexibility of storing and archiving diverse file types and maintaining their dependency directory structure is essential.

Usage

Usage is two-fold: use of the actual site and use of the HVWC platform. A website was created to show four different historical spatial narratives of Lhasa.[26] The main narrative is the photographic account of Lhasa made by London *Times* special correspondent Perceval Landon in 1904. His tour around the historical city is retraced within an interactive 3D model with accompanying photos and notes by the author. A guided tour is provided, but an independent tour can also be undertaken. This resource is open for public use and exploration.

The intent of the HWVC was to release the code and a completed platform to the scholarly community for use with 3D content generated by other institutions. At second phase funding, the project needs more work and refinement before it can be effectively released. The existing code is open for community access on GitHub as noted above.

Usage Recommendation

This is not quite a case of "Once you have the data you can do anything." But it is a case in which not only is the end product platform a useful tool, but so are the individual models and assets that went into creating the content for the platform. Making these available, as well as the end product, is recommended.

TABLE 2.A.1
Sources-based case study: preservation intervention points (PIPs)

PIPs	Actions
Planning PIPs	**Actions**
Build partnerships.	Divide responsibilities and agree on standard recording method.
Identify data sources and forms.	Determine need for databases or digital asset management tools to organize information.
Collection PIPs	**Actions**
Research existing information.	Utilize existing standards for research asset collection and consult a librarian for strategies.
Processing PIPs	**Actions**
GIS—annotate building footprint and add polygon feature.	Define and enter features into the database (e.g., elevation data, images, historical maps, roads, water bodies, and vegetation).
Export GIS for CityEngine.	Retain a copy of the original GIS file and the exported shape files.
CityEngine—add additional data.	Record data added and location of source.
Add custom 3D models.	Note which models were auto-generated and which were custom. Record source for decisions made during custom 3D creation.
Unity 3D.	Retain a copy of CityEngine files and exported files that were imported into Unity 3D. Maintain file structure for grouping together exports.
Platform integration.	Retain completed application. Document GitHub repository.
Curation and Long-Term Access PIPs	**Actions**
Review files to determine preservation needs.	For this project, the following file types were preserved. GIS data: • ArcGIS & CityEngine: .shp shape files, .tif geoTIFF (SRTM), .gdb geodatabase • WebGL ArcGIS webviewer Database: • Tibetan and Himalayan Library Places Database—PostGres database 3D content: • 3D Studio Max: .obj, .fbx, .dae • CityEngine: .fbx • Unity 3D: .unityproj, export to .fbx, WebGL Web content: • Drupal dependencies

PIPs	Actions
Archive with each partner institution.	Work with institution to maintain file structure—necessary for file dependency. Preserve assets as well as end product.

High-Resolution X-ray CT (HRXCT) of Burmese Amber

Case Study: High-resolution X-ray CT of Burmese amber
Project Dates: February–July 2019
Author: Jessica Maisano, The University of Texas at Austin

Acquisition Method

High-resolution X-ray computed tomography (HRXCT): Uses polychromatic X-rays to nondestructively detect differences in materials within an object based on their density and atomic number. It produces a 3D digital map of the object's structure that can be digitally resliced, **rendered**, segmented into its constituent parts, analyzed quantitatively, and used to produce a surface model that can be rapid prototyped at any scale.

Good/Better/Best Assessment

This workflow exemplifies a Better implementation of the tiered Good/Better/Best approach to 3D data preservation. It exceeds Good primarily because raw data are preserved, and these and the final data products are preserved in perpetuity. It falls short of Best because iterative versions of the data (from raw to client deliverables) are not preserved, and the data will not be placed in a publicly accessible repository whereby they could be repurposed.

Object

Specimens are pieces of Burmese amber (Cretaceous) from various mines around Myanmar. These pieces contain everything from crab claws and ammonites to bird wings and lizards. Most preserve only a mold of the original organism, but some do preserve bone or internal structure. Client is a commercial gem dealer, collaborating with an academic client of the University of Texas High-Resolution X-ray CT Facility (UTCT).

Acquisition Process

Planning

Client has examined each piece and described contents and region of interest to UTCT for scanning. Specimens are shipped or hand-carried to UTCT.

Scanning parameters depend on size of specimen, size and location of region of interest, and scientific importance of specimen. Sometimes multiple specimens can be acquired in a single scan depending on their shape and density. Client confers with UTCT as to desired resolution, enabling UTCT to determine the appropriate scanning geometry (i.e., **volumetric**, in which 1,000 or 2,000 slices are acquired in a single rotation of the stage; or helical, in which the stage also translates on its vertical axis during acquisition to enable high-resolution imaging of high-aspect-ratio samples).

Planning Recommendation

Client provides a unique identifier and locality information for each specimen, as well as an indication of the region of interest for scanning and desired resolution.

Collection

Most of these amber specimens are scanned on UTCT's North Star Imaging (NSI) scanner, although some requiring high-resolution close-ups are scanned on UTCT's Zeiss MicroXCT 400. Scanning parameters are determined at acquisition. For example, those amber pieces containing bone are rarer and thus more scientifically important, but often the presence of bone cannot be determined until digital radiographs are acquired of the specimen during scan setup. Data are then acquired, using acquisition quality standards corresponding to the scientific importance of the specimen while keeping within the client's budget (i.e., UTCT charges by the hour; in general, the longer the acquisition time, the better the data quality or signal-to-noise ratio). The resulting README text, called contents.doc at UTCT, accompanies the data to the client, to our archives, and to any repositories.

Below are contents.doc files (client/company name changed) representing the three possible imaging modalities for these amber specimens at UTCT: NSI volume acquisition; NSI helical acquisition; and Zeiss volume acquisition.

/Data from README/

NSI volume acquisition:

University of Texas High-Resolution X-ray CT Facility Archive

Smith:

Ref-29602: Scan of amber with insect (Ref-29602; Cretaceous, Myanmar) for Mr. Smith of Smith International. Specimen scanned by Matthew Colbert on 24 June 2019.

16bitTIFF: Scan parameters: NSI scanner. Fein Focus High Power source, 120 kV, 0.14 mA, no filter, Perkin Elmer detector, 0.25 pF gain, 1 fps, 1x1 binning, no flip, source to object 133.85 mm, source to detector 1316.703 mm, continuous

CT scan, 2 frames averaged, 0 skip frames, 3000 projections, 5 gain calibrations, 0.762 mm calibration phantom, data range [-20.0, 130.0] (grayscale adjusted from NSI defaults), beam-hardening correction = 0.1. Post-reconstruction ring correction applied by Jessie Maisano using parameters oversample = 2, radial bin width = 21, sectors = 32, minimum arc length = 8, angular bin width = 9, angular screening factor = 4. Voxel size = 9.72 µm. Total slices = 1755.

8bitJPG: 8bit JPG version of the above images.

Specimen Photos: JPG photos of the specimen.

/End Data from README/

/Data from README/

NSI helical acquisition:

University of Texas High-Resolution X-ray CT Facility Archive

Smith:

Ref-31068: Scan of a lizard in amber (Ref-31068; 47.350 ct) for Mr. Smith of Smith International. Specimen scanned by Matthew Colbert on 25 June 2019.

16bitTIFF: Scan parameters: NSI scanner. Fein Focus High Power source, 130 kV, 0.24 mA, no filter, Perkin Elmer detector, 0.25 pF gain, 2 fps, 1x1 binning, no flip, source to object 133.85 mm, source to detector 1316.703 mm, helical continuous CT scan, vertical extent 48.6 mm, pitch 8.1 mm, 6 revolutions, 3 sets, helical sigma 0.0, no frames averaged, 0 skip frames, 15000 projections, 5 gain calibrations, 0.762 mm calibration phantom, data range [-10.0, 200.0] (grayscale adjusted from NSI defaults), beam-hardening correction = 0.1. Voxel size = 9.72 µm. Total slices = 4155.

8bitJPG: 8bit JPG version of the above images.

Specimen Photos: JPG photos of the specimen.

/End Data from README/

/Data from README/

Zeiss volume acquisition:

Smith:

Ref-30248A: Close-up scan of a fang in amber (Ref-30248; Cretaceous, Myanmar, Kampti Mine) for Mr. Smith of Smith International. Specimen scanned by Jessie Maisano on 17 May 2019.

scan parameters: Xradia. 4X objective, 80kV, 10W, 4s acquisition time, detector 42 mm, source -37 mm, XYZ [-1238, 42501, 784], camera bin 2, angles ±180, 1261 views, no filter, dithering. End reference (45 frames, each for 4s). Reconstructed with center shift -1.5, beam hardening 3, theta 0, byte scaling [-40, 1100], binning 1, recon filter smooth (kernel size = 0.5). Total slices = 944.

REF-30248A rcp: Xradia recipe with scan parameters.

16bit: 16bit TIFF images reconstructed by Xradia Reconstructor. Voxels are 3.15 microns.

8bitJPG: 8bit JPG version of the reconstructed images.

/End Data from README/

Acquisition parameters to consider documenting include: kV/W of the source (determines penetrating capability/flux; higher energy required for higher atomic number/denser samples, whereas lower energy preferable for delicate samples or those with small differences in attenuation between their constituent parts); filtering of the X-ray beam (to block the lower end of the energy spectrum, thereby minimizing beam hardening and ring artifacts); binning the detector to decrease acquisition time/resolution (detectors are 2000 × 2000 pixels, but can be binned once [to 1000 × 1000 pixels] or, in the case of the Zeiss, twice [to 500 × 500 pixels]); distance between source, detector, and specimen (determines resolution, or voxel size, of the resulting dataset); number of radiographs to acquire (the greater the number, the greater the signal-to-noise ratio of the resulting dataset—but also the longer the reconstruction time); and time spent acquiring each radiograph (combination of time per view/averaging of views, again determines signal-to-noise ratio of resulting data).

Recording the scanning parameters also permits those with knowledge of HRXCT to interpret artifacts they see in the resulting data (e.g., if the README file indicates that no X-ray prefilter was used, and the data are beam-hardened, then the reader knows to prefilter the X-ray beam if scanning a similar specimen).

In addition to HRXCT scanning, UTCT takes digital photographs of biological and paleontological specimens for two purposes. First, these photographs document the condition of the specimens when received. Second, rendering programs may load HRXCT slice stacks in opposite orders; in some cases this may result in a 3D rendering of the digital volume that is the mirror image of the actual specimen. Most fossils are strongly asymmetrical, so it is important to have a photographic record of the specimen to ensure that renderings are not mirrored.

Collection Recommendation

Record all sample mounting and scanning parameters necessary for someone to replicate the scan. The client may have multiple samples that require the same scanning conditions for the data to be directly comparable. Also, it is not unusual for the client to return months or even years later with new samples requiring the same scanning protocol. Photograph the specimen if it is asymmetrical to check that 3D renderings are not mirrored.

Processing

Post-acquisition processing may include a beam-hardening correction, reorientation of the data, changing byte scaling, and a post-reconstruction ring correction, among other things. On UTCT's scanners, the first three are applied to the raw projection images using the proprietary scanner software.

A beam-hardening correction addresses the hardening (increasing mean energy) of the polychromatic X-ray beam as it passes through the specimen; left uncorrected, this artifact will make the specimen appear artificially more attenuating peripherally than in the center in the reconstructed slices. It is best to prevent beam hardening by filtering the X-ray beam (using glass, calcium fluoride, aluminum, brass, or steel of varying thicknesses), but most scanners have a polynomic correction built into their reconstruction software. It is important to record the correction applied, and it is possible to overcorrect.

Reorientation of the data involves rotating the volume so that it is more orthogonal to the specimen, if necessary. This can be done in the scanner reconstruction software or after data reconstruction in a program like ImageJ. Reorientation is important to note so the client can relate the data deliverable (a stack of 16-bit TIFF slices) to the raw projection images.

Byte scaling refers to the scaling of the HRXCT data to fill the available 16-bit grayscale space. Data that are 16-bit present a range of 2^{16} possible grayscale values from black (value 0) to white (value 65,535). If voxels (3D pixels) in the HRXCT data volume have a value of 0 or 65,535, this means that variation in those voxels has been discarded. Thus, it is recommended to byte scale to leave space at each end of the histogram of grayscale values. Data can always be rescaled to increase contrast post-reconstruction using a program like ImageJ.

The post-reconstruction ring correction is a program written in IDL to remove ring artifacts from reconstructed HRXCT slices. Ring artifacts are common in CT data, because any slight differences between channels or pixels in the detector will manifest as rings in the reconstructed slices (except in systems like Zeiss, where the specimen stage shifts by a few microns on *x*, *y*, and *z* from one projection to the next—dithering—to prevent ring artifacts). UTCT's RingFree post-reconstruction IDL ring correction program includes parameters such as bin width, number of sectors, minimum arc length, angular screening factor, and restriction of the grayscale range to process, to tailor results to samples of varying shapes and densities and artifact intensity. These parameters should be recorded because, if care is not taken, the ring correction can introduce new artifacts into the reconstructed data.

The documentation of some of these processing parameters is underlined in this README text:

/Data from README/

NSI volume acquisition:

University of Texas High-Resolution X-ray CT Facility Archive

Smith:

Ref-29602: Scan of amber with insect (Ref-29602; Cretaceous, Myanmar) for Mr. Smith of Smith International. Specimen scanned by Matthew Colbert on 24 June 2019.

16bitTIFF: Scan parameters: NSI scanner. Fein Focus High Power source, 120 kV, 0.14 mA, no filter, Perkin Elmer detector, 0.25 pF gain, 1 fps, 1x1 binning, no flip, source to object 133.85 mm, source to detector 1316.703 mm, continuous CT scan, 2 frames averaged, 0 skip frames, 3000 projections, 5 gain calibrations, 0.762 mm calibration phantom, data range [-20.0, 130.0] (grayscale adjusted from NSI defaults), beam-hardening correction = 0.1. Post-reconstruction ring correction applied by Jessie Maisano using parameters oversample = 2, radial bin width = 21, sectors = 32, minimum arc length = 8, angular bin width = 9, angular screening factor = 4. Voxel size = 9.72 μm. Total slices = 1755.

8bitJPG: 8bit JPG version of the above images.

Specimen Photos: JPG photos of the specimen.

/End Data from README/

Processing Recommendation

As with recording data collection parameters, it is critical to record data processing and correction parameters so that the final result may be replicated. Also, it is possible to overcorrect data, especially for beam-hardening and ring artifacts, so recording the parameters used enables the data consumer to determine whether this has occurred.

Curation and Long-Term Access

Curation at UTCT is achieved via a Microsoft Access database. The database is organized by project and records metadata like taxon name, museum accession number, specimen storage location, arrival and departure dates, and locality information. The database also records scanning parameters, from which the README contents.doc files accompanying each scan are directly exportable.

Long before the National Science Foundation required it, UTCT had a robust long-term data management policy. The goal is to never *have to* re-scan a specimen—unless the scanning technology has improved so much that the data will show significant improvement (this technology turnover occurs approximately every ten years). The cost to maintain HRXCT data and data products generated is covered in part by an archiving fee charged to the client, based on the total gigabytes of a project. This means that if the client loses their data at any time in the future, they can come back to UTCT for them.

UTCT archives all raw HRXCT data and client deliverables (e.g., reconstructed 16-bit .tif slice stacks, data analysis spreadsheets, data visualizations, etc.) to duplicate external hard drives; one is kept on-site, and the other is taken off-site. All client deliverables are maintained on redundant online servers at UTCT and backed up to the off-site Texas Advanced Computing Center (TACC) at UT's research campus and to the cloud-based UTBox.

Long-term access is available on the internal UTCT network to all of these resources. Short-term access is provided to the client via virtual server FTP, UTBox, or both. Occasionally for large projects, data are sent to the client via external hard drive.

Long-term access is also provided to a subset of biological and paleontological datasets via DigiMorph.org, the Digital Morphology Library. This NSF-funded Digital Libraries Initiative project resides at UTCT. And, via a recently funded NSF grant called oUTCT, approximately 9 terabytes of HRXCT data representing approximately 1,500 fossil and Recent vertebrate taxa will be uploaded to MorphoSource.org, where they will be more easily discoverable and repurposed. In addition, various UTCT datasets are reposited as supplemental information in journals upon publication.

Curation and Long-Term Access Recommendation

At a minimum, the deliverables to the client (at UTCT, the 16-bit .tif and 8-bit .jpg slice stacks, any derivative image processing or data analysis, and README contents.doc file) should be retained, for whatever duration is possible. These should be retained in multiple locations in case of hard drive or server failure. In the best case, all data generated during a project (raw scanner files, iterative versions as data are corrected, and those components already mentioned) should be retained.

The philosophy at UTCT is to keep the accessibility (and generally, expense) of a particular data type in line with the likelihood of revisiting it. For example, the raw scanner files are rarely revisited, so they are archived on inexpensive external hard drives. The client deliverables, on the other hand, are kept in duplicate on online servers because they are revisited much more frequently.

Output

Output from HRXCT scanning at UTCT ranges from raw scanner files to visualizations to packaged data analyses. These file types will vary from facility to facility.

The raw files produced by the HRXCT scanner prior to 2008 were individual .raw sinogram images. With the addition of the Zeiss scanner in 2008 came the proprietary .txrm and .txm raw file formats, which are a collection of digital radiographs. With the upgrade of the original scanner by NSI in 2013 came raw projection images in .tif format with associated .nsi proprietary files.

The analysis and visualization programs employed at UTCT also produce their own proprietary file formats: .hx (Avizo); .vgl and .vgp (VGStudioMAX); .ORSSession and .ORSObject (Dragonfly); and .sav (Blob3D).

Data acquisition deliverables to the client include a README file (Microsoft Word document) and 16-bit .tif and 8-bit .jpg versions of the HRXCT slices. The README file includes all of the relevant scanning parameters and data corrections applied to replicate the scans if necessary.

Data visualization deliverables to the client may include .tif or .bmp frames of animations, .avi, .mov and/or .mp4 movies, and/or surface models (.obj, .stl, etc.) to use in FEA analysis, for rapid prototyping, or both.

Data analysis deliverables to the client may include .xlsx spreadsheets generated from various analytical packages (e.g., Blob3D, ImageJ).

List of file types: .avi, .hx, .jpg, .mov, .mp4, .nsi, .obj, .ORCObject, ORSSession, raw, .sav, .stl, .tif, .txrm, .vgl, .vgp, .xlsx.

Output Recommendation

Not all file types have longevity. UTCT decided to deliver data to clients as 16-bit TIFF slice stacks because the TIFF file format has been around since the mid-1980s and is readable by all major data rendering and analysis programs. That said, 16-bit images cannot be handled properly or loaded at all by many programs, so it may be necessary to convert them to 8-bit. This transforms the scaling of the data from 65,536 to 256 possible grayscale values spanning black to white. The 16-bit data are much more information-rich, but the user can throw information away by converting them to 8-bit—the opposite is not true.

Usage

Most UTCT clients are academic researchers seeking information about the internal structure of their specimens for scientific research and publication. A smaller percentage are commercial clients seeking to reverse-engineer their specimen or determine what defects it may have.

For the typical academic client, usage will involve the publication of the HRXCT dataset and/or visualizations and analysis derived from it as figures or supplementary information; these data should be made available so that scientific peers can critically evaluate them, just as any other data. For commercial clients, usage will involve the identification of defects within the sample so they can be corrected during the manufacturing process, or reverse engineering to produce a superior product; these data typically are not seen outside of the commercial entity and often not archived by UTCT at the request of the client. In commercial cases involving historical objects (e.g., violins), usage may involve using the HRXCT data to identify unnecessary restoration that can be removed. In other commercial cases, the data may be used to prove or disprove the soundness of the scanned object in a legal setting. In most of these commercial cases, the data will not be seen outside of the commercial entity.

Secondary usage, by those other than the original client, is diverse. It may include K–12 and postsecondary educators who wish to incorporate HRXCT data into their lesson plans and labs or use these data to rapid prototype models for their classroom. It may also include academic researchers seeking to repurpose data for their own research. Artists represent another class that seeks to repurpose HRXCT data, for either digital modeling or rapid prototyping. Finally, publishers and producers often inquire after images and animations derived from HRXCT data for books and documentaries.

Usage Recommendation

If at all possible, make data available for repurposing. This accomplishes a number of objectives: it (1) minimizes specimen handling that would be required for multiple scans, which is especially important for precious natural history specimens; (2) extends the impact of the funds used to acquire the data, which often are provided by the taxpayer; and (3) allows the data to be used to address new and different questions outside the scope of the original research objective.

There are several options for making HRXCT data available for repurposing. For specimens scanned at UTCT there is DigiMorph.org—however, this library has not had active NSF funding since 2007, so it is necessary to charge the client on a cost-recovery basis to construct new DigiMorph pages. That said, existing DigiMorph pages represent data that have been cited in more than 300 scientific publications—most of which are examples of repurposing.

Other options include general data-sharing sites like Figshare, Dryad, MorphoBank, Nature Scientific Data, and OSF. Most recently, UTCT has decided to reposit approximately 9 terabytes of data into MorphoSource (Duke University), a repository with the specific target of volumetric data.

Wherever one chooses to reposit data to make them available for discovery and re-purposing, keep in mind any requirements regarding data ownership and copyright that the lending or scanning institution may have (see chapter 5, "Copyright and Legal Issues Surrounding 3D Data"), as well as ensuring proper acknowledgment of the sources that funded the data acquisition.

TABLE 2.A.2

X-ray CT: preservation intervention points (PIPs)

PIPs	Actions
Planning PIPs	**Actions**
Retrieve information from client.	Determine and record scanning objectives based on needs and budget of client. Record unique specimen identifier from client.
Collection PIPs	**Actions**
Scan specimen using parameters finalized at acquisition stage.	Create a README text (contents.doc) that documents the scanning and processing parameters and the types of files generated, which is shared with the client.
Processing PIPs	**Actions**
Beam-hardening correction.	Record the correction applied.
Reorientation of the data.	Record reorientation of the data so they can be related back to the raw projection images.
Post-reconstruction ring correction.	Record the correction applied.
Curation and Long-Term Access PIPs	**Actions**
Curate at UTCT.	Save relevant data into the in-house Microsoft Access database, including client information, specimen metadata, and scanning parameters.
UTCT file preservation policy.	Archive raw HRXCT data and client deliverables (.tif slice stacks, data spreadsheets, data visualizations, etc.).
Client deliverables.	Deliver via short-term FTP access or UTBox. Where applicable, upload to DigiMorph or MorphoSource.

Large-Scale Laser Scanning

Case Study: Large-scale scanning of University Hall, University of Virginia
Project Dates: September 2018–May 2019
Author: Will Rourk, University of Virginia

Good/Better/Best Assessment

As mentioned earlier (see the Sources-based case study) the UVA Library employs a Best method of preservation for 3D data due to the flexibility of both open-access and long-term preservation platforms in use. Dataverse and Academic Preservation Trust (APTrust), respectively, can both ingest a wide variety of data formats including esoteric file formats generated by 3D laser scanning technologies. This includes the preservation of proprietary 3D data file formats associated with the processing of raw 3D scanner data. The data from a FARO Focus 3D scanner are not a single file but a directory network of file dependencies that includes the 3D geometric and color data as well as scanner settings used during data acquisition governed by a single .fls formatted data file. The metadata for acquisition settings and conditions are included only in the project file and, since these data will be crucial for reproduction of the data in the future, the UVA Library accepts this file format and its dependencies as primary data for archiving and data preservation despite the proprietary, binary nature of the data. The raw project file can be exported to other more exchangeable formats such as .ptx or .e57, which are included in the data preservation package.

The 3D laser scanning process yields several 3D datasets necessary to fully document all surfaces and dimensions of a site or subject. The raw data processing of these datasets involves the registration and precise placement of all individual datasets in spatial relationship to each other to create a single comprehensive and unified dataset. The UVA Library preserves all individual raw datasets in addition to the registered data in the project file. Derivatives of the raw data can be exported to provide more exchangeable data formats that are more easily implemented in other software through immediate open access as well as useful for long-term preservation such as human-readable, ASCII-text-formatted files, which include .pts, .obj, .ply and .x3d. A Best assessment of recommended preservation is given here because the UVA Library has the capability of preserving esoteric as well as widely usable file formats for a single 3D documentation project.

Acquisition Method

Laser scanning—large scale
FARO Focus 3D model scanner (S120, X130)

Terrestrial 3D laser scanning technologies acquire data using laser light to capture the surface geometry and onboard color cameras to capture color data of the subject. Also known as lidar (light imaging detection and ranging), this method uses light bouncing off a surface to calculate the distance between the laser light source and the subject. The resulting data comprise a mass of points (commonly called a point cloud) that represents the geometric surface conditions of the subject as the time of documentation. Most 3D laser scanners are devices of metrology and are used for measuring the dimensions of a subject. In this case study, terrestrial 3D laser scanning is used because the laser scanning device functions on a tripod located on the ground.

Object

University Hall (U-Hall) was once the central indoor sports facility on grounds at the University of Virginia. It was built in 1965, and the design was based on a very large dome with a compression ring wrapped around concrete piers connected to concrete roof structure ribs. The building was replaced in 2006 and demolished in May 2019. The building was considered to be beyond restoration especially due to massive amounts of asbestos covering the arena ceiling, which accounts for 99 percent of the exposed ceiling structure. Ten months prior to demolition, a group was assembled to perform 3D documentation of the building. The documentation teams included robotic 3D scanning of the interior (Department of Computer Science and Engineering), aerial and terrestrial photogrammetry (UVAS Institute for Advanced Technology in the Humanities—IATH), 3D CAD modeling (Department of Architecture), architectural history (Department of Architecture), and 3D laser scanning of the interior and exterior (UVA Library Scholars' Lab). The 3D laser scanning captured the most data. The focus of interior documentation was to capture major public spaces, including the arena, main hallways and passages, and any ancillary spaces that defined the extents of the building interior. The focus of exterior documentation was to capture as much exterior surface as possible and rely on aerial photogrammetry for rooftop data.

Acquisition Process

Planning

The first stage began with conversations with the university lead conservator, which led to a wider conversation with the Office of the University Architect to determine what resources might be called upon within the university to perform a thorough 3D documentation of the site. It was determined that data collected could be managed by the library and made available to the scholarly community via the library open-access repository. A team was formed with specialization in the areas described above. Documentation of the site needed to coincide with activity still taking place at U-Hall. The arena had been sealed off from the rest of the interior to prevent any potential hazards from the asbestos-covered ceiling. It was decided that the main documentation effort would take place during winter break after the 2018 fall semester, but before asbestos abatement scheduled to begin in the early part of 2019.

A preliminary site visit made a couple of months before project initiation included a concurrent visit with all documentation team members. It was decided that the building was too complex to capture every single room given the project schedule. Architectural plans were prepared for a preliminary walkthrough of the site to determine which spaces were essential for documentation. These spaces included the main arena, the main hallways and passages, public meeting spaces, and ancillary spaces that describe the full extent of the building including utility closets and key office spaces.

U-Hall was quite well known in the region as a major basketball venue, so the architectural historians decided that a narrative could be developed along the path the basketball team would travel from the locker room to arena. Thus the space where the basketball team lockers once resided was included as an essential space, as was the former green room, where performers once prepared for shows in the arena.

The team was allowed into the arena for approximately an hour to discuss strategies for documenting the expansive space—roughly 300 feet in diameter and 100 feet in height. It was decided that 3D laser scanning would capture most of the data but would be supplemented by an initial autonomous robot scanner and terrestrial photogrammetry. Aerial photogrammetry was decided against inside the arena to prevent any disturbance of asbestos.

The teams planned to do a test scan of the arena to ensure that the technologies on hand were capable of efficiently documenting such an expansive space. For laser scanning, the architectural plans were edited to highlight areas to be documented and approximate scanning positions to determine how many scans would be needed. These scan positions helped to determine the amount of time it would take to actually capture the building data, which could then help determine how much time it would take to process the data for a completed set of documentation.

Planning Recommendation

Make sure all parties involved understand the nature and intent of the 3D data. Consider client wishes for levels of access to data.

Use architectural site documentation to plan scanning positions. If such documentation is not available, create hand-drawn sketches of the spaces to be documented. These are part of the documentation record, supplemental to the data.

Collection

The equipment used for laser scanning U-Hall was two FARO Focus 3D X130 and one FARO Focus 3D S120 laser scanners. There were two stages for the documentation:

1. Preliminary test

   Due to the expansive space, a few laser scans were taken in advance of the actual scanning to test the sensitivity of equipment and quality of data gathered. This testing was concurrent with the robotic scanner testing, whose team ended up being captured as data by the laser scanning test. This proved to be advantageous as they provided scale figures in the final data to show the immense scale of the arena. Two FARO Focus X130 scanners were used for the preliminary test.

The following procedure is followed whenever starting a new scan project:

1.1. Choose scanners and tripods that are most effective for the task at hand:

 Tripods include Manfrotto MT055XPRO3, Gitzo carbon fiber tripod, Nedo survey grade tripod.

 Two FARO Focus 3D X130 scanners were chosen for the preliminary test with Manfrotto MT055XPRO3 and Nedo survey grade tripods.

1.2. Project name: After scanners are booted up and mounted on tripods, a new project needs to be created in the settings on the scanner. The file naming protocol is in the form ProjectName_scanner#_YYYYMMDD, e.g. UHall_s2_20181224. Scans will be iterated starting with either 00 or 01.

1.3. Profile: A scanning profile is chosen to set parameters for data capture. The default setting for an interior space that is greater than 10 meters was chosen for the arena. This is approximately a 15-minute scan that captures 44 million points of data plus color photographic texture.

1.4. Scanner location: Choose locations that are optimal for collecting data that effectively document the site. In the preliminary test, data were recorded in key positions on the floor, first tier, and top tier of the stadium seating. Eight scans recorded enough information to show the project managers the extent to which the space could be recorded. Scanner location is marked on the "scanplan" or architectural plan of the site. The naming convention for marking a scanner location is in the form scanner#-scan#, e.g., S3-01. This number is consistent with the naming of each scan dataset that was input at the project naming phase above. A data and location number is also included in the title of the scanplan, which will also help identify datasets during the processing phase.

1.5. Level the scanner: Once the scanner is in its location and will not be moved, it must be leveled. The FARO Focus scanners have a built-in inclinometer that is accessed from the Sensors pane of the Manage window. This sensor acts like a digital dual axis level so that the scanner is completely level prior to scanning. The scanner will not scan if it is too far out of level, and leveling it also helps the data processing stage.

1.6. Prepare for scanning: Ensure that no people are in the way of the scanner and the surfaces to be recorded. Also clear out anything that might be obscuring surfaces. The most important features of a space are the areas that show the extents, like wall corners and tops and bottoms of walls, and any architectural details important to the historical narrative, like cornices, fireplace hearths, and window and door surrounds.

1.7. Start the scanner: Initiate scanner by pushing the green Scan button on the Focus 3D screen. Try not to place yourself in the scan or obstruct any surfaces needing to be documented.

1.8. Finish the scan: When the scan is finished, return to the home screen on the scanner and prepare for the next scan.

2. Actual documentation

A week was planned for complete capture of data. This week coincided with winter break to maximize building access. The laser scanning was split into interior and exterior space documentation stages. Interior spaces were prioritized according to access.

2.1. Interior

Arena: The team was given a small window of time, just a few days, to enter the arena space. The team members then coordinated schedules with each other so as to not get in each other's way. Two scanners were used to capture data, set diametrically opposite and moved around the circular space. A third scanner was brought in to scan the press box and entryway. Due to the asbestos threat, a certified P100 grade respirator was worn while working in the arena space.

Hallways and other interior spaces: Scanners were set up on both levels of the building to capture the major hallways. A third scanner was used to capture stairwells and utility spaces.

2.2. Exterior

Two scanners were used to capture the exterior, placed at opposite sides and moving in sync around the building. A third scanner captured data at all major entrances. Spherical targets were placed at these positions to help with the registration of interior and exterior spaces.

Total data collection took over forty hours of scanning time yielding over 100 individual scan datasets.

A final dataset was collected five days prior to the implosion of U-Hall, just before charges were set and nearly 85 percent of the building structure had been manually demolished. This provided a dataset of the raw structural system of the area and dome roof. Scanning time was nearly three hours resulting in sixteen scan datasets that were added to the original set of data.

Collection Recommendation

Scanning positions should be optimal to capture as many surface data of the subject as possible for thorough documentation.

Data resolution is a balance between sufficient data collection for thorough documentation and time needed to perform data collection.

Always do full 360-degree scans to capture more data than needed unless pressed by time constraints. Extra data are better than missing data.

Processing

Once the scanning is finished, all the scan data must be downloaded and processed. The main goal of the processing phase is the registration of individual datasets (scan readings) into one comprehensive dataset. FARO Scene (v. 2018) is the main software used for registration of data from the FARO Focus 3D scanners. The main objective of registration is to ensure that all individual datasets register as accurately as possible to maintain data integrity.

This stage has distinct steps:

1. Download the data: Each scanner records data onto an SD card, which must be downloaded onto the processing workstation.
2. Create a FARO Scene project: Create a project with naming convention SiteName_CollectionDate_scanproj, e.g., UHall_20181226_scanproj.
3. Import data into Scene: After all the data have been downloaded, the individual datasets can be dragged and dropped into the Scene project. Each dataset is not a single dataset but a directory of dependencies that includes the raw data. The topmost directory is considered the scanner data file, and this is what is imported into Scene. The data are in the form of raw point clouds. Color photo data are also included and automatically matched to the point cloud geometry.
4. Organize data in Scene: It can be difficult to process all of the data at once, so the datasets must be organized into what Scene refers to as "clusters." These are special folders that allow processing of sets of data rather than

the entire data volume, and they comprise scan datasets that were collected in relative spatial proximity to each other. For example, all the arena data would be put in one cluster, the first floor data in another, second in another, and so forth. It is recommended to further subdivide data organization into subclusters for more efficient registration. For example, the arena may be divided into sections like the main floor; north, south, east, and west quadrants; and the press box. The more atomized the datasets into clusters, the easier it is for Scene to process and register them. The scanplan will help determine which scans are within proximity to each other, which is why it is important to carefully record scanning positions during the collection stage.

5. Registration
    - 5.1. Choose an algorithm: Once the data are organized into clusters, registration can begin. This involves selecting a cluster and applying the Place Scans Operation. There are three algorithms that Scene provides for registering datasets:
        - 5.1.1. Target-Based—Use only if targets were used during scan collection stage. Targets include spheres, checkerboards, flat circular targets, and so on. Identify targets in the scan data, and name them so they are consistent between scans. Scene will then register datasets using at least three targets that are consistent between datasets.
        - 5.1.2. Top View—Scene can register datasets without targets using this algorithm by "viewing" the data from above and finding spatial similarities. The result is usually a rough association requiring further registration.
        - 5.1.3. Cloud to Cloud—This is another "targetless" registration method where Scene compares data points in the datasets to find similarities by which datasets can be joined together. This process is usually used to supplement the rough registration provided by the Top View algorithm.
        - 5.1.4. Correspondence View—Scene provides this method for manually moving datasets within proximity of each other. A view of all datasets in a cluster is provided, and individual datasets can be selected, moved, and rotated in place. This is usually used during the Top View registration phase when Scene has difficulty relating datasets on a rough draft level.
    - 5.2. Check registration accuracy: After each Place Scan registration, Scene reports back numerical statistics on how closely the datasets registered within millimeters of each other. Scene will green-light

registration if it is within about 5–6 mm of accuracy. A consistently yellow- or red-light registration result indicates that Scene is having difficulty finding relationships between scan datasets. Manual transformation of datasets using Correspondence View is usually the solution to obtaining a green-light result.

5.3. Iterative registration: Datasets generally need multiple iterations of registration (usually Cloud to Cloud) to obtain the best results based on the registration report.

6. Color the data: Once the data have been accurately registered, the color imagery can be applied to the geometry, creating a more photorealistic view of the data. Sometimes data are more easily interpreted in grayscale, in which case it is not necessary to color the data. Data from the U-Hall project were colorized for photorealistic appearance.
7. Create scan point clouds: This process is native to FARO Scene and is a way of optimizing the point cloud so it can be viewed in different ways, including virtual reality or video screen capture.
8. Apply clipping box: Creating scan point clouds also allows for the use of a clipping plane to help frame specific views of the data. The clipping box is a virtual cube in Scene that hides any data that are outside the box.
9. Create architectural views: The clipping box can be used to create architectural views of the data including plan, section, elevation, isometric, and perspective views.
10. Render architectural views: Once the clipping box is implemented to provide architectural views, these can be saved as screen captures for characterizing the 3D dataset outside of the Scene program.
11. Export data derivatives: The final phase of processing involves exporting datasets that will be included in the data archive as well as for immediate use. The dataset can be exported in total, which is recommended for archive, or from the clipping box, which is more efficient for immediate use in other programs.

Processing Recommendation

Use processing methods that allow for checking the accuracy of registration between datasets.

Curation and Long-Term Access

3D data at the University of Virginia are currently archived using three different methods:

1. Immediate backup via cloud storage and physical hard drive: This method is for safe backup of all files, both active and archivable, and is primarily for individual access.
2. Dataverse: This is an immediate access solution. This is an open-source, open-access repository that allows all assets to be fully accessed by the public. Data that are published in Dataverse are made automatically searchable and discoverable in Virgo, the UVA Library's catalog.
3. APTrust: This is a dark archive into which only the project file and first flush derivatives are placed for long-term storage.

Curation and Long-Term Access Recommendation

Consider different levels of archive access on a spectrum from open access to dark archive.

Output

Exported file formats for point cloud data include the following:

- .lsproj—This is the FARO Scene native project file that retains all metadata and detailed information from the scanner and from processing.
- .pts—This is an ASCII text format that preserves the data in their most basic form as seven values: *x*, *y*, *z*, *i*, *r*, *g*, and *b*. The values *x*, *y*, and *z* are 3D geometry values; *r*, *g*, and *b* are color values; and *i* describes the way light interacts with the surface.
- .ptx—This is a binary format that is a directory of file dependencies that includes photographic information as well as geometry.
- .ply—This format can be ASCII or binary and retains all color information with geometry. ASCII is preferred for archiving purposes, whereas binary is preferred for immediate use.
- .e57—This is currently used as an open-exchange format for point cloud data, although it is a binary file. It would be included for short-term archives.

Output Recommendation

Data formats should reflect immediate as well as long-term data usage. ASCII formats are predicted to be more useful in the future because content is human-readable and easily parsed. Propriety raw data formats will probably not be as useful in the future due to the need for specific software to interpret binary data, which will most likely be deprecated or dead in the near- to long-term future. However, these raw data files, usually project files, contain most of the metadata and detail information that are lost when exported to other formats. In a best-case scenario the project file should be archived to preserve metadata from the original capture of the data.

Usage

- 3D print
- virtual reality
- 3D modeling, CAD
- BIM (building information modeling)
- GIS

Usage Recommendation

The data are the most important part of the process. Anything can be done once the data are responsibly collected and processed.

TABLE 2.A.3
Large-scale laser scanning: preservation intervention points

PIPs	Actions
Planning PIPs	**Actions**
Build team of experts to document the site.	Determine responsibilities of each team member. Document this information and include required levels of access to the data by the client.
Conduct preliminary site visit.	Walk through the site with architectural plans and determine which spaces would be captured with which methods. Retain this documentation as supplemental data.
Collection PIPs	**Actions**
Conduct test scans.	Document equipment, parameters, and procedure used during test scans. Use test scans to determine extent of scanning and time required. Retain any documentation of agreements between parties.
Conduct documentation level scans.	In addition to the information above, document scanning positions and data resolution. Ideally, this level of documentation will include additional data beyond the minimum needed for the project. Retain original scans as raw data.
Processing PIPs	**Actions**
Import data for registration.	Retain a copy of the scene file into which original scans are imported.
Organize the data into clusters.	Save iterations of working files as needed. If useful, record clusters used for registration process.

PIPs	Actions
Register clusters.	Create a protocol that accounts for the registration algorithm and iterative registration in a consistent manner or documents deviations from that protocol at a reasonable level. Retain a copy of the registered data and any additional processing applied (such as coloration, scan point clouds, or clipping boxes).
Architectural views and rendering.	Create and render views as needed. Retain a copy of the rendered views and screen captures.
Derivative data.	Export the data for archiving and further use. Archival preservation should include all the data regardless of a clipping box.
Curation and Long-Term Access PIPs	**Actions**
Archive using University of Virginia standards.	Immediate backup via cloud storage and physical hard drive, Dataverse, and APTrust.

Multimodal

Case Study: Warm Springs Bath Houses 3D Site Documentation
Project Dates: April 2016
Author: Will Rourk, University of Virginia

Good/Better/Best Assessment

Because laser scanning was used in this project, the explanation for a Best method of data preservation is already given above (see Large-Scale Laser Scanning). The project illustrated here adds the photographic source material, processing, and 3D content results from aerial photogrammetry methods using a quadcopter and high-resolution camera. In this case, the individual captured source photos intended for archive are the .jpg files from the quadcopter camera. The photogrammetry project built using Agisoft PhotoScan (v. 1.3), resulting point cloud, and color texture mesh 3D content are also included in the archive. 3D content exported to similar formats as the laser scanning, such as .obj, .ply and .x3d, are included. The laser scanning data and photogrammetry data were combined using RealityCapture (v. 1.0), by Capturing Reality, to produce a comprehensive set of data from both types of 3D data.[27] The RealityCapture project file and exported 3D data formats are also intended for preservation. As mentioned above (see Sources-based), the project files are important for reproducibility of data in the long term because they capture all data acquisition as well as data processing metadata. Due to the proprietary nature of the project file formats, it is understood that these files will be difficult to access over time as software becomes deprecated. It is hoped that data preservation methods, such as data containers and emulation environments being researched by the Software Preservation Network, will increase accessibility to the information in project files over the long term.[28]

Acquisition Method

Terrestrial laser scanning
Aerial photogrammetry
Ambient audio recording

Terrestrial laser scanning is a form of lidar technology that uses light reflected on the subject surface to determine surface geometry and dimension as well as color imaging to capture surface texture (see Large-Scale Laser Scanning case study for more details). Data are acquired from a ground location. Aerial photogrammetry uses 2D images acquired from a UAS (Unmanned Aerial System) that are processed to create 3D data with color photorealistic texture (see Photogrammetry case study for more details). Ambient audio recording is achieved with a device that records audio in two-channel stereo or more channels to capture ambient sounds occurring at an existing location. Multimodal is a technique of combining two or more documentation technologies (modes) for a comprehensive spatial record of the subject.

Object

The Jefferson Pools, also known as the Warm Springs Bathhouses, are located in the western part of Virginia, close to the West Virginia border. Two structures exist on the site, each housing a pool of sulfuric water fed by natural springs. The northern structure (for men) has an estimated origin in the late eighteenth century CE, while the larger southern structure (for women) was built in the 1870s. In 2016 an assessment of their condition was performed due to decay and structural deterioration. It was requested that a full-site 3D documentation be performed to help with the assessment and provide for the historical record. A team from the University of Virginia was assembled to acquire data using terrestrial 3D laser scanning and aerial photogrammetry with a quadcopter UAS. Because water is an important factor for the bathhouses, ambient audio was also recorded at key points where water interacts with the site.

Acquisition Process

Planning

The UVA team, comprised of terrestrial and aerial documentation teams, coordinated with a member of UVA Facilities Management who was part of the preservation efforts at the Jefferson Pools. The team also worked with architect Terry Ammons of StudioAmmons in Petersburg, Virginia, who was charged with site investigation and generation of a Historic Structure Report (HSR), which served as the record of historical and present-day site conditions as well as recommendations for preservation.

Equipment chosen for data collection included two FARO Focus 3D X130 and one FARO Focus 3D S120 laser scanners for terrestrial scanning and a DJI Inspire quadcopter for the aerial data. Equipment preparations for the 3D laser scanners included cleaning the scanner optics, printing checkerboard targets, charging batteries, and making sure the tripod was secured. Aerial photogrammetry preparation included cleaning the HD camera optics on the quadcopter, charging enough batteries for at least two hours of flight, and making sure to pack extra quadcopter parts, such as extra rotors and a battery charger, as well as making sure the iPad tablet for the camera feed was charged and flight software updated.

The UVA team travelled to the site where they worked with maintenance crews who maintain the site for the Homestead Resort hotel, owned by Omni Hotels, in nearby Hot Springs, Virginia. Upon arrival the two documentation teams did a preview of the site to note key features to be documented and potential impediments such as locked rooms, unstable structural areas, water hazards, and tourist access. Timing of aerial and terrestrial documentation processes was key to ensuring that one team did not impede the other. The decision was made to perform terrestrial scanning inside the structures while the aerial team worked outside to gather images of the entire site.

Initial site preparation included the placement of checkerboard targets throughout the interior and exterior of the site for laser scanning registration. Checkerboards were also placed facing upward on the terrain to coordinate with the quadcopter data collection. Green tennis balls were staked throughout to also be used as targets for aerial data. Site plans were drawn of each building as well as the overall site to record scanning positions.

Planning Recommendation

In cases where the data acquisition process includes multiple teams, document expectations for each group, including the time frame when the work will be done. Take into account the need for shared space and resources so that multiple groups don't need access to the same resources at the same time.

Collection

Aerial data were collected by a two-person team, one navigating the quadcopter and one operating the HD camera on a three-axis gimbal. Each task required operating individual control systems. To capture the entire site, the quadcopter was flown in rows across the site to capture terrain data; then flights were made encircling the structures at varying levels in accordance with photogrammetry methods. Single-shot aerial photos were taken at higher altitudes restricted by FAA rules on flying UAS devices. Terrestrial photos were also taken using the quadcopter camera for views from the ground.

Laser scanning involved positioning three scanners around the site. Initial data collection began inside each bathhouse, with scanners placed in key positions to capture the most effective spatial data in each location. Major spaces were documented, but some minor rooms were excluded due to time constraints. Scanning positions were noted on the plans for each building with a naming convention of "scanner number-scan number." For example, 2-3 next to a square locator symbol on the plan indicated the position of scan 3 from scanner 2. Recording scan positions is key to successful registration of scan data in the processing stage.

Audio was recorded at the site after the 3D documentation stage. Recordings were taken at key points in the site exterior where water interacts with the structures and flows through the site. Recordings were acquired using a Zoom H1 stereo audio recorder.

Collection Recommendation

Treat individual collection methods as normal. Maintain clear and documented file structures for the data that takes into account the different acquisition methods. Each set of data should be a self-contained set of material that can be cross-referenced with the additional datasets.

Processing

After the data were collected by the two different techniques, each set of data was processed independently and then combined.

1. Aerial photogrammetry
   - 1.1. Download digital photos from the DJI Inspire 1.0 quadcopter SD card.
   - 1.2. Cull photos for bad shots, e.g., blurry photos, photos out of scope, misfired photos, etc.
   - 1.3. Collect photos into directories for each major site feature, e.g., women's bathhouse, men's bathhouse, full site.
   - 1.4. Main software for photogrammetry processing was Agisoft PhotoScan Pro on a 12-node computing grid. The PhotoScan process is a very detailed process; the following is a basic outline of the processing procedure:
     - 1.4.1. Create a new PhotoScan project for each major site feature.
     - 1.4.2. Drag and drop quadcopter photos into the PhotoScan Workspace.
     - 1.4.3. Create sparse point cloud.
     - 1.4.4. Optimize cameras and refine point cloud.
     - 1.4.5. Create dense point cloud.

1.4.6. Generate color texture.
1.4.7. Generate 3D mesh content.

2. Terrestrial laser scanning
   2.1. Download scan datasets from each scanner's SD card.
   2.2. Main software for registration of datasets is FARO Scene (v. 6.2). (For detailed workflow see the Large-Scale Laser Scanning case study.)
   2.3. After registration, color texturing, and point cloud generation, export .ptx file.
3. Multimodal processing
   Main software for processing laser and photogrammetry data is RealityCapture created by Capturing Reality. The basic process for combining 3D laser with photogrammetry data is as follows:
   3.1. Import laser data .ptx file generated by registration in FARO Scene.
   3.2. Import photogrammetry data from Agisoft PhotoScan.
   3.3. Generate comprehensive 3D dataset coordinated by establishing control points between the disparate datasets.
4. Audio processing
   4.1. Audio files were downloaded from the Zoom recorder's micro SD card and stored on a local drive.
   4.2. The main processing software was Audacity (v. 2.1.2).
   4.3. Audio files were clipped appropriately to create looping files in a 3D program.
   4.4. Files were exported to .mp3 for use in 3D software such as Unity 3D (v. 5.6).

Processing Recommendation

Develop a protocol for each method individually as well as a protocol on how to combine the materials.

Curation and Long-Term Access

3D data at the University of Virginia are currently archived using three different methods, which are described in the Curation and Long-Term Access section of the Large-Scale Laser Scanning case study.

Curation and Long-Term Access Recommendation

Multimodal data curation should take into account file dependencies and maintain a file structure that clearly marks how the datasets integrate with each other.

Output

Exported data files are contingent upon the process implemented as well as 3D formats that are best suited for near term and future use.

1. Aerial photogrammetry
   a. Photos: .rawformat .dng, high-resolution .jpg
   b. 3D mesh file formats: .obj, .ply, .dae, .x3d
   c. Project file: .pscan PhotoScan project file
2. 3D laser scanning
   a. 3D point cloud file formats: .ptx, .pts, .e57, .ply. x3d
   b. FARO Scene project file: .lsproj
3. Multimodal exports from RealityCapture
   a. 3D mesh formats: .obj, .ply, .dae, .x3d
   b. RealityCapture project file: .rc
4. Sound recordings
   a. .wav, .aiff, .mp3

Output Recommendation

Consider a variety of data formats that address the variety of data types included in the project.

Usage

- Preservation: The point cloud dataset was used by Richmond architecture firm 3North to help its design work on mediating the structural integrity and maintaining the historical construction and design of the Jefferson Baths during the preservation reconstruction process.
- A VR model of the Jefferson Baths was created in the Unity 3D game engine and made openly available in the UVA Clemons Library public VR space. The virtual model of the Jefferson Baths was used as a multi-participant meeting space for the CHIVR (Cultural Heritage Informatics VR) initiative in the spring of 2018 between UVA, James Madison University, and the University of Mary Washington. The model was enhanced by including 3D audio objects based on the sound recordings. Sound was associated at key points in the site and could be treated as 3D objects in Unity 3D with spatial characteristics such as attenuation and sphere of influence. Files were set to loop for continuous ambient sound.

Usage Recommendation

Multimodal projects have a variety of assets that can be reused and remixed for a variety of purposes. Careful consideration of possible reuse scenarios can assist in determining the level of preservation.

TABLE 2.A.4

Multimodal site documentation: preservation intervention points (PIPs)

PIPs	Actions
Planning PIPs	**Actions**
Choose technology specialists appropriate for site documentation needs.	Determine responsibilities of each team member.
Schedule site access with project coordinators.	Secure permissions from all groups involved.
Preliminary site preview.	Walk through the site with architectural plans and determine which spaces will be captured with which methods. Retain this documentation as supplemental data.
Team coordination.	Decide what areas need to be scanned when and by whom so as not to impede each other's progress or potentially block data acquisition.
Place targets for registration and control points.	When needed, place targets in areas that will overlap between scanning locations for registration and highly visible areas for control points.
Collection PIPs	**Actions**
Conduct test scans.	Document equipment, parameters, and procedure used during test scans. Use test scans to determine extent of scanning and time required. Retain any documentation of agreements between parties.
Conduct documentation level scans.	In addition to the information above, document scanning positions and data resolution. Ideally, this level of documentation will include additional data beyond the minimum needed for the project. Retain original scans as raw data.
Laws and regulations.	Be mindful of laws and regulations regarding the technologies employed for documentation, e.g., FAA rules on flying UASs.
Record scan positions.	Selection of key scanning locations should be backed up with proper documentation.
Supplement data with media.	Audio and video recordings can supplement data acquisition and become helpful assets in the final archive.

PIPs	Actions
Processing PIPs	**Actions**
Import data for registration.	Raw data files were imported from the two teams' equipment: terrestrial team—data was downloaded from the laser scanners' SD storage cards as FARO formatted .fls files; aerial—images downloaded from the quadcopter's micro SD storage card in .raw format with JPEG backup files.
Vet content to be used for creating data.	Raw data files from laser scanners do not need vetting and are ingested into the registration project file directly; aerial images from the quadcopter are reviewed on a workstation and bad images (blurry, poorly captured, do not include the subject) are thrown out;
Registration organization.	Terrestrial laser scan datasets are organized in the registration software manually, according to the spatial relationships of where the scans were taken on site. Clusters of datasets that are spatially similar are registered first dure to the overlapping of similar, captured data; for the photogrammetry registration, aerial photos with similar overlapping image information are collected together in the processing software. Images taken in circuits of similar camera view angles are organized for more effective processing.
Control points.	Checkerboard targets that were placed on site are identified where necessary to register the laser scan datasets in software; Photogrammetry software uses Structure from Motion (SfM) algorithms to find points of similarities between photos to register data together.
Derivative data.	Point cloud data generated by laser scan registration can be clipped down to expel data outside the range of the subject matter. This clipped data is exported as .PTS, .PTX or .LAS file formats for archive and immediate use.
Curation and Long-Term Access PIPs	**Actions**
Archive using University of Virginia standards.	Immediate backup via cloud storage and physical hard drive, Dataverse, and APTrust.

Photogrammetry

Case Study: Photogrammetry of artifacts from the Santa Cruz Museum of Art and History
Project Dates: 2016–2018
3D Data Creator: Abigail Crawford
Author: Marcia McIntosh, University of North Texas

Good/Better/Best Assessment

This workflow demonstrates a Good implementation of level of documentation, data preservation points, and long-term access and a Better implementation of preserved file formats.

Acquisition Method

Photogrammetry

Photogrammetry is "the science of making reliable measurements by the use of photographs and especially aerial photographs."[29] It is a nondestructive and affordable option for creating 3D data. Photogrammetrist and Sketchfab Community Manager Abigail Crawford completed this case study.

Object

Almost 200 artifacts housed at the Santa Cruz Museum of Art and History, including but not limited to a doctor's bag, a Screaming Hand sculpture, and a basket.[30]

Acquisition Process

Planning

Crawford consulted with the museum to find out the purpose of the models (educational/outreach), in what format the museum wanted the final products, and how the museum would like the files to be named. She also gained permission from museum staff to do the photography and, in some cases, contacted the artists to make them aware of the model's creation and let them know they could use the data if they desired.

Planning Recommendation

Match the intent for the model with the creation process. Intent can be based on the audience. Knowing the data's intended purpose will affect the manner in which they are documented, the final outputs, and where the data will go for preservation.

Collection

Abigail Crawford's photographic process for photogrammetry

- Camera: Crawford started with a 5 MP camera and later switched to a 24 MP camera with a fixed focal length lens (50 mm).
- Other equipment: She used a tripod and a remote control for the camera to decrease blurring of the images and a turntable for the objects being photographed.

- Setup: She placed targets (paper with text or lines) under the subject to help align photos in the processing phase. The targets were not moved during the shooting session to avoid photo alignment issues later. A circular polarizing filter (CPF) was used when shooting shiny or reflective objects to reduce bright spots. When shooting outdoors, she preferred to shoot on overcast or cloudy days. If shooting inside, she aimed to use diffuse 360-degree lighting when possible. It was not always available at the museum.
- Subjects: She avoided selecting subjects that were untextured, completely flat, very thin, transparent, shiny, or reflective. She generally selected items that are solid, matte, and textured.
- Capture: Shots were taken in .raw format. Depending on the subject, images were captured every 10 to 15 degrees horizontally and vertically with roughly 50 to 60 percent overlap between photos. Before shooting, the subject was studied to see what areas or features would require more images for an accurate model beyond the standard rotations. She generally took more photos than needed, particularly when the object could not be revisited, and then selected those that were used in processing.

Collection Recommendation

Organize files according to institutional or departmental source, object, and status (raw, final, etc.).

Document the setup (camera, degrees per shot, etc.) and equipment settings, in a README.txt, .csv, spreadsheet, or similar file to the extent that would be useful for the intended audience. Scholars and researchers will want this information to accompany data submitted to a publication, and it is needed for submission to a repository.

Processing

When Crawford used Agisoft PhotoScan (now Metashape), she converted files to TIFF in Adobe Lightroom or Photoshop. When processing in RealityCapture, she left the images in .raw format. The files were organized based on the naming convention given by the museum; she would use accession numbers or descriptive names based on the organization's preference.

Example of file naming system:

- (folder) Santa Cruz MAH >
  - (folder) 4-11-18 >
    - (folders) Red Hat, Boardwalk Clown, Projector >
      - (files) projector1.dng, projector2.dng, etc.
        - Final model format: Projector.obj, Projector.jpg, Projector.mtl

Crawford output the models to .obj format. She edited the mesh in Blender, which can import the .obj file and export either it as either an .obj (for PhotoScan) or a .dae (for RealityCapture) file.

She periodically saved to prevent loss. Blurry or unfocused photos were removed to avoid alignment problems. Depending on the demands of the software, the images may have been sent through a time-intensive process of masking before alignment. The mesh was then created. Once it was satisfactory, it was simplified to reduce the size to the number of polygons requested by the museum. The mesh was then imported into Blender as an .obj file. There, extraneous geometry was removed, and holes and other geometry problems were fixed. The mesh was next exported from Blender and imported as a .dae when using RealityCapture or as an .obj file when using PhotoScan. It was unwrapped and textured. After texturing, it was examined to make sure it looked accurate and exported into a .zip file. Crawford then sent the data to the clients through Dropbox. A pre-decimated mesh was kept in the project file.

Processing Recommendation

All photogrammetry data creators will want to save periodically during the processing steps to prevent data loss from software crashing or mistakes.

Cultural heritage preservation and scientific workflows should also document what processing choices are made for additional transparency and to assist in recreation of their models. This documentation may be accomplished by exporting a processing report from the photogrammetry software.

Curation and Long-Term Access

The files were put on an external hard drive along with the documentation and correspondence from the museum. The files were uploaded to the organization's Sketchfab pages.[31] The files were scheduled to be reviewed for data management and culling approximately every 10 years.

Curation and Long-Term Access Recommendation

At the least, have an internal backup of the raw data and final data model. It is recommended to make the data available online if possible. If used for cultural heritage or scholarly purposes, data products should be submitted to a repository.

Output

The clients received .obj, .mtl, texture images (these can be .jpg, .tif, or .png depending on the situation), and, optionally, the .raw photos.

Output Recommendation

Depending on the intent of the model, .obj, .mtl, texture images, README file, processing report, and other descriptive metadata may result in a completed photogrammetry dataset.

Usage

The intended usage for the Santa Cruz Museum models was outreach and educational purposes.

Usage Recommendation

The data can be consumed by the intended audience or others once accessible online.

TABLE 2.A.5

Photogrammetry: preservation intervention points (PIPs)

PIPs	Actions
Planning PIPs	**Actions**
Consult with stakeholders.	Determine audience and use to guide documentation and file format needs. Acquire rights and permissions as needed.
Collection PIPs	**Actions**
Set up space.	Record use of targets and equipment to create a better environment or shot (such as circular polarizing filter or additional lighting).
Capture images.	Save .raw images. Record camera settings and interval of photos.
Processing PIPs	**Actions**
Option 1: Process images for Agisoft PhotoScan.	Organize files using given naming convention. Save captured .raw files as .tif.
• Process model in Agisoft PhotoScan.	Save working files periodically. Record mesh reduction options. Export Agisoft metadata file. Preserve final PhotoScan file.
• Process model in Blender.	Retain copy of .obj exported from PhotoScan. Record if extraneous data were removed, if holes were filled, or other types of processing done.
• Finalize model.	Retain copy of textured export from PhotoScan.
Option 2: Process images for RealityCapture.	Use .raw photos for processing.
• Process model in RealityCapture.	Process and save periodically in RealityCapture.

PIPs	Actions
• Process model in Blender.	Export and use .obj in Blender to fix holes or extraneous geometry. Export from Blender as .dae and retain a copy of this file.
• Finalize model.	Retain a copy of the textured export from RealityCapture.
Curation and Long-Term Access PIPs	**Actions**
Personal archiving.	Save relevant files from previous stages with documentation, including correspondence from the museum.
Institutional archiving.	The clients receive .obj, .mtl, texture images (.tif, or .png depending on the situation), and, optionally, the .raw photos as per original agreements.
Online public access.	Upload derivatives to organizational Sketchfab pages as agreed.
Create a review period.	Review retention needs every 10 years.

Structured Light Scanning

Case Study: Lithic collection from the Early Upper Paleolithic Site of Tvarožná X, Czech Republic
Project Dates: 2016–2019
Author: Kristy Golubiewski-Davis, University of California, Santa Cruz

Acquisition Method

Structured light scanning (SLS) via a single-camera David SLS Scanner, now Hewlett-Packard (HP) SLS Scanner

Structured light scanning is achieved by projecting a series of known patterns onto an object using white light. The software measures the distortion of the pattern from a known point, which is calibrated at the beginning of the process. These measurements are used to create a point cloud, which, in the case of HP 3D Scan Pro software, is interpreted into a polygon mesh for processing and export.

Good/Better/Best Assessment

This project exemplifies a Better implementation of the tiered Good/Better/Best approach to 3D data preservation. The data are archived in the UMN institutional repository and will also be archived in tDAR for long-term storage. The choice to follow Better was intentional and reflected limitations in available data storage. Raw data were preserved, but neither the intermediate data nor the proprietary files associated with the project were preserved in perpetuity. A copy of these files is accessible through a departmental shared drive as an intermediate-term storage solution. This drive is backed up nightly but is not a long-term archival solution.

Object

This project was part of an effort to create 3D documentation of the lithic artifact collection from the Early Upper Paleolithic open-air site of Tvarožná-Za školou (Tvarožná X) to supplement traditional excavation methods.[32] The 3D dataset represents the 187 retouched tools, complete flakes, and cores that were over 2 cm in length.[33]

Acquisition Process

Planning

The primary concern in the planning stage was to determine the optimal parameters for the scans without having the lithic artifacts present, given that the artifacts were in the Czech Republic and the researchers conducting the scanning were in the United States. Two methods of 3D data acquisition were considered: photogrammetry (see description in Photogrammetry case study) and structured light scanning. The research team opted to bring both sets of equipment to the scanning location and spent the first day testing.

Prior to going to the field, the research group decided on the desired parameters for scanning. These included how to determine the selection criteria for scanning (tools, cores, and refits that are greater than 3 cm along the widest access); which scanning methods would be tested for use (structured light and photogrammetry); and the resolution, number of scan perspectives, and rotation degree of the turntable that would be used for each method. The 3D data would be acquired in the field and processed at a later date in the lab, necessitating the collection of additional data to compensate for scanning errors. A repository was identified (the institutional repository + tDAR) but not consulted at this stage. Consulting a tDAR representative at this time would have benefited the project, as it affected the final derivative file format and the method through which tDAR collects project metadata.

Planning Recommendation

When planning to acquire 3D data of material that is unavailable for testing and has unique characteristics, build in time for on-site testing. As possible, pilot test scans prior to travel and determine the functional requirements of the project and associated scanning parameters. Build in time beyond the expected time for the procedures to allow for scanning issues. Discuss the project with any repository resources early to build its limitations and requirements into the workflow, especially if the repository does not have an existing workflow specific to 3D data deposition.

Collection

After testing, it was decided that the best scanning option given staff expertise and time was the HP SLS scanner. A scanning protocol was created during the planning stage and refined during on-site testing. The parameters of the protocol are listed in the README file associated with the data and are included below for reference. As each artifact was scanned, information about it was entered into a Google Form, including any deviations from the protocol. For the scanning stage, the following information was captured:

- unique identifier (associated with the artifact's accession number)
- times: start time, end time, total scan time
- additional scans taken to capture the edge
- total number of scans captured
  - additional notes as needed
  - additional perspectives taken
  - issues during the scan
  - known refit and conjoin information

/Data from README/

The artifact was scanned on a David SLS 2 scanner using the 30mm calibration plate and the David automatic turntable.

The object was lit using the room lighting.

The artifact was placed vertically and flipped for 2 perspectives unless otherwise stated in the metadata file.

For each perspective, the artifact was automatically turned every 20 degrees 18 times for a complete 360 degrees and scanned at each angle.

If needed to capture the edge, the flake was manually turned so that the camera viewed the edge directly. A scan was taken and the edge was turned a maximum of 5 degrees twice in either direction to capture a total of 5 scans per edge.

/End Data from README/

Collection Recommendation

Determine a set of known parameters that will be the baseline for 3D data capture. Use a consistent method of recording any deviations from those parameters. Google Forms worked well for this project, as it was easy to use in the moment, created consistent data, and provided a spreadsheet output that was repurposed for the structured metadata that are currently associated with the dataset.

Processing

All of the 3D data were scanned first and processed at a later date due to time restrictions at the artifact location and computer processing power. As with collection, all of the artifacts went through the same protocol in the processing stage, with deviations collected via a Google Form. The protocol was made available alongside the dataset in the repository and can be found below. Multiple save points were kept during the working phase, with four points in the workflow identified as points to collect data via Google Form and ensure a save copy for internal preservation. Each form included a series of hand edits, which were recorded as present or absent for each artifact. Definitions for those edits were provided in the README file to provide transparency on the editing process. The following data were collected via Google Form during those points:

- unique identifier (to link the data from other forms)
- times: start time, end time, total scan time
- number of alignment attempts
- number of perspectives
- vertex spacing—after merge and final
- merge resolution
- hole filling—present/absent and program used
- if hand editing was performed in the following manners:
  - complete mesh removal
  - edge deletion
  - edge duplication
  - original edge reintroduced
  - partial mesh removal
  - picked point alignment
  - texture errors removed
  - labels
  - seams
  - mesh defects
  - defeatured (Geomagic X)
  - smoothed (Meshmixer)
  - flattened (Meshmixer)

/Data from README/

Initial alignment of the scans was completed in David 4 software using the automatic align feature.

Additional cleanup, global alignment, and hard to align scans were completed in Geomagic Design X. Specific processes used during this phase are noted in the Pre-Merge Data Editing Processes column of the file TvaroznaX_3D_metadata.CSV.

The photo textures were processed using Lightroom 5.5 and Photoshop CS6 to better approximate the original texture. This processing included using auto tone and auto balance to brighten the image, removing the background, and refilling the background with a color based on the midrange color of the artifact.

The final merge was completed in HP 3D Scan Pro (formerly David 5). Scans were set to merge with an estimated vertex spacing of .075 mm or less. The merged result was exported as an .obj and imported into Geomagic Design X.

Next, scan texture was converted to vertex color. The scans were centered around the origin, and oriented based on standard conventions for lithic analysis.

The mesh was copied, and manual editing operations were performed as noted in the Post-Merge Data Editing Processes column of the metadata sheet. If Meshmixer was used, the scan was exported from Geomagic, the required operation were performed, then the new mesh was reimported.

A global remesh operation was run with the target average edge length set to .075 mm. Next, the healing wizard operation was run to mitigate mesh defects. Color information was copied from the original mesh imported into Geomagic.

The mesh was then exported from Geomagic as an ASCII PLY file.

/End Data from README/

Processing Recommendation

Identify points in the process to retain working copies of the 3D data. The HP SLS scanner creates shelled data. The working copies for this project included four consistent processing save points, which effectively acted as PIPs: first-round edits just before merge; final edits just before merge; just after merge; and final object. Collect documentation on deviations from the protocol in a manner consistent with the collection stage.

Curation and Long-Term Access

Three preservation strategies were followed for this project. Each strategy reflects different use case needs for the data and limitations of the strategies available at the time of publishing the dataset. The metadata and paradata collected with the project were collected into both a README text file and structured .csv file that are associated with the data in each instance. The use of Google Forms early in the process expedited the process of creating both the structured .csv data, which collected the most relevant information, and the unstructured README text file. The data in tDAR and DRUM include only the final derivative data and neither the raw data nor the working files, reflecting the funding available for preservation and preservation strategies discussed with both repositories.

The differences in file format reflect the requirements of the repositories. The original derivative data were in .ply format and were exported again to .obj for tDAR.

The three preservation strategies followed for this project were

1. Maintain a copy of identified preservation working points, raw data, and final derivatives on departmental drives. This allows the researchers to share the raw data if requested and the working files that are most likely to be used if reprocessing of the data is required. This is not a long-term preservation strategy, but the research team determined it was the best option available for the full set of data associated with the project.
2. Preserve a copy of the final derivative data in .ply format via DRUM (Data Repository for U of M), the institutional repository of the PI. Documentation is provided at the project level via a README file and a .csv file.
3. Preserve a copy of the final derivative data in .obj format via tDAR (currently in process). Documentation will be provided at both the project level, in a manner similar to the DRUM repository, and at the object level as README files.

Curation and Long-Term Access Recommendation

Discuss preservation strategies with the intended repository early, including file format options. Have multiple strategies and consider options for preserving raw data and mid-process data that capture important workflow decisions.

Output

Derivatives for the project include .ply files and .obj files. Both file types include textured images. Proprietary file formats include the original scan files (.davidproj) and the Geomagic .x3d files that were created during the processing stage.

Output Recommendation

Ensure that any raw data or derivatives for the project are available in both proprietary and nonproprietary formats.

Usage

The intended usage for the scans is for scholarly and research purposes. Specifically, the models can be used to view identifying information such as the bulb of percussion and ripples and provide models detailed enough to allow for meaningful measurements to be taken from the 3D data.

Usage Recommendation

Ideally, the preservation and public accessibility stage of this project would have included the raw files. Though that was not possible for this case, the raw data are being kept on networked departmental drives that are backed up regularly. Additionally, detailed notes on the 3D creation and processing stages were kept to document the work done to the data.

TABLE 2.A.6

Structured light scanning: preservation intervention points (PIPs)

PIPs	Actions
Planning PIPs	**Actions**
Determine functional needs of the project.	Record expected parameters and protocols. Develop a method to consistently acquire data.
Consider space planning needs.	SLS requires control over lighting conditions and access to power. Incorporate possible issues into the expected timeline and protocols.
Collection PIPs	**Actions**
Scan capture.	Record information that deviates from protocol or expected parameters. Recommended information to record with SLS: number of scans; number of scan orientations; and object data not already collected elsewhere that affect the scan. Retain a copy of the raw files in both proprietary and nonproprietary formats.
Processing PIPs	**Actions**
First round pre-merge processing.	The first round of processing should stick as closely to the protocols as possible. Record any variables and retain a working copy at this stage. If the initial test merge is good, move to texture processing.
Second round pre-merge processing.	The second round of processing should include additional hand edits required to achieve a satisfactory merge. Record additional work as needed. Retain a working copy of the data at this stage.
Texture processing.	If texture editing is part of the protocol, retain a copy of the textures both before and after any image processing.
Merge.	Retain a working copy of the data just after the merge point in the workflow. Document any merge settings used that aren't already documented in the protocol.
Post-merge processing.	Document any post-merge processing done to the data. Retain a copy of the final data in both proprietary and nonproprietary formats.

PIPs	Actions
Curation and Long-Term Access PIPs	**Actions**
Internal preservation strategy.	Determine an appropriate location to store raw, working, and access-level data. Clean up the file directory prior to closing out the project.
Public access preservation strategy.	Long-term preservation in publicly accessible locations should include documentation from earlier stages in a structured format accompanied by a document describing the structure of the data. Ensure data are accessible via both proprietary and nonproprietary file formats.

Notes

1. Edward M. Corrado and Heather Moulaison Sandy, *Digital Preservation for Libraries, Archives, and Museums*, 2nd ed. (Lanham, MD: Rowman & Littlefield, 2017).
2. Consultative Committee for Space Data Systems, *Reference Model for an Open Archival Information System (OAIS)*, Recommended Practice, Issue 2, CCSDS 650.0-M-2, ISO 14721:2012 (Washington, DC: Consultative Committee for Space Data Systems, June 2012), https://public.ccsds.org/pubs/650x0m2.pdf.
3. Consultative Committee for Space Data Systems, *Audit and Certification of Trustworthy Digital Repositories*. Recommended Practice, Issue 1, CCSDS 652.0-M-1, ISO 16363:2012 (Washington, DC: Consultative Committee for Space Data Systems, September 2011), https://public.ccsds.org/pubs/652x0m1.pdf.
4. PREMIS Editorial Committee, *PREMIS Data Dictionary for Preservation Metadata*, version 3.0. (Washington, DC: Library of Congress, June 2015, last modified November 2015). https://www.loc.gov/standards/premis/.
5. Mark D. Wilkinson et al., "The FAIR Guiding Principles for Scientific Data Management and Stewardship," *Scientific Data* 3:160018 (March 15, 2016), https://doi.org/10.1038/sdata.2016.18.
6. GO FAIR, "FAIR Principles," https://www.go-fair.org/fair-principles/.
7. Archaeology Data Service, *VENUS Preservation Handbook* (Archaeology Data Service, 2008), https://archaeologydataservice.ac.uk/research/venus.xhtml; Kieron Niven, ed., "Data Selection: Preservation Intervention Points," Guides to Good Practice, Archaeology Data Service and Digital Antiquity, 2011, https://guides.archaeologydataservice.ac.uk/g2gp/ArchivalStrat_1-3.
8. Consultative Committee for Space Data Systems, *Audit and Certification*.
9. CoreTrustSeal home page, https://www.coretrustseal.org/.
10. Michael Witt and Jake R. Carlson, "Conducting a Data Interview," Libraries Research Publications, Purdue University, paper 81, December 2007, https://docs.lib.purdue.edu/lib_research/81/.
11. OAIS Reference Model (ISO 14721) home page, https://www.oais.info/.
12. Library of Congress, "Sustainability of Digital Formats: Planning for Library of Congress Collections," https://www.loc.gov/preservation/digital/formats/sustain/sustain.shtml.
13. Angus Whyte, "Five steps to decide what data to keep," Digital Curation Centre, October 31, 2014, https://www.dcc.ac.uk/resources/how-guides/five-steps-decide-what-data-keep.
14. Consultative Committee for Space Data Systems, *Reference Model* , p. 1-9.
15. Digital Preservation Coalition, "4.2.1.4.2 Preservation Description Information," Digital Preservation Coalition Wiki, last modified October 5, 2015, 12:50, https://wiki.dpconline.org/index.php?title=4.2.1.4.2_Preservation_Description_Information.
16. Cultural Heritage Imaging, "Digital Lab Notebook," http://culturalheritageimaging.org/Technologies/Digital_Lab_Notebook/.

17. Library of Congress, "Recommended Formats Statement," https://www.loc.gov/preservation/resources/rfs/.
18. FADGI Guidelines, http://www.digitizationguidelines.gov/.
19. University of Virginia Library, "Libra: Search and Submit," https://www.library.virginia.edu/libra/; Dataverse Project, "About," https://dataverse.org/about.
20. Academic Preservation Trust home page, https://aptrust.org/.
21. University of California, Los Angeles, Humanities Virtual Worlds Consortium, http://virtualworlds.etc.ucla.edu/.
22. Unity home page, https://unity.com/.
23. Tibetan and Himalayan Library, https://www.thlib.org/.
24. André Alexander, *The Lhasa House* (Chicago: Serindia, 2019); André Alexander, *The Temples of Lhasa* (Chicago: Serindia, 2005); Knud Larsen and Amund Sinding-Larsen, *The Lhasa Atlas* (Boston: Shambhala, 2001).
25. "HVWC/HVWC-Multiplayer-Platform," GitHub, last modified February 10, 2016, https://github.com/HVWC/HVWC-Multiplayer-Platform.
26. University of Virginia, "Lhasa Historic Tours," https://wordpress.its.virginia.edu/LhasaVR/.
27. CapturingReality home page, https://www.capturingreality.com/.
28. Software Preservation Network home page, https://www.softwarepreservationnetwork.org/.
29. *Merriam-Webster*, s.v. "photogrammetry," accessed November 6, 2021, https://www.merriam-webster.com/dictionary/photogrammetry.
30. Sketchfab, "Santa Cruz Museum of Art and History," https://sketchfab.com/santacruzmah/models; Santa Cruz Museum of Art and History home page, https://santacruzmah.org/; Santa Cruz Museum of Art and History, "Doctor's Bag," 3D model, May 13, 2016, Sketchfab, https://sketchfab.com/3d-models/doctors-bag-4846fe7d8fd4499ca1e62ea24b895b09; Santa Cruz Museum of Art and History, "Screaming Hand," 3D model, August 5, 2016, Sketchfab, https://sketchfab.com/3d-models/screaming-hand-15f218f80cde4286929d3b9c4c66107d; Santa Cruz Museum of Art and History, "Basket," 3D model, September 1, 2016, Sketchfab, https://sketchfab.com/3d-models/basket-dc3d8c07769e4226833cbd01a2936939.
31. Sketchfab, "Santa Cruz Museum."
32. Petr Škrdla et al., "Tvarožná-Za Školou: The Results of 2008 Excavation Season," *Přehled výzkumů* 50 no. 1–2 (2009): 11–24.
33. Gilbert Tostevin et al., "Lithic Collection from the Early Upper Paleolithic Site of Tvarožná X, Czech Republic," dataset, April 8, 2019, Data Repository for the University of Minnesota, https://doi.org/10.13020/hmrj-a558.

Bibliography

Academic Preservation Trust home page. https://aptrust.org/.

Alexander, André. *The Lhasa House: Typology of an Endangered Species*. Chicago: Serindia, 2019.

———. *The Temples of Lhasa: Tibetan Buddhist Architecture from the 7th to the 21st Centuries*. Chicago: Serindia, 2005.

Apple Inc. iPhone 8 Specifications. https://www.apple.com/iphone-8/specs/ (page discontinued).

Archaeology Data Service. *VENUS Preservation Handbook*. Archaeology Data Service, 2008. https://archaeologydataservice.ac.uk/research/venus.xhtml.

CapturingReality home page. https://www.capturingreality.com/.

Consultative Committee for Space Data Systems *Audit and Certification of Trustworthy Digital Repositories*. Recommended Practice, CCSDS 652.0-M-1, ISO 16363:2012. Washington, DC: Consultative Committee for Space Data Systems, September 2011. https://public.ccsds.org/pubs/652x0m1.pdf.

———. *Reference Model for an Open Archival Information System (OAIS)*. Recommended Practice, Issue 2, CCSDS 650.0-M-2, ISO 14721:2012. Washington, DC: Consultative Committee for Space Data Systems, June 2012. https://public.ccsds.org/pubs/650x0m2.pdf.

CoreTrustSeal home page. https://www.coretrustseal.org/.
Corrado, Edward M., and Heather Moulaison Sandy. *Digital Preservation for Libraries, Archives, and Museums*, 2nd ed. Lanham, MD: Rowman & Littlefield, 2017.
Crawford, Abigail. Abby Crawford Sketchfab Page. https://sketchfab.com/abby.
———. "How to Set Up a Successful Photogrammetry Project." *Sketchfab Community Blog*, June 18, 2015. https://sketchfab.com/blogs/community/how-to-set-up-a-successful-photogrammetry-project/.
Cultural Heritage Imaging. "Digital Lab Notebook." http://culturalheritageimaging.org/Technologies/Digital_Lab_Notebook/.
Dataverse Project. "About." https://dataverse.org/about.
Dataverse Project home page. https://dataverse.org/.
Digital Preservation Coalition. "4.2.1.4.2 Preservation Description Information." Digital Preservation Coalition Wiki, last modified October 5, 2015, 12:50. https://wiki.dpconline.org/index.php?title=4.2.1.4.2_Preservation_Description_Information.
DOI System (ISO 26324) home page. https://www.doi.org/.
Federal Agencies Digital Guidelines Initiative. Guidelines. http://www.digitizationguidelines.gov/.
Figshare home page. https://figshare.com/.
GO FAIR. "FAIR Principles." https://www.go-fair.org/fair-principles/.
"HVWC/HVWC-Multiplayer-Platform." GitHub. Last modified February 10, 2016. https://github.com/HVWC/HVWC-Multiplayer-Platform.
Larsen, Knud, and Amund Sinding-Larsen. *The Lhasa Atlas: Traditional Tibetan Architecture and Townscape*. Boston: Shambhala, 2001.
Library of Congress. "Recommended Formats Statement." https://www.loc.gov/preservation/resources/rfs/.
Library of Congress. "Sustainability of Digital Formats: Planning for Library of Congress Collections." https://www.loc.gov/preservation/digital/formats/sustain/sustain.shtml.
Merriam-Webster online dictionary. https://www.merriam-webster.com.
MorphoBank home page. https://morphobank.org/.
MorphoSource home page. https://www.morphosource.org/.
Niven, Kieron, ed. "Data Selection: Preservation Intervention Points." Guides to Good Practice. Archaeology Data Service and Digital Antiquity, 2011. https://guides.archaeologydataservice.ac.uk/g2gp/ArchivalStrat_1-3.
OAIS Reference Model (ISO 14721) home page. https://www.oais.info/.
ORCID home page. https://orcid.org/.
PREMIS Editorial Committee. *PREMIS Data Dictionary for Preservation Metadata*, version 3.0. Washington, DC: Library of Congress, June 2015, last modified November 2015. https://www.loc.gov/standards/premis/.
Santa Cruz Museum of Art and History. "Basket." 3D model, September 1, 2016. Sketchfab. https://sketchfab.com/3d-models/basket-dc3d8c07769e4226833cbd01a2936939.
Santa Cruz Museum of Art and History. "Doctor's Bag." 3D model, May 13, 2016. Sketchfab. https://sketchfab.com/3d-models/doctors-bag-4846fe7d8fd4499ca1e62ea24b895b09.
Santa Cruz Museum of Art and History home page. https://santacruzmah.org/.
Santa Cruz Museum of Art and History. "Screaming Hand." 3D model, August 5, 2016. Sketchfab. https://sketchfab.com/3d-models/screaming-hand-15f218f80cde4286929d3b9c4c66107d.
Scientific Data. Nature Portfolio. https://www.nature.com/sdata/.
Sketchfab home page. https://sketchfab.com.
———. "Santa Cruz Museum of Art and History." https://sketchfab.com/santacruzmah/models.
Škrdla, Petr, Gilbert Tostevin, Daniel Nývlt, Lenka Lisá, Ondřej Mlejnek, Antonín Přichystal, and Daniel Richter. "Tvarožná-Za Školou: The Results of 2008 Excavation Season." *Přehled výzkumů* 50 no. 1–2 (2009): 11–24.
Software Preservation Network home page. https://www.softwarepreservationnetwork.org/.
tDAR: The Digital Archaeological Record home page. https://www.tdar.org.
Thingiverse home page. https://www.thingiverse.com/.
Tibetan and Himalayan Library home page. https://www.thlib.org/.

Tostevin, Gilbert, Petr Škrdla, Gillane Monnier, Kristina Golubiewski-Davis, and Samantha Porter. "Lithic Collection from the Early Upper Paleolithic Site of Tvarožná X, Czech Republic." Dataset. April 8, 2019. Data Repository for the University of Minnesota. https://doi.org/10.13020/hmrj-a558.

Unity home page. https://unity.com/.

University of California, Los Angeles. "Humanities Virtual Worlds Consortium." http://virtualworlds.etc.ucla.edu/.

University of Virginia. "Lhasa Historic Tours." https:// pages.shanti.virginia.edu/LhasaVR/.

University of Virginia Library. "Libra: Search and Submit." https://www.library.virginia.edu/libra/.

Whyte, Angus. "Five Steps to Decide What Data to Keep." Digital Curation Centre, October 31, 2014. https://www.dcc.ac.uk/resources/how-guides/five-steps-decide-what-data-keep.

Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, et al. "The FAIR Guiding Principles for Scientific Data Management and Stewardship." *Scientific Data* 3:160018 (March 15, 2016). https://doi.org/10.1038/sdata.2016.18.

Witt, Michael, and Jake R. Carlson. "Conducting a Data Interview." Libraries Research Publications, Purdue University, paper 81, December 2007. https://docs.lib.purdue.edu/lib_research/81/.

Chapter 3

Management and Storage of 3D Data

Doug Boyer, Rachel Fernandez, Monique Lassere, Marcia McIntosh, Jennifer Moore, Francis P. McManamon, Albert Rozo, Todd P. Swanson, and Kate Webbink

ABSTRACT

The challenge of managing 3D data for long-term ***preservation*** *is one that many* ***repository managers*** *are grappling with as the technologies for creating 3D products become more accessible. This chapter aims to discuss the unique features of 3D data management and how repository managers are currently wrestling with 3D. Discussion of current data management standards is placed alongside responses collected through an international informational survey of 3D data managers and creators in the winter of 2018–2019, and the chapter concludes with recommendations for 3D data preservation. The survey respondents provided details on the systems and platforms used to store 3D data, how preservation packages are composed, what cost models are used to finance their creation and storage, and other areas of interest, including data types and retention methods. Readers will take away from this chapter an understanding of how the management of 3D data differs from management of other types, how*

creators can take an active role in the preservation of their 3D data, how existing managers are managing their data, and what recommendations serve as best practices in the work of preserving these assets.

Introduction

The management and storage of 3D data does not eclipse preexisting best practices for **digital preservation**. As digital libraries emerged onto the internet scene in the early 1990s and 2000s, they sought to create access to locally held, created, or managed digital content, particularly in the scholarly communication field. From this emphasis on access and the need for preservation of digital content, **digital repositories** were created. As the mission of digital repositories stretches to transcend traditional resources such as articles and books to incorporate formats such as research data, audio and video, creative works, and 3D data, so too must the skills and know-how of repository managers. This next phase in repository development is to provide access and safeguard these more demanding media types with systematic curation methods.

Data preservation is an essential outcome of the data curation process. The process of curation includes many key **preservation intervention points** (PIPs) in the 3D data stewardship life cycle. Essential treatments for long-term preservation are file inventory and validation, documentation and **metadata** creation, file format **transformations**, secure storage, migrations, and **checksums**. These steps absolutely apply to 3D data, but the best methods of accomplishing them are not necessarily clear. The unique needs for 3D data preservation in a repository overlap and compound to present a quandary many have stumbled over. According to survey data collected for this chapter, the greatest issues described include deciding what to **archive**, the size of that selection, where to preserve those items, and the overall storage requirement that result from having made those decisions.

The approach to these challenges can be shaped by the goals and mission of the hosting institution or platform. Commercial and independent repositories, such as Sketchfab and the Digital Archaeological Record (tDAR), will have different priorities in comparison to each other and other government, institutional, or academic repositories such as MorphoSource. In an effort to determine recommended courses of action, the CS3DP Management and Storage work group survey examined aspects—data organization, submission packages, file formats, repository technology, storage, funding, and staffing—and how they are being managed by individuals and repositories. This chapter summarizes those findings and offers recommendations for implementation.

Digital preservation as described by Corrado and Sandy is a triad of management, technology, and content-related activities.[1] It is key for a data manager to have firm awareness of all three preservation aspects when stewarding a digital preservation plan. This chapter interprets and structures itself around that triad into management, technology, and sustainability.

Survey Overview

The work group used two surveys to gather information on the issues of data organization, preservation, and sustainability. Respondents identified themselves as a 3D data creator, a repository manager, a creator and repository manager (creator-manager), or another perspective (other). The full results, available in appendix A, speak to how those in possession of 3D data are putting into practice data management methods.

The research began with an initial survey, called Survey One, that was circulated among the attendees of the first CS3DP forum in May and June of 2018. It contained open-response questions relating to repository information and scope, ingestion requirements, discoverability services, accessibility services, preservation/active curation, delivery/download services, sustainability, and additional information. This first survey acquired eight respondents. The results formed the presentation delivered by management group members at the second CS3DP forum in August 2018. After receiving feedback and suggestions on areas for further exploration, work group members began crafting what would become CS3DP 3D Data Management Survey Two.

Survey Two circulated in fall 2018 and spring 2019. The informational questionnaire consisted of four sections of multiple-choice and open-ended questions that centered on respondents' interaction with 3D data. The 53 respondents, who self-identified as creators, repository managers, creator-managers, and other, had separate questions relevant to their perspectives. The survey subsections focused on contextual information, repository infrastructure, usage and permissions, risk management, data management, and demographic information. Full results are viewable in the appendix A and the complete survey text in appendix B.

Sixty-four respondents began the survey, and 53 completed at least 25 percent: 14 creators, 20 repository managers, 17 creator-managers, and 2 other (see table 3.1). The largest group, 23, came from the library and archive domains, followed by 16 from university departments and centers, and 12 from museums. Nonprofits, commercial organizations, and a governmental organization were also represented. The most selected type of 3D data creation was **photogrammetry**, followed distantly by **born digital*** and 3D **laser scanning**. The majority of respondents from all groups

TABLE 3.1
Survey Two respondents

Group	Number of Respondents
Creators	14
Repository managers	20
Creator-managers	17
Other	2
Total	**53**

Note: These respondents completed at least 25% of the survey.

* Through our research, we have learned that the term "born digital" is ambiguous in the context of 3D data. We do not recommend use of this term.

began creating, managing, or archiving 3D data in 2015–2019 with 27, and 14 in 2010–2014. Only 5 began working with 3D data before 2000 (appendix A, table 3.A.4).

The following sections describe the managerial, technical, and sustainability duties of preserving 3D data.

Management

The management, or what could be more broadly described as documentation and policy-making, portion of 3D data preservation is a step that should not be taken lightly. Another key PIP occurs when designing **workflows** in the preservation process that can help circumvent issues that may arise later in the process. Corrado and Sandy stress the importance of documenting adherence to established best practices in the development and maintenance strategy of a digital preservation initiative.[2] One example of such documenting policies and completed work when dealing with data is found in the Trusted Digital Repository Checklist (TDR). The checklist's three major features are Organizational Infrastructure (governance, financial sustainability, and legal issues), Digital Object Management, and Infrastructure and Security Management.[3] 3D data management areas that may require documentation, or at the least decision-making, include certification, services, file states and submission packages, file formats, embargoes, and rights and licenses.

Certification

With an ever-increasing production and use of digital data over the past twenty years, there has been a greater impetus to use checklists to create international certifications for digital repositories. A digital repository storing 3D data should adhere to the professional best practices as depicted in OAIS (see the section "Curation and Long-Term Access [SIP, AIP, and DIP]" in chapter 2, "Best Practices for 3D Data Preservation"). This framework was designed to be accessible, in terminology and function, and is not specific to a particular community.[4] It is the basis on which trusted repository certification checklists, such as TDR/ISO 16363, are made and used for auditing and certifying digital repositories as trusted repositories.

Two of the main international certifications for all types of digital repositories are the CoreTrustSeal (formerly known as Data Seal of Approval) and the Nestor Seal for Trustworthy Digital Repositories.[5] The CoreTrustSeal involves an intensive review of repository functions and services, while the Nestor Seal involves a self-audit and provides an extended certification when coupled with the CoreTrustSeal. The Center for Research Libraries (CRL) also has a certification system, which is focused primarily on research libraries and requires a multistep audit using TRAC and ISO checklists.[6]

This certification involves a self-audit, analysis and a site visit by CRL, followed by the release of a final audit report. Although the number of repositories that have been certified by these systems remains low, the extensive lists of requirements provide valuable principles that repositories can follow to properly preserve and protect data, while making them accessible and reusable.*

When repository managers were asked what certifications their repositories have pursued, of the 5 total, 2 respondents selected CoreTrustSeal, 1 TRAC, and 2 responded that they were aiming to gain CoreTrustSeal and TRAC respectively (appendix A, table 3.A.37). Though a well-managed repository may be able to successfully preserve 3D data, repository managers are encouraged to seek certification or publish the results of a self-audit to increase documented reliability of their repositories.[7] The Good/Better/Best recommendations for certification are described in table 3.2.

TABLE 3.2
Certification recommendations

Tier	Description
Good	Aspire toward certification or publicly documented self-audit.
Better	Currently undergoing certification or publicly documented self-audit.
Best	Achieve certification or have published documented self-audit.

Services

Examining the certifications and checklists mentioned above, all digital repositories should offer certain services and perform a set of activities no matter what type of repository they are or what type of material they are preserving. When considering curation activities related to preserving 3D material, some of the most pertinent services that should be offered include file integrity and checks, record relation, **administrative metadata**, access, and file migration.

As the topic of services and repository activities was not mentioned within the survey, work group members sent participants follow-up questions concerning if and how these services are implemented within their particular repository. The questions pertained to the services listed in table 3.3, and out of the 27 respondents who were contacted, 5 replied from varied geographical regions and institutions. While 1 institution mentioned it was able to do the majority of the services tailored to 3D files, other institutions stated that they have not been able to offer any of the services in regard to 3D due to lack of expertise, use, and resources. These responses offer a glimpse into not only the challenges of working with 3D, but also into what areas there needs to be more research.

* As of September 2020, CoreTrustSeal has certified 163 repositories and the Center for Research Libraries has certified six repositories.

One manner in which repositories as service providers can prepare for future depositors is to have answers to the following questions:[*]

1. What is the cost, if any, for depositing data?
2. How are data deposited in the repository?
3. Does the repository offer additional curation services? What are the fees?
4. Is there appropriate metadata that meets the needs of the data?
5. Is the repository sustainable?
6. How is the material accessed?
7. What steps does the repository take to keep files secure?
8. What, if any, certification has the repository undergone?

The answers to these questions should be easily found on a website or through direct inquiry. Other recommendations for repository services can be found in table 3.3.

TABLE 3.3
Repository services recommendations

<table>
<tr><th>Description</th><th>Tier</th><th>Description</th></tr>
<tr><td rowspan="3">Recommended services to offer for 3D data:
• file integrity and checks
• record relation
• administrative metadata
• access
• file migration</td><td>Good</td><td>Working toward offering some of these services</td></tr>
<tr><td>Better</td><td>Offering some of these services</td></tr>
<tr><td>Best</td><td>Providing services in all of these areas</td></tr>
</table>

File States and Submission Packages

Before addressing the specific methods for dealing with 3D data preservation, it's important to first address the issue of data selection. Using the foundational guide on digital preservation of digital assets, the *Reference Model for an Open Archival Information System (OAIS)*,[8] **data curators** have adopted the practice of creating information packages for different purposes, which begins with the producer submitting data (**Submission Information Packages**, SIPs), followed by the curator treating the data and packaging them further for sharing (Distribution Information Packages, DIPs) and archiving (**Archival Information Packages**, AIPs). The OAIS states that SIPs are negotiated between the producer and the curator, but producers are generally encouraged to supply both their raw and final data to the greatest extent possible.

In several cases of 3D data creation, the information collected proceeds through several stages: a raw, just-captured state; a full data final version; and a reduced access

* List of repositories can be found at Re3data.org (Registry of Research Data Repositories), https://www.re3data.org/.

version. Creators have the opportunity to save the data in each of these stages. It is even best practice when creating data to save files periodically in order to avoid data loss during processing. Or in the case of sources-based or creative 3D data, there are native, often proprietary 3D modeling formats, as well as several possible derivatives of varying operability with other software and human-readability. Thus, when the time arrives to move the data or project to an archival stage, creators submit several versions of their data or several derivatives of different types to their repository. This practice, however, can be costly, and one of the largest challenges submitted by surveyed creators, repository managers, and creator-managers is the size and storage of 3D data (appendix A, tables 3.A.43, 44, and 45).

Survey results revealed that the majority, 69 percent of repository managers and 36 percent of creator-managers, from all perspectives preserve, or upload, the raw and derivative files (appendix A, tables 3.A.18 and 19). Another popular option is to upload only the derivatives while keeping the raw files offline. Some of the comments stated the files are dependent on what is provided by the creators at upload (appendix A, table 3.A.18). The survey responses suggest that despite the cost, creators and managers think having both a raw and processed version available is an imperative in the data's preservation. It could also explain why storage is reported as being the most challenging aspect of 3D data management.

The trend toward keeping raw files as the basic step in the data's preservation is supported by the Archaeology Data Service (ADS)/Digital Antiquity *Laser Scanning for Archaeology*.[9] It recommends several separate datasets be archived for laser scans, including, at a minimum, the individual raw scans and the final registered **point cloud**. The project, scan, and **registration** metadata should also be associated in the digital archive with the data. The guide also strongly advises that repositories include additional related products: the interim dataset used to create the final product and all transformation matrices, along with the metadata associated with these products.

Those working with **volumetric data**, such as CT or MRI volume data, should archive raw data and deliverables (TIFF slice stacks, data spreadsheets, data visualizations, etc.). They are also encouraged to reduce the size of their data as suggested by MorphoSource submission guidelines to crop down their data to the necessary volume and consider reducing bit-depth of their TIFF stacks.[10] See the case study "High-Resolution X-ray (HRXCT) of Burmese Amber" in the appendix of chapter 2, "Best Practices for 3D Data Preservation," for a detailed example. CT data submitters may also be asked to zip their data files and further reduce their space requirements.

When appraising data's value there are several characteristics one can consider, such as their scientific or historical value, non-replicability, and uniqueness and the economics of keeping them.[11] Given this publication's intent to focus on the long-term preservation of 3D data, the chapter's authors recommend that the selection of files included in a Submission Information Package be based on the intended use and preservation purposes of the data while also prioritizing their most original state. 3D technology has come a long way, but it still has opportunity for improvement. Therefore, keeping the raw data that could one day be reprocessed and produce an even better final object should be anticipated.[12] The Good/Better/Best recommendations for SIP contents are in table 3.4.

TABLE 3.4
Submission package tiers

Tier	Recommendation	Examples
Good	The raw data and an access version or image	Unprocessed capture, access copy or image (2D TIFF or JPG)
Better	Raw data, full-resolution derivative copy, access image	Photogrammetry: TIFF images, full-sized processed model Laser scanning: Unprocessed scan data, full-sized processed model, 2D image of the data
Best	Raw data, full-resolution derivatives, and access copies	Photogrammetry: TIFF images, full-sized processed model, final optimized model Laser scanning: Unprocessed scan data, full-sized processed model, final optimized model, 2D image of the data

File Formats

Given the many methods for creation of 3D data and their many consumption environments, the number of file format types and their capabilities are wide-ranging. Photogrammetry, for example, may be one of the easier raw formats to preserve because the data consist of 2D—hopefully, TIFF images, a file type already often used and preserved in the digitization world. A photogrammetry-produced derivative would also need to have a 3D format, such as ASCII versions of DXF, OBJ, or X3D.[13] A creator selects a 3D file format based on its support of certain features and its ability to function within a given software. In determining the best formats for archival preservation, a basic understanding of 3D data, formats, and viewers is useful.[14] McHenry and Bajcsy categorize the features of 3D data into three sets: appearance, geometry, and scene.[15] Geometry is the data points (vertices) and surface information (polygons or faces) and the ability to edit that geometry after exporting. Appearance is texture that is mapped to the surface of the model, its surface (diffusion of color, transparency, and reflection), lights (color and position), and environment. And lastly, a scene includes camera, light sources, and other 3D models.[16] Repository managers who encounter the need to convert one format to another should examine the features that are supported by the format type and check to see what features may be lost in the conversion. Secondly, they should also check to see if the conversion can be made to a format type that is interoperable between a wide variety of 3D software.

The interoperability of a file format allows for greater sharing opportunities among 3D data workers and their software of choice. According to the 3D-printing-focused platform All3DP, the most popular file formats of 2019 include STL, OBJ, FBX, COLLADA, 3DS, IGES, STEP, and VRML/X3D.[17] The PLY format is also reported popular within the academic and research world.[18] Many, but not all of these are program-neutral. Table 3.5 lists the open-source or proprietary nature of each format.

TABLE 3.5

Interoperability of popular 3D format types

(Adapted from Dibya Chakravorty, "8 Most Common 3D File Formats," All3DP, accessed June 25, 2019, https://all3dp.com/3d-file-format-3d-files-3d-printer-3d-cad-vrml-stl-obj [page content changed].)

3D File Format	Interoperability
STL (.stl)	Neutral
OBJ (.obj)	ASCII variant is neutral, binary variant is proprietary
FBX (.fbx)	Proprietary
COLLADA (.dae)	Neutral
3DS (.3ds)	Proprietary
IGES (.igs/.iges)	Neutral
STEP (.stp)	Neutral
VRML/X3D (.wrl/.x3d)	Neutral
PLY (.ply)	Neutral

Questions 21 and 23 on the management survey listed nine different 3D file formats, an "other" option, and "all of the above" to determine which file formats respondents use or accept into their repositories. Of the 14 creator respondents, there were 11 (79%) reports of using OBJs, and 10 (71%) mentions of PLYs in the creation of 3D data (appendix A, table 3.A.21). Repository managers also skewed toward accepting more OBJs and PLY file types when they did not already accept all format types (79 percent said they accepted all of the formats listed; appendix A, table 3.A.23. The prevalence of OBJ and PLY formats among creators is also supported by the Sketchfab 2019 survey results.[19] Their ability to preserve geometry and visual surface properties of a 3D object—if not scene light sources—also make OBJ and PLY the recommendation of ADS alongside COLLADA and X3D for more complex datasets and visualizations.[20] As of September 2020, the Library of Congress also listed STL, RTI, PLY, and OBJ as acceptable formats for 3D object outputs from photogrammetry.[21]

Among 3D data preservationists, the file format X3D, developed by the Web3D Consortium, has been recognized as an exemplary format.[22] This ISO-certified format can embed metadata and supports both individual 3D objects and virtual reality.[23] Its ability to embed metadata could ensure greater contextual understanding, **provenance**, and possibly citation of its data by future users.* The format does not come native to many software programs, however, and requires conversion from other more available format types, a process Web3D has documented on its website.[24] There is also less awareness of it among all perspectives surveyed. Out of 42 respondents, 21 (50%) were unfamiliar with and did not use the format, while 3 (7%) were familiar with and used it. Given its limited reach within existing software, the result is not unexpected. Still, the potential benefits of the format make its discussion and promotion worthwhile.

Repository managers should follow established best practices of the NDSA Levels of Digital Preservation of ensuring file , data integrity, and file format preservation.[25] For file **fixity** and data integrity, repository managers should aspire to conduct routine fixity checks for bit rot, institute overwriting safeguards, implement virus checks, and detect and/or repair corrupted data. For file format preservation, managers should provide input of preferred formats, inventory file formats in use, monitor obsolescence issues, and perform migrations and **emulation** as needed.

For 3D data, these format preservation actions would tangibly mean that a repository should quality-control the files to inspect for the data appearance and completeness. It should also clearly state somewhere the kinds of 3D data formats it prefers or recognizes (less preferred) and document the normalization paths (conversion of file formats before or after ingestion) that are used for submitted data. It is important to review converted files to ensure original data were not corrupted or lost unexpectedly during conversion. Similar to "hash checking" to ensure that a file was not corrupted during transfer, there may be similar automatable checks for quality control or validation of data between converted formats. It should be noted that limiting the variety of formats in a repository would result in less monitoring, or format watching, and thus require fewer future migrations to keep the data preserved and accessible. This sort of standardization can assist in controlling costs, which will be discussed later in the sustainability section. The initial selection of the file format will depend on the creator, but the repository manager can recommend or prefer a format such as the types listed in table 3. 6.

* The PLY format also allows embedded metadata. Both the binary and ASCII versions have an ASCII head that allows unlimited comment lines. The header may or may not be preserved by various software.

TABLE 3.6

3D preservation file format recommendations
(Chapter 2, "Best Practices for 3D Data Preservation," table 2.4c)

Tier	Recommendation	Examples
Good	A program-neutral format	.obj, .stl, .ply, .fbx, .tif
Better	An open-source, program-neutral format that is indexed by PRONOM	.obj, .stl, .tif
Best	An open-source, program-neutral format, indexed by PRONOM, that allows for structured and customized embedded metadata (ISO Standards encouraged) + The original file format	Open-source: .dae, .x3d, .ply Proprietary: .skp, .vue, .psx

Embargoes

Both creators and repository managers will grapple with embargoes, rights, and licensing issues when content is submitted or ingested into a repository. This holds true for all types and formats of data, 3D not excluded.

When choosing a repository for their data, creators will want to evaluate the embargo periods and terms. The timelines for embargo vary between digital systems and policies created by the institutions that host them. Creators should have reasonable expectations for how long data should be embargoed in order to serve potential users and should aim for the widest access allowable for the data.

Repository managers should strive to be transparent and flexible in establishing embargo periods. They can be transparent by stating when a dataset will be available and flexible by understanding that the data archiving needs of one creator may not fit other creators' needs in other fields. A wide range of variable time periods of embargo offerings could better suit one's submitters than a single, set standard.

Suggested standards for embargos can be found in table 3.7.

TABLE 3.7

Embargo recommendations

Creators	Repository Manager Tiers	
Creators should possess reasonable expectations for embargo period and access.	Good	Have an embargo period policy
	Better	Criteria from Good recommendation + published policy for potential submitters
	Best	Criteria from Better recommendation + multiple embargo period options

Rights and Licenses

When submitting data to a repository, creators may be asked to select or apply a rights statement or license to the submitted data. Chapter 5, "Copyright and Legal Issues Surrounding 3D Data," discusses these rights in more depth. It is recommended that creators thoroughly review all statements before agreeing and uploading their data.

Creators with current or newly established relationships with a repository will find that the repository's policies will govern how rights and licenses are handled. Most of the repository managers surveyed stated that they have submitters affirm that they have rights to the data they are submitting (appendix A, table 3.A.30). The majority also stated submitters will not be asked to transfer copyright upon submission (appendix A, table 3.A.31). Six of 14 in both repository manager and the creator-manager perspectives said they state the licenses (permitted uses) of their repository or recommend some (appendix A, table 3.A.33). Five out of 9 repository/creator-managers reported not requiring users to acknowledge the licenses (appendix A, table 3.A.34). Last, 9 of the 14 respondents stated that their repositories are open-access, with three having varying policies depending on the item (appendix A, table 3.A.35). While trends exist—affirming rights to data before submission, not requiring transfer of rights at submission, and stating use licenses—and it is encouraged that they continue, how to implement rights and licenses should be dependent on the collection, repository, and dataset (see chapter 5, "Copyright and Legal Issues Surrounding 3D Data").

The recommendation for rights and licenses can be found in table 3.8.

TABLE 3.8

Rights and licenses recommendations

Type	Description
Creators	• Should thoroughly review all statements before agreeing and uploading their data.
Repository managers	• Should affirm that creators have rights to the data before submission. • Should not require transference of rights to the repository at submission. • Should state use licenses for repository users. • Application of rights and licenses should be dependent on the collection, repository, and dataset.

Technology

Tools and Systems Landscape

Since 2010, data storage options available to independent creators as well as repositories for 3D data have changed drastically in terms of structure, scale, and cost. Average storage space cost per gigabyte has fallen from $10 in 2000 to $0.11 in 2009 to nearly $0.02 in 2017 for consumer-grade stand-alone hard drives on the scale of 1 TB to 8 TB.[26] Space is an important measure, but evaluating overall quality of storage in terms relevant to long-term preservation involves other dimensions, including system architecture, security, and management policies such as those recommended by the OAIS reference model (ISO 14721).[27]

Current (2019) networked storage requires more complex server setup and maintenance, but in return it can offer more robust data preservation and accessibility. Projects and institutions often coordinate with technology specialists to manage in-house storage, but new technologies change these relationships rapidly. Storage setup and maintenance are increasingly available through cloud services, ranging from storage in combination with more user-friendly applications (storage as a service, e.g., Dropbox, Google Apps) to hosted underlying storage infrastructure (infrastructure as a service, e.g., Amazon Web Services, Nasuni). Online consumer-grade storage options can currently accommodate demands for storing and transferring data on the order of gigabytes—enabling their use in 3D projects involving complex or large files—but creators need to be aware of the potential lack of long-term support and preservation.

Survey results indicate that creators differ from repositories in terms of their predominant storage methods (see table 3.9). While hard drives are the most prevalent storage method among creators (10), no repositories reported stand-alone (non-networked) hard drives as a storage method. Instead, most repositories (7) use server storage in combination with open-source (e.g., Samvera, Islandora, Archivematica) or proprietary (e.g., Preservica, EMu) database software. Database software in use included platforms for collections management (e.g., EMu) versus those more specialized for digital asset management (e.g., Preservica), which may include more robust file-preservation tools. The system architecture underlying those database platforms also included both open-source (Linux) and proprietary (Windows) operating systems.

Use of storage via online third-party services (MorphoSource, Dataverse, AWS, Figshare, Sketchfab [AWS], Drive, Box) was reported by 4 repositories and equal numbers of creators and creator-managers.

TABLE 3.9

Question 17—Preservation Platform: What system(s) are you using to store/ manage your data? Response comments are available for viewing in appendix A, table 3.A.17.

Group	3rd-Party Vendor	Server Setup	Hard Drives	Open-Source Option	Software Packages	Other
Creators	4	6	10	1	1	2
Repository managers	4	6	0	4	3	1
Creator-managers	5	9	5	1	0	2
Other	0	1	0	0	0	0

The OAIS reference model recommends that an archival system should at minimum be able to record the following for a given dataset.[28] (Prompts are included below to help repositories document system policies and standards regarding these functions. Also see chapter 2, "Best Practices for 3D Data Preservation.")

- *Identifiers*—How does the system record and assign **persistent identifiers**?
- *Related objects and context*—How does the system reference other objects or context (e.g., as in a relational database)? (This may be critical for 3D projects that involve data from real physical objects or places.)
- *History of ownership*—How does the system document and verify provenance information?
- *Fixity (data integrity) checks*—How does the system periodically check files to ensure they have not changed (e.g., due to bit rot or other corruption)? Are audit trails recorded? Are corrupted files flagged for database administrators or recoverable from periodic backups?
- *Access rights and restrictions*—How does the system allow for delayed or restricted access? (e.g., Can notifications be scheduled, or user groups and permissions defined?)

It is important for repositories to define and document system policies and standards in addition to the functions listed above. Doing so clarifies expectations for creators, users, and the repository itself. To facilitate longer-term maintenance, tracking usage over time can help estimate future needs for a system as well as maintain current policies (e.g., see "Repository Size Considerations" under the next section, "Sustainability"). Beyond the minimum recommendations, the specific system appropriate for a given repository may be dependent on staff and funding available for implementation and maintenance.

Software Archiving

Archiving stabilized copies of proprietary software can provide a repository with the future option of accessing proprietary obsolete formats by running the corresponding software in emulation environments. While this can be beneficial from a data preservation standpoint, it is important to be aware that the nascent domain of software preservation brings its own evolving technical and legal complexities. Creators and repositories should also consider the potential costs of relying on customized or proprietary **rendering** software. To help avoid these costs, the Library of Congress assessment of stability for media types highlights the benefits of using standard, nonproprietary formats that can be preserved long-term and handled in open software.[29] For example, collections that include proprietary raw output (e.g., CTR files from Breuckmann **structured light** surface scanners or respective proprietary formats from other 3D scanning instruments) may need to preserve a copy of the raw dataset in a nonproprietary format such as OBJ, paired with MTL file for texture information and other text documentation of the scanning instrument settings and other metadata (see chapter 2, "Best Practices for 3D Data Preservation").

Some 3D data collections may be unable to avoid relying on specific or proprietary rendering software or methods to make the data viewable, for example, collections of creative works where 3D data might not be convertible to nonproprietary formats. One archive, the Canadian Centre for Architecture (https://cca.qc.ca), reported that it is exploring emulation to help handle legacy software and formats in its collections.[30] Academic or in-house repositories considering similar approaches might be covered under fair-use arguments explored in the Software Preservation Network's *Code of Best Practices*,[31] but it is important to consult legal counsel on specific use cases and licensing questions. To comply with (or minimize) special licensing requirements, also consider recommendations on software preservation from the Magenta Book[32]—for example, documenting software versions, dependencies, and system requirements may be especially important when archiving the software itself is not an option.

If a repository is more deeply involved with the development of software, in addition to its preservation, then employing approaches like Unix and GNU philosophies could help lengthen the lifespan of software alongside the data by making it technically and legally easier to maintain.[33] Whether a creator or repository is aiming for open or restricted access, the Unix design philosophy advocates for frugal code that is easier to maintain by virtue of its simplicity, brevity, and modularity. The GNU licensing philosophy advocates for open code that is free to run, modify, and share. In these more complex cases where it is necessary to preserve a package of data, workflow documentation, and software together, it becomes all the more important for repositories to clearly document how data should be submitted for preservation (see "File Formats" section above and chapter 2, "Best Practices for 3D Data Preservation" for SIP preparation).

API Automation

For repositories that aggregate and otherwise publish data online, an API can help structure metadata that accompany a 3D asset for discoverability and accessibility, but data standards differ among domains and communities. For example, the Global Biodiversity Information Facility (https://www.GBIF.org) Integrated Publishing Toolkit (IPT) is specialized for publishing biodiversity data as a Darwin Core Archive, which can include corresponding multimedia data in "Audubon Core" format.[34] By contrast, Dryad (https://DataDryad.org) is a more general repository of publications across academic disciplines and requires that submissions include a set of metadata fields based on DataCite's metadata schema rather than a particular community's ratified data standard.[35] Data encoding methods like JSON-LD (JavaScript Object Notation for Linked Data) may also offer routes for coordination and discoverability across domains but might not address issues specific to accessing 3D data. The technology recommendations are listed in table 3.10. For more information about 3D-specific APIs and software libraries, refer to the section "Interoperability" in chapter 6, "Accessing 3D Data."

TABLE 3.10

Technology recommendations

Tier	Recommendation	
Good	A system with documented policies for these recording repository services:	• Identifiers • Related objects and context • History of ownership • Fixity (data integrity) checks • Access rights
Better	A system that records changes to the above and facilitates the following:	• Data ingestion • Data migration • Data publication • Interoperability
Best	An open-source system supporting the all of the above.	All of the above

Sustainability

Repository Size Considerations

One of the continuing realities of digital preservation is the amount of storage it requires. When using 3D data, whether for documentation or for analysis, both creators and repositories must also manage large file sizes. It's not a new problem. In the 1999 white paper *Creating Digital Archives of 3D Artworks,* Levoy and Garcia describe the challenges of their 500 GBs of laser scan data from their Digital Michelangelo Project.[36]

While storage costs have decreased, 3D data's size and storage requirements are still areas of trepidation for data creators and managers. 3D file sizes for objects can range from gigabytes to several terabytes of data. File size is dependent on a number of factors including model resolution, source data resolution, and the model creation method. Typically, source data take up more disk space than the model that uses them, so they should be considered when deciding on retention of data.

When asked about the challenges of working with and managing 3D data, the answer most often given by both creators and repository managers was storage and file sizes. File formats and costs came in second (file format archiving decisions—file state and type—and the cost associated with storage). Of the 13 creators, 9 of them mentioned size or storage space of data; for repository managers, 4 of 6; and creator-managers, 3 of 10 (see appendix A, tables 3.A.43, 44, and 45 for detailed responses). Interestingly, more creators than repository managers identified size requirements as a challenge. This is perhaps a result of managers' experiences with securing enough storage for their repositories on a regular basis.

Although cases of 3D data creation can vary wildly, to add some perspective, the average size of the front of a digitized 8.5-by-11-inch document scanned at the standard 400 ppi is around 40 to 45 MB. Therefore, a complete two-page document (four total images capturing the back and front) could be around 160 to 180 MB. A raw scan of a miniature book that can fit in the palm of one's hand can be around 142 MB. Its reduced 3D printable STL could be around 40 MB; its textured zip OBJ, 70 MB. Then add in access images, metadata, an optional GIF, and one could come out with an archival package of around 250 MB.

This size is small in comparison to video files, where one minute can range from 2 MB to 84 MB depending on the resolution. However, the 3D data's sizes are highly dependent on what is being modeled or captured and its resolution as well as in what submission states files are kept. Several states for file formats of a statue or building would inevitably dwarf the miniature example above and exponentially inflate that 250 MB to unsettling amounts.

The challenges of working with large 3D file sizes described by the respondents included ensuring that the files can be ingested properly in a repository system, as well as being archived and migrated to new formats when needed. A variety of factors affect the size limit repositories may impose on the files they accept. Apart from costs, the manner in which files are ingested or retrieved will affect file size limits. Depending on a repository's underlying systems and software, working with large 3D files can prove difficult.

When asked about size limits for data ingest, 70 percent of the respondents acknowledged that there were limits in place. Reasoning for these limits included software and browser capabilities, along with issues with uploading and pages timing out. One respondent stated that files over 20 GB had to be uploaded in a different manner and there were limitations within their repository with supporting files over 500 GB (table 3.11).

Since file size limitations can vary depending on a number of factors, there is no de facto recommendation as to the appropriate size for creating 3D files. However, when considering the storage of the files in repositories, some things to consider are the following:

TABLE 3.11

Question 28: File Sizes: What do you consider to be a large data file size in your current repository? (*N* = 16) Response comments are available for viewing in appendix A, table 3.A.28.

Files Sizes	Count
< 4 GB	5
5–20 GB	3
21–499 GB	3
500 GB	2
1 TB	2
N/A	1

- *Uploading data*—What is the maximum file size that can be uploaded into the repository for an individual file?
- *Downloading data*—What is the maximum file size that can be downloaded from the repository?
- *Cost*—What is the cost of storage?
- *Growth*—What is the estimated annual growth of digital collections?
- *Culling data*—If culling is an option within department or institutional policy, and archival practices are taken into consideration, what are the important file components and associated documents that must be included in the submission package? (Refer to the section "Curation and Long-Term Access [SIP, DIP, and AIP]" in chapter 2, "Best Practices for 3D Data Preservation.")

Having identified these aspects, repository managers can begin to anticipate future storage issues and plan for them. Some ways large repositories have navigated the storage issue include the following:

- *Geographically distributed support network*—The 3D repository MorphoSource maintains data "replicated at two geographically separated storage facilities, on a third continuously updated tape archive and on a fourth set of tapes that are removed to off-site storage every sixty days."[37] Inspired by the LOCKSS (Lots of Copies Keep Stuff Safe) system, 3D repository MorphoSource plans to recruit storage partners to assist in maintaining part of the repository's data requirements. Secondary plans, should the partners not emerge or external funding not be secured, include charging for user registration or submission.[38]
- *Storage replacement fund*—The University of North Texas Digital Projects' TRAC documentation appendix O includes the digital libraries division's approach to managing storage costs.[39] Operating under assumptions including a long-term commitment to preserving and archiving all items it acquires, maintaining two complete copies of the content, and that storage costs will

> decrease in five years and the technology perform better than existing storage, the department established an Archival Storage Replacement Fund. In each funding cycle, it contributes a calculated amount to the fund using this equation:
>
> Base Storage Amount * Cost-Per-Year Multiplier = Yearly Amount
>
> This fund comes from aggregating the institutional support of the UNT digital collections, grants, local accounts, and funding pools.[40] The storage capabilities and demands will depend on the system used to house the data.

Since creators and repository managers will encounter file sizes and storage requirements that vary wildly, a general recommendation is not possible. The authors suggest, however, that repository managers plan ahead for larger data requirements than for 2D information storage (table 3.12).

TABLE 3.12
Size and storage

Type	Description
Creators	Storage capabilities and demands will depend on the system used to house their data.
Repository managers	Plan ahead for above-average storage costs due to the large file sizes of 3D data.

Retention

Abandoned data and orphan works are a problem for many academic and public repositories. Over time, tracking down long-lost creators or rights holders for proper citation, licensing, or other documentation can be tedious if not impossible. One way for repositories to prevent the issue going forward is to state a retention schedule and require user agreements up front when a creator is submitting data. At that point, a creator could sign an agreement that licenses a copy of the data for the repository to use, manage, preserve, and, if necessary, cull under explicitly documented terms.

The question of retention and when it is okay to cull data may arise in managing 3D data. Best practices and recommendations for the retention of 3D objects vary widely across use cases and disciplines. While some advocate for 3D data curators to migrate files to open, nonproprietary standards when possible and retain original files when necessary,[41] others more generally advise for the preservation of all data in as many forms as possible.[42] These varied approaches reflect an inability to define a creating community within 3D data applications, as William Kilbride explains in a 2017 article on the preservation of 3D data.[43] Further, they reflect the many formats and tools we use to create 3D data and make them accessible. The breadth of creators within the research and 3D data communities, coupled with the proliferation of often proprietary technologies, in many ways necessitates discipline- and application-specific approaches to retention.

Discipline-specific approaches to retention often come down to the software that is used to create those data. **Computer-aided design** (CAD) files provide a relevant example of this point. Repositories without the resources to maintain a peripheral archive of corresponding software and emulation environments could coordinate with a separate software archive for emulation needs or set explicit up-front retention schedule criteria based on availability of required software so that data submitters will understand when and why the repository may cull certain files or file types.

Current recommendations within this chapter regarding the retention of 3D data may not apply to all 3D data for reasons specific to discipline, institution, funding, and so on, as seen in table 3.13.

TABLE 3.13

Question 40. Culling Considerations: What affects preservation priority?
(e.g., value of raw scan data versus 2D or 3D derivatives) (Select all that apply) (N = 21)
Response comments are available for viewing in appendix A, table 3.A.40.

Priority	Count
Institutional/departmental policy	15
Rarity of data	7
Disciplinary factors	6
Legal terms	6
Data are kept forever	5
Submitter's desire	4
Other	4
Funder	1

The survey data provided a glance into when repository managers consider culling. As indicated in question 40, most decisions to cull data are affected by institutional and departmental policies (71 percent). Rarity of data (33 percent), disciplinary factors (29 percent), and legal terms (29 percent) were also significant factors. The prevalence of the notion that data should be kept forever was a culling priority for 24 percent. One respondent listed funders as a deciding priority, and 19 percent of respondents identified submitter's desires as a factor in the decision to cull data. Of the 19 percent that listed Other, respondents cited a variety of priorities: from keeping data as long as needed to be consumed on the web or for teaching and learning to risk of loss.

Creators indicated that data size/storage capacity is a significant challenge in the preservation of 3D data (appendix A, table 3.A.43). As referenced in question 19 (appendix A, table 3.19), "[Creator] Saved Files: What files do you preserve?" 3D data creators generally preserve their data in both raw and derivative files, while access for derivatives varies. Of those who responded, 60 percent answered that they keep both raw and derivative files, 20 percent indicated that they keep derivatives or final versions, and 13 percent keep raw files offline. Based on these results, creators often retain several versions of 3D objects, thus greatly increasing the size of data that must be preserved.

The relationship between file formats and retention is further exposed when examining respondents' answers to the following question "Do you see particular risks related to 3D files that need to be taken into account as part of risk management planning?" Respondents cited proprietary and raw data as well as file format, software, and hardware obsolescence as specific risks that threaten the accessibility of those data over time (appendix A, table 3.A.38). For example, proprietary raw formats generated by 3D scanning instruments such as laser scanners and structured light scanners (SLS) may be readable only with the hardware and packaged processing software. Without standardization or documentation, these datasets are at high risk of becoming inaccessible if not maintained alongside their required software and hardware. In this way, retention of data is inherently dependent on the ability to confirm that the object will be able to be opened in several years. The less knowledge of or confidence in a 3D file format and the tools to use it with, the more data we cull. Table 3.14 summarizes the retention recommendations.

TABLE 3.14

Retention recommendations

Type	Description
Creators	Be aware of format longevity and discipline-specific culling and repository policies.
Repository managers	• Good—Implement a retention schedule to prevent new backlog of orphan works. • Better—Document and implement a retention schedule following relevant discipline-specific approaches. • Best—State a retention schedule and require user agreements up front when creators submit data.

Funding

A majority of respondents (34, or 56 percent) who answered question 12, "How is your data archiving funded?" listed institutional support as a source of funding for archiving 3D data, presumably in an institutional digital data repository their university, museum, or other institution maintains (table 3.15). The second-most-frequent reply to this question was grants, presumably referring to a portion of the grant funds that were used to create the 3D data funded the archiving.

TABLE 3.15

Question 12. Cost Considerations: How is your data archiving funded? (Select all that apply) Response comments are available for viewing in appendix A, table 3.A.12.

Group	Grants	Institutional Support	Personal Investment/ Self-funded	Other
Creators (*n* = 14)	5	8	5	1
Repository managers (*n* = 14)	6	13	2	0
Creator-managers (*n* = 15)	4	13	3	1
Other	0	0	0	1

In question 14, the survey also asked about what plans repositories or data creators had to meet the long-term costs of curating their 3D data (table 3.16). In general, respondents plan to rely on institutional support for long-term data archiving. Thirty-four of the respondents (63 percent) expect institutional support to meet long-term archiving needs. An additional 5 respondents identified "endowment" as the planned means of long-term support.

TABLE 3.16

Question 14. Cost Considerations: How do you plan to pay for long-term archiving costs? Response comments are available for viewing in appendix A, table 3.A.14.

Group	Charge at Cost (Data Contributor Pays)	Charge at Cost (Data Downloader Pays)	Institution Support	Endowment	Other
Creators (*N* = 14)	1	0	10	1	4
Repository managers (*N* = 14)	3	0	12	2	2
Creator-managers (*N* = 15)	1	1	12	2	2
Other (*N* = 1)	0	0	0	0	1

While the answers to these survey questions provide us with information about how current levels of 3D data creation and preservation for wider access and use are funded, they don't provide insights into what the costs are to carry out the archiving and data access functions. Answers to a couple of other survey questions about cost estimates and cost charging by a few repositories provide some light on this aspect of financial sustainability.

There were 34 responses to survey question 10, "Does your repository charge to upload 3D data?" Most of the respondents (85 percent) do not charge for deposits at this time (table 3.17).

TABLE 3.17
Question 10. Cost Considerations: Does your repository charge to upload 3D data?

Group	No	Sometimes	Yes
Repository managers	14	1	2
Creator-managers	15	2	0
Totals	**29**	**3**	**2**

When the work group collected responses about financial costs and models related to curation and preservation of 3D data in the preliminary version of the survey, Survey One, 3 responding organizations reported that they charged fees for these services. Two of these repositories—the Archaeology Data Service (ADS) and the Center for Digital Antiquity (Digital Antiquity), also are included in the current survey and reported the same information.

For those repositories or creator-managers who responded positively to this question, the most common answer to question 11, "What is your pricing model?" was "It depends." Table 3.18 summarizes the comments received.

TABLE 3.18
Question 11. Cost Considerations: What is your pricing model? (*N* = 3)

Group	Response	*N*
Repository manager	Basic charge is $5/file with 10 MB allotted for each file.	1
Creator-manager	[Price] depends if we deposit with UK Archaeology Data Service as well. We use the ADS calculator.	1
Creator-manager	[Pricing varies for] internal v. external 3D files, also depending on funding source.	1

ADS is one of the repositories that charges for the deposit of 3D data. Its "Charging Policy" (updated in 2016) includes 3D data among the more complex data formats for which it applies a bespoke, or individually customized, approach to pricing:

> Archives comprising less well-known file formats and types, those with more than 300 files or with larger file sizes (c.100Mb or more) require a bespoke costing. The charging policy for these more complex archives is applied according to the individual needs of the project and estimates are provided by the Collections Development Manager.[44]

In response to Question 11, Digital Antiquity, which runs tDAR (the Digital Archaeological Record), noted that the standard charge for depositing data is $5 per file and each file is assigned 10 MB of space. This pricing, which enables Digital Antiquity to plan for long-term curation and storage for deposited files, seems to be satisfactory for many of the file formats created by archaeological and cultural heritage investigations and resources. However, this simple pricing model has not been particularly successful with complex datasets and file formats of 3D scans, which require much larger storage space, frequently by orders of magnitude, than digital documents, analytical datasets, and images.[45]

tDAR contains 158 3D scanned resources.[46] The aggregated size of these files is about 39,000 MB or 39 GB when compressed into .zip files; uncompressed, the file size would be two to five times greater. Most of these scans were deposited as part of a cooperative project with the Center for Advanced Spatial Technology in developing the tDAR software to support deposits of 3D and other scanned data and to develop metadata categories and standards for these kinds of files. By contrast with all other file formats for digital data in tDAR, these 3D scan files are substantially larger. The scanned image file requires 795 MB of storage space when zipped in a compressed file. Uncompressed, the storage space is slightly over 1 GB. Using the $5 per file at 10 MB per file rate, Digital Antiquity's long-term curation charge for this item would be $500. This price might not be too high for a project that created a single scan image or a few scans. However, for projects that create tens or hundreds of scans, the **digital curation** charges using this cost formula have not been practical or acceptable in a number of specific cases.

The survey and responses reported in this chapter have not provided a specific model or simple formula for estimating the cost of long-term archiving and curation of 3D data.[47] The "customize" approach taken by ADS, which may have the longest experience of archiving these kinds of archaeological data, probably is the best way to proceed at present.

The ADS/Digital Antiquity Guide to Good Practice "The Project Archive: Storage and Dissemination" makes important general points that need to be kept in mind when estimating long-term costs for individual projects. Most importantly, long-term archiving is not a simple matter of data storage cost.

> …"storage" covers not only the size of the media on which data is stored and backed up, but also encompasses the ongoing periodic process of data refreshment (the movement of datasets to new hardware or software environments). While the cost of physical storage continues to decrease, that of refreshment and long-term curation—key factors in continuing to make data accessible and available—does not. In addition, in order to take advantage of technological advances and decreasing costs in certain areas, archives have to periodically upgrade systems or parts thereof.[48]

In addition to storage costs, there are also personnel costs, discussed in the section "Staffing Considerations" below, for the ongoing administration and management of the repository content which ensures that the content remains discoverable, accessible, and usable.

Obtaining a Better Understanding of Digital Curation Costs for 3D Data

As mentioned previously, the generation and use of 3D data are increasing rapidly in archaeology, bioarchaeology, physical anthropology, other cultural heritage fields, and natural resource fields as well. Laser scanning, lidar, computer visualization, and other scientific research applications produce large, complex datasets. However, there has been little focus on or understanding of the implications for cost and good practice in data preservation, dissemination, reuse, and access. It is time to focus more attention on developing case studies in order to draw out best practice guidelines on digital curation for 3D data.

A substantial body of work exists—and more is developing—that investigates the financial sustainability of digital data repositories.[49] Most of this has focused on repositories that curate digital documents, still images, audio, and video. However, some of these best practices and general good guidance from these studies can be applied to the digital curation for 3D data. Such studies also might serve as models and provide methodologies to research costs and best practice for 3D curation. Funding source recommendations are summarized in table 3.19.

TABLE 3.19
Funding recommendations

Type	Description
Creators	• Good—Grants • Better—Grants and institutional support • Best—Institutional support (particularly for archiving one's data)
Repository managers	• Good—Grants and institutional support • Better—Institutional support • Best—Institutional support supplemented by another income stream (endowment, charging for storage or services,* etc.)

* If charging for services or storage, recommend customized pricing.

Staffing Considerations

Similar to 2D data preservation, staff involvement in the preservation process of 3D data is best determined by specific workflows. Some processes work better by requiring the preservation process to occur directly after each item is digitally captured. Other processes work better by assigning others to collect and process digital captures for preservation at specific times throughout the process. Due to the many disparate fields that incorporate 3D digital captures into their work processes—and their highly technical nature—the digital capture manager needs to decide when and by whom this work best fits into their standard workflow. For CAD-based preservation, a working knowledge of the software used is essential to know what specific types of files are required for proper preservation.

Staff considerations for a data repository will be determined by its goals and capacity. The types of services a repository provides will determine its staffing numbers. If a specific repository is closed to the public and concentrates on one specific 3D data type from a few entities, a relatively small staff should be effective. Should a repository be open to the public and both store and share 3D data, preserve many types of data streams, or do high-volume data ingest, more staff will be needed to effectively manage the platform and provide customer service.[50] This type of repository requires a more robust software intake and output platform and may need more staff to oversee and maintain the infrastructure, along with dealing with public inquiries.

Software used for 3D modeling and preservation will also affect staffing numbers. For preservation, using an out-of-the-box package such as DSpace versus developing an in-house platform with open-source components will have both short- and long-term staffing ramifications. The use of proprietary software packages generally includes user support and a prescribed repository preservation procedure. There may be many features, some of which you may not need but will still pay for. Training sessions, free or for a fee, may be provided so that the proprietary software is used efficiently. This will require work downtime while staff are involved in training sessions. On the other hand, creating a custom, open-source platform will have higher up-front costs to hire software developers and architects but should result in a finished platform that is custom-tailored to your needs and procedures. Staff will also need to be trained, but at least key personnel can be involved in the development process and generate training procedures as the software is developed.

Training redundancy should be a high priority to ensure there are always individuals capable of operating special equipment, processing complex sets of instructions, or utilizing creative or artistic 3D modeling tools. This duplication ensures that as individuals with these skills leave the organization, there will be others to take their place. Ongoing training and procedure review is essential so as not to interrupt current workflows. Related to this, a detailed set of work procedures should be maintained and be continually reviewed for relevance.

Conclusion

The addition of archiving 3D data does not change the fundamental goals of digital repositories, though the shape and demands of the preserved items have changed. Repository managers are approaching the preservation of 3D data with established preservation practices in mind. Although the specific audience needs of a repository hosting 3D data may differ depending on the institution, some universal challenges emerge across disciplines, including long-term curation decisions (PIPs) and file states and format selection, embargoes, licensing and rights, technology, and storage. In exploring these issues through CS3DP Management Survey Two, the work group aspires to spotlight these challenges and offer recommendations in table 3.20.

TABLE 3.20
Recommendations table

3D Area	Current Practices	Recommendations
Certification	Of the small sampling of five respondents, three have pursued one of these certifications: TRAC or CORE. The other two would like to or intend to.	Good—Aspire toward certification or publicly documented self-audit. Better—Currently undergoing certification or publicly documented self-audit. Best—Achieve certification or have published documented self-audit.
Repository services	Some repositories are able to offer a majority of the recommended services,[a] while others are aware of the need but are currently unable to do so.	Recommended services to offer for 3D data • file integrity and checks • record relation • administrative metadata • access • file migration Good—Working toward offering some of these services Better—Offering some of these services Best—Providing services in all of these areas
Submission packages	Creators—Most save both raw and derivative files. Creator-managers—Most save both raw and derivative files. Repositories—Large majority deliver (allow for download) both raw and derivative files.	Good—Raw data and an access version or image Better—Raw data and full-resolution copies Best—Raw data, full-resolution copies, and access copies
File formats	Creators—Use varies; top three reported OBJ, PLY, and STL. Creator-managers—Also wide variety. Most often used or accepted OBJ, STL, and PLY. Repositories—Most often accepted OBJs, all formats listed, and STL/PLY tied for third place.	Good—A program-neutral format Better—An open-source, program-neutral format that is indexed by PRONOM Best—An open-source, program-neutral format, indexed by PRONOM, that allows for structured and customized embedded metadata + the original format, if applicable

3D Area	Current Practices	Recommendations
Embargoes	Creators—No data were collected on this topic. Repositories/creator-managers—No data were collected on this topic.	Creators • Possess reasonable expectations for embargo period and access. Repositories/creator-managers • Good—Have an embargo period policy. • Better—Criterion from Good recommendation, plus policy for potential submitters is publicized. • Best—Criteria from Better recommendation, plus multiple period options.
Rights and licensing	Creators • Dependent on access/affiliation with a platform, 3rd party, or archiving repository. Repositories/creator-managers • Affirm submitter has rights to data before submission. • Do not require transfer of rights at submission. • State use licenses.	Creators • Thoroughly review all statements before agreeing and uploading their data. Repositories/creator-managers • Continue current trends. • Application of rights and licenses should be dependent on the collection, repository, and dataset.
Systems	Creators and repositories use both open (Fedora, Samvera) and proprietary systems (Preservica, EMu), as well as 3rd-party services (MorphoSource, Dataverse).	Good—A system with documented policies for these recording repository services: • Identifiers • Related objects and context • History of ownership • Fixity (data integrity) checks • Access rights Better—A system that records changes to the above and facilitates the following: • Data ingestion • Data migration • Data publication • Interoperability Best—An open-source system supporting all of the above
Size and storage	Creators—Most use hard drives; some servers and 3rd-party platforms (e.g., Dataverse, MorphoSource). Repositories/creator-managers—Most use servers (no stand-alone hard drives, except among creator-managers).	• Creators' storage capabilities and demands will depend on the system used to house their data. • Repository managers should plan ahead for above-average storage costs due to the large file sizes of 3D data.

3D Area	Current Practices	Recommendations
Collection retention	Creators often store multiple file versions and formats, which greatly increases the size of data in need of preservation. Repositories/creator-managers often prioritize which data to preserve based on institutional policy or may have a mandate to preserve in perpetuity.	Creators • Be aware of format longevity and discipline-specific culling and repository policies. Repositories • Good—Implement a retention schedule to prevent new backlog of orphan works. • Better—Document and implement a retention schedule following relevant discipline-specific approaches. • Best—State a retention schedule and require user agreements up front when creators submit data.
Funding	Creators—They intend to use the following funding sources: • For data creation—grant + institutional support • For data archiving—grant + institutional support and personal investment/self-funding • For long-term planning—institutional support Repositories—They intend to use the following funding sources: • For data archiving—institutional support • For long-term planning—institutional support Some charge for services or storage of data.[b]	Creators • Good—Grants • Better—Grants and institutional support • Best—Institutional support (particularly for archiving one's data) Repositories • Good—Grants and institutional support • Better—Institutional support • Best—Institutional support supplemented by another income stream (endowment, charging for storage or services,[b] etc.)

a. Consultative Committee for Space Data Systems, *Reference Model for an Open Archival Information System (OAIS)*, Recommended Practice, Issue 2, CCSDS 650.0-M-2 (Washington DC: Consultative Committee for Space Data Systems, June 2012), https://public.ccsds.org/pubs/650x0m2.pdf.
b. If charging for services or storage, recommend customized pricing.

APPENDIX 3A
CS3DP 3D Data Management Survey Results

TABLE 3.A.1

Question 1: What kind of 3D data do you produce or host? (Select all that apply.)

Group	3D Laser	BornDigital	CT Scanning	Lidar Scanning	Photo-grammetry	Structured Light	Other
Creators	5	3	6	4	12	4	1
Repository managers	4	13	2	3	9	1	2
Creator-managers	10	6	7	6	21	9	1
Other	2	1	1	2	1	1	1
Totals	21	23	16	15	43	14	5

TABLE 3.A.2

Question 2: How do you define your role in relation to 3D data?

Group	Number of Respondents
Creators	14
Repository managers	20
Creator-managers	17
Other	2
Total	53

Note: These respondents completed at least 25% of the survey

TABLE 3.A.3

Question 3: Repository Institution Type: What type of institution houses/maintains the repository? (Select all that apply.)

Response	Repository Managers	Creator-Managers
Archive	13	10
Commercial organization	1	1
Governmental organization	0	1
Museum	5	7
Nonprofit	2	1
University department/center	4	11
Other	0	1

TABLE 3.A.4

Question 4: Repository Information: In what year did you or your institution begin creating, managing, and/or archiving 3D data?

Group	Before 1990	1990–1994	1995–1999	2000–2004	2005–2009	2010–2014	2015–2019
Creators	0	0	0	0	2	4	10
Repository managers	1	1	0	1	1	5	7
Creator-managers	0	1	1	2	1	4	10
Other	1	0	0	1	0	1	0
Totals	2	2	1	4	4	14	27

TABLE 3.A.5

Question 5: Repository Information: Is your repository 3D-specific?

Group	No	Yes
Repository managers	16	3
Creator-managers	17	4

TABLE 3.A.6

Question 6: Repository Information: Is your repository subject-specific?

Group	No	Yes
Repository managers	10	7
Creator-managers	15	5

TABLE 3.A.7

Question 7: Repository Information: Approximately how many 3D datasets do you have in your collection (where one dataset corresponds to one metadata record)?

Group	0	1–5	6–10	11–25	26–75	76–150	150+
Creators	0	0	2	2	4	0	7
Repository managers	2	5	2	2	0	0	7
Creator-managers	1	3	0	2	1	2	10
Other	0	0	0	0	0	0	1
Totals	3	8	4	6	5	2	25

TABLE 3.A.8

Question 8: Are these datasets shared online or through a repository? (N = 14)

Creator's Response	Count	Percentage
No	4	28.57
Some are	7	50
Yes	3	21.43
Total	14	100%

No comments:

- Under construction, currently stored in an online database.
- Not published yet.

Some are comments:

- Some are in a repository at UCSD.
- One is at Open Topography, the other, still looking for a place to store/share them.
- Some shared on Sketchfab, Internet Archive, or own website.
- Some are shared on MorphoSource.
- Different curators and different datasets are shared differently. Some datasets are available on cloud drives; others only on physical drives.
- Some are available on Sketchfab, but most are for internal use only

Yes comments:

- Through a website (see https://sites.lib.jmu.edu/mac3d/3d-models/).
- Datasets containing published data are on MorphoSource.

TABLE 3.A.9

Question 9: User: In what areas does your 3D data meet user needs? (Select all that apply.)

Group/Data Use	Creator	Repository Manager	Both	Other	Totals
Architectural drawings	3	3	2	1	9
3D printing	5	5	14	1	25
Art	4	3	8	0	15
Augmented reality	3	0	9	1	13
Commercial	0	1	1	0	2
Conservation/preservation	7	4	12	1	24
Education (higher ed.)	6	8	13	1	28
Education (K–12)	1	3	6	0	10
General public	4	10	11	1	26
Government/policy	2	1	1	0	4
Recreational/hobbyist	3	3	5	0	11
Research	13	9	15	1	38
Tech. documentation (e.g., engineer)	3	3	2	1	9
Virtual reality	4	7	12	1	24

Group/Data Use	Creator	Repository Manager	Both	Other	Totals
Other	2	1	1	0	4

Other comments:
- Museum preps and exhibition support
- Campus museum, demonstration
- Noncommercial applications

TABLE 3.A.10

Question 10: Cost Considerations: Does your repository charge to upload 3D data?

Group	No	Sometimes	Yes
Repository managers	14	1	2
Creator-managers	15	2	0
Totals	29	3	2

TABLE 3.A.11

Question 11: Cost Considerations: What is your pricing model? (N = 3)

Group	Response	*N*
Repository manager	Basic charge is $5/file with 10 MB allotted for each file.	1
Creator-manager	[Price] depends if we deposit with UK Archaeology Data Service as well. We use the ADS calculator.	1
Creator-manager	[Pricing varies for] internal v. external 3D files, also depending on funding source.	1

TABLE 3.A.12

Question 12: Cost Considerations: How is your data archiving funded? (Select all that apply.)

Group	Grants	Institutional Support	Personal Investment/ Self-funded	Other
Creators (*n* = 14)	5	8	5	1
Repository managers (*n* = 14)	6	13	2	0
Creator-managers (*n* = 15)	4	13	3	1
Other	0	0	0	1

Other comments:

Creator
- We have not started archiving 3D data. We disseminate through a web CMS.

Creator-Managers
- Sometimes included on grant funding, but not standard.

Other
- Unfunded.

TABLE 3.A.13

Question 13: Cost Considerations: How is your data creation funded? (Select all that apply.)

Group	Grants	Institutional Support	Personal Inv./ Self-funded	Other
Creators	9	11	2	0
Repository managers	8	9	1	0
Creator-managers	10	11	2	2
Other	1	0	1	0

Other comments:
- Sometimes included on grant funding, but not standard.
- Industry partnerships.

TABLE 3.A.14

Question 14: Financial Model: How do you plan for long-term costs? (Select all that apply.)

Group	Charge at Cost (Data Contributor Pays)	Charge at Cost (Data Downloader Pays)	Institution Support	Endowment	Other
Creators (*N* = 14)	1	0	10	1	4
Repository managers (*N* = 14)	3	0	12	2	2
Creator-managers (*N* = 15)	1	1	12	2	2
Other (*N* = 1)	0	0	0	0	1

Other comments:

Creator
- No plan!
- Data creation is charged at-cost, archiving is not.
- No long-term plan (I do not own/oversee facility).

Repository Manager
- University overhead costs—eventually I hope.

Creator-Manager
- Not sure.
- We host our data publicly on GitHub so it is free forever (we hope!). If not, we can easily move it to another git repository.

Other
- No planning yet.

TABLE 3.A.15

Question 15: Staffing: How many persons/FTE (full-time equivalent) do you employ to maintain the repository and its functions?

Group	<1	1–3	4–9	10+
Creators	3	2	4	5
Repository managers	3	7	4	0
Creator-managers	7	4	3	1
Other	0	0	0	1

TABLE 3.A.16

Question 16: Staffing: How many people create 3D in your organization/department?

Group	<1	1–3	4–9	10+
Creators	3	2	4	5
Repository managers	4	6	1	2
Creator-managers	1	8	3	2
Other	0	0	0	1

TABLE 3.A.17

Question 17: Preservation Platform: What system(s) are you using to store/manage your data?

Group	3rd-Party Vendor	Server Setup	Hard Drives	Open-Source Option	Software Packages	Other
Creators	4	6	10	1	1	2
Repository managers	4	6	0	4	3	1
Creator-managers	5	9	5	1	0	2
Other	0	1	0	0	0	0

Comments:
Hard Drives

- Network Attached Storage
- External hard drives
- Computers where photogrammetry and display occur
- NAZ
- External, internal
- Time Capsule
- Two synced external hard drives
- Personal hard drives
- xyz
- Buy large externals at Costco

- Various external HDs
- Hard drives
- Redundant Backups on External Hard Drives
- N/A
- On central PC

Server setup

- Network Attached Storage
- Internal servers
- AWS
- Linux; Python venvs; U of I–hosted cPanel VM for scripts to process, archive, and upload 3d models
- Institutional servers and services such as Drive and Box
- Linux
- Apache2
- Stored on research data repository
- AWS
- Cloud storage through the institution
- N/A
- We've currently deployed a hybrid-cloud (VPC) enterprise NAS solution to store, retain, and back up our data. There are also additional servers to handle the web services.

Software package(s)

- Python
- Islandora
- postgreSQL database
- EMu—to manage 3D data in parallel to its corresponding specimen/object data. R-scripts to check/compare files.

3D-party vendor (e.g., Dataverse, MorphoSource, etc.)

- Figshare
- U of I Box instance for archive storage
- Sketchfab
- Sketchfab, Internet Archive
- Sketchfab
- Dataverse
- Dropbox and Sketchfab
- Sketchfab.com (http://sketchfab.com/)
- Constructed 3D models and corresponding processed image stacks from CT scans are uploaded to/accessible online via MorphoSource

Open-source option (Fedora/Samvera)

- Fedora
- Islandora
- Archivematica
- Online access to collections is maintained through a Drupal site. We plan to transition to the Atlas of Living Australia (ala.org.au) web framework.

Other

- We use Preservica for flat items and hope to use it eventually for 3D objects.
- Preservica
- GitHub, Zenodo, GitLab
- GitHub
- Drobo

TABLE 3.A.18

Question 18: Data Package Structure: What items are contained in the Delivery Information Packages (the downloadable files) offered by your repository? (Select all that apply.)

Group	Both Raw and Derivative Files	The Derivatives	The Derivatives or Final Versions	Raw Data Files	Other
Repository managers	8	1	1	2	1
Creator/ managers	5	4	4	0	1

Other comments:

- Whatever files and information are deposited.
- Raw files are kept but not distributed.
- Whatever the researcher wants to keep. Sometimes only the final versions, sometimes the raw and final versions.

TABLE 3.A.19

Question 19: [Creator] Saved Files: What files do you preserve? (Select all that apply.)

Group	Both Raw and Derivative Files	The Derivatives	The Derivatives or Final Versions	Raw Data Files	But Keep Raw Files Offline	Other
Creators	9	2	3	3	2	1

Other comments:

- We are not preserving our 3D data at this point. Option 3 [The derivatives but keep raw files offline] is likeliest.

TABLE 3.A.20

Question 20: File Naming: What file naming conventions/systems do you use? [multiple selection] (N = 14)

Group	Date	Text-Based Title of Model	Unique Identifier	Other
Creators	2	8	8	2
Creator-managers	2	3	7	4

Other Comments:

Creators

- Taxon and specimen number
- Specimen number and sex

Creator-Managers

- Holding Unit, Collection #, part description, etc.
- The file naming used by the researcher if the 3D scanning was not done in house. Otherwise, we use barcoding if it's related to a physical artifact.

- Depends on deposit place
- Simple human-readable file/folder naming scheme
- Spreadsheet of all scans (includes parameters and specimen info), with PDF of entry kept in dataset folder

Example comments:
- PDF sent to [a survey administrator]
- https://github.com/nomadproject/objects/tree/gh-pages/collection

TABLE 3.A.21

Question 21: File Formats [for Creators]: What file formats do you use? (Select all that apply.) (N = 14)

File Type	Creators
3DS	2
BLEN	1
DXF	1
E57	1
geo Tiff	3
LAS	2
LAZ	1
NRRD	1
OBJ	11
PLY	10
PTS	1
STL	6
VOL	1
WRL_VRML	1
XYZ	4

TABLE 3.A.22

Question 22: File Formats: Do you have specific format requirements?

Group	No	Yes
Repository managers	7	6
Creator-managers	8	7

TABLE 3.A.23

Question 23: File Formats [for Repository Managers and Creator-Managers]: What file formats does your repository accept? (Select all that apply.)

File Type	Repository Managers (*N* = 11)	Creator-Managers (*N* = 6)
glTF	1	2
OBJ	5	4
PLY	3	3
PTS	0	1
PTX	0	1
STL	3	4
WRL-VRML	0	1
X3D	0	1
ZLT	0	1
Other	3	1
All of the Above[a]	4	0

Other comments:
Repository Manager
- We have put no limitations on 3D formats yet, but have limits for other file types
- .zip
- fbx

Creator-Manager
- CTM

a. *All of the Above* includes 3DS, 3MF, BLEN, DAE, DXF, geo TIFF, glTF, LWO, OBJ, OFF, PLY, PTS, PTX, SC1, SCL, SKP, STL, TRI, V3D, WRL - VRML, X3D, X3DV, XSI, ZTL, XYZ

TABLE 3.A.24

Question 24: File Formats: Describe your familiarity with ISO standard X3D file format. (N = 42)

Type of Familiarity	Count
Familiar, and use the file format	3
Familiar, but do not use	18
Unfamiliar, do not use	21

TABLE 3.A.25

Question 25: File Sizes: Do you have a size limit for ingest of data packages? (N = 25)

Response	Count
No	18
Yes	7

TABLE 3.A.26

Question 26: File Sizes: What are the size limits for ingest of data packages? (Select all that apply.) (N = 6)

Options	Count
>2 GB	1
2–5 GB	0
5–10 GB	2
10–20 GB	0
Other	3

Other comments:
- 50 MB
- 1 TB (may be larger)
- <10MB

TABLE 3.A.27

Question 27: File Sizes: What is/are the limiting factor(s)? [selected choice] (N = 3)

Options	Count
Software	1
Browser[a]	1
Other[b]	1

a. GUI uploading/timing out
b. Repository

TABLE 3.A.28

Question 28: File Sizes: What do you consider to be a large data file size in your current repository?

Files Sizes	Count
<4 GB	5
5–20 GB	3
21–499 GB	3
500 GB	2
1 TB	2
N/A	1

Responses
- Don't know
- 4 gb
- Per file, anything over 20 GB has to be uploaded a different way; per package, anything approaching more than 500 GB would give me pause. We can support a couple of packages of that size but not too many atm.

- 150 MB
- 10 gb
- 1 GB
- 100 gb
- 500 MB
- 1 GB
- 0.5TB
- 1 tb
- 50 GB per ingest (which can contain very many files); 3–5 GB per file
- 20 GB +
- 5 GB
- 1TB
- 10 GB

TABLE 3.A.29

Question 29: Metadata Quality: Do you require or capture anything beyond basic metadata (e.g., title, ID, date created, creator, etc.)?

Group	Yes	No	*Yes* Descriptions
Creators (*n* = 14)	6	8	• Object description • There is a whole database behind the 3D data bringing in everything from the archaeological excavation, e.g., timestamp, creator, but not yet designed explicitly for metadata. • Standard .pca file data is archived. • Info about the specimen, who it's for, collection it came from, etc. Kept in spreadsheet. • We record all scan data for CT scans. • Scan settings, other data related to specimen/organism scanned (specimen number, museum/herbarium, etc.)
Repository managers (*n* = 9)	6	3	• Software requirements; hardware requirements (if the submitter knows them); and information on what is in each file (if possible). • We use a custom standard. • We have a MODS application profile that we use. • Date created, author, size, & image features. • Archivematica captures PREMIS metadata. • Metadata fields can be viewed at https://tdar-arch.atlassian.net/wiki/spaces/TDAR/pages/557072/Data+Dictionary#DataDictionary-SensoryData-Fields. • Associated object/specimen occurrence ID, signed media agreement with preferred attribution/credit line; README TXT description of technical details for photogrammetry; Scan Sheet PDF of scanner settings for CT scans.

Group	Yes	No	*Yes* Descriptions
Creator-managers ($n = 13$)	8	5	• s.si.edu/2PDZFEY • Subject matter data for works of art. • Reach dialogue tech manual expands on this. • Technical administrative and descriptive metadata. • License. • Specimen information from museum catalog. • Stakeholders, location of shoot, rig/setup, equipment used, etc. I use the CHI Digital Lab Notebook.

TABLE 3.A.30

Question 30: Copyright/Licensing: Do you require depositors to affirm their ownership of the submitted data?

Group	Yes	No
Repository managers ($n = 12$)	11	1
Creator-managers ($n = 14$)	6	8

TABLE 3.A.31

Question 31: Copyright/Licensing: Are creators/submitters required to transfer copyright (where applicable) to the repository?

Group	Yes	No
Repository managers ($n = 11$)	1	10
Creator-managers ($n = 6$)	0	6

TABLE 3.A.32

Question 32: Copyright/Licensing: Do you ask them to choose a license (e.g., Creative Commons)?

Group	Yes	No
Repository manager ($n = 1$)	0	1

TABLE 3.A.33

Question 33: Licensing/Usage Permitted: Are licenses (permitted uses) specified by your repository?

Group	Yes	No	Other	*Yes/Other* Comments
Repository managers ($n = 11$)	5	3	3	Yes • We require a Creative Commons license. • https://creativecommons.org/ • https://www.tdar.org/about/policies/contributors-agreement/ • https://www.fieldmuseum.org/field-museum-natural-history-conditions-and-suggested-norms-use-collections-data-and-images

Group	Yes	No	Other	*Yes/Other* Comments
				Other • Depends on license assigned to the dataset. • We use a boilerplate licensing agreement at upload. • Embedded in the metadata but not shown in the repository.
Creator-managers ($n = 14$)	6	4	4	Yes • https://umorf.ummp.lsa.umich.edu/wp/terms/ • Can't find link just now. • https://creativecommons.org/licenses/by-nc-sa/4.0/ • http://diginole.lib.fsu.edu/repository/copy-right-and-legal-agreements • http://diginole.lib.fsu.edu/repository/copy-right-and-legal-agreements • https://umorf.ummp.lsa.umich.edu/wp/regis-ter/ Other • We recommend Creative Commons. • All data in our repository is institution-collected data (or contractor -collected where rights are not retained by the contractor). Data of institu-tion items falls under SI terms of use (https://www.si.edu/termsofuse), other data is on a case-by-case basis. • We use Sketchfab as a public repository but do not engage in their licensing options. • These vary. Licenses for external objects are managed by the object owners.

TABLE 3.A.34

Question 34: Licensing: Do you require users downloading datasets to acknowledge these licenses? (N = 9)

Group	Yes	No
Repository managers and creator-managers	4	5

TABLE 3.A.35

Question 35: Policy: What is your policy for accessing data? (N = 14)

Policy Type	Count
Closed access	1
No policy	1
Open access	9

Policy Type	Count
Varied access	3

Open Access Responses:

- Can't find link just now, behind firewall.
- https://creativecommons.org/licenses/by-nc-sa/4.0/
- http://diginole.lib.fsu.edu/repository/copyright-and-legal-agreements

Varied Access Responses:

- Data is made open and accessible under SI terms of use (https://www.si.edu/termsofuse) and the only restrictions placed on access are for copyright, cultural sensitivities, etc. For the complete policy see SD-609 (https://www.si.edu/content/pdf/about/sd/SD609.pdf).
- Our institution supports open access, but where the case requires the data to be closed, we will maintain private research access only.
- CC-NC-BY for internal work. External varies depending on permissions specified by copyright holders.

TABLE 3.A.36

Question 36: Planning: What measures have you taken to ensure your 3D data's integrity? (Select all that apply.) (N = 24)

Planning Measure	Response Count
A risk management plan	9
Catastrophic migration plan	5
Documentation of repository infrastructure (back-end architecture, staff, roles, etc.)	14
Require/encourage checksums upon submission	9
Complete checksums over time	11
Utilize version histories	5
Other (please explain below)	6

Other Responses:

- As our current "repository" is simply a folder structure on network attached storage, we do not do any file fixity checks, but we do have the entire system backed up to redundant storage and also to tape.
- Open and check the files post upload [*sic*] when possible; not always possible due to proprietary [*sic*] software.
- Bagit workflow with checksum and manifest.
- As we use GitHub, our data is versioned, meaning we can roll back if necessary.
- As per existing digital repository.
- Institutional informatic services.

TABLE 3.A.37

Question 37: Certification: What, if any, trusted digital repository certifications has your repository pursued? (Select all that apply.) (N = 5)

Certifications	Response Count
TRAC (Trusted Repositories Audit and Certification)	1
Core (CoreTrustSeal)	2
Nestor (Network of Expertise in Long-Term Storage of Digital Resources)	0
Other (please explain below)	2

Other Responses:
- Would like to do Core
- Aiming for TRAC in long-term plan.

TABLE 3.A.38

Question 38: 3D Planning: Do you see particular risks related to 3D files that need to be taken into account as part of risk management planning?

Risks	Count
Long-term preservation	4
File formats	3
File submission state	2
Lack of guidelines	2
3D buzz factor	1

Responses:
- Biggest issue would be with proprietary file format for raw or derivative data. While this is unavoidable for some scanning technologies, photogrammetry and spherical laser scans both use (or can use) open, durable formats. That's not to say they should be preferred, but that the durability of the data should be taken into consideration when selecting a capture technology. For 3D models derived from scan data, as long as those files are saved as .ply, .obj, or .x3d (depending on your needs), you should be good.
- Software and hardware obsolesce and access restrictions due to proprietary file formats.
- Thus far, most of our risk assessments have focused on rights issues, though we acknowledge problems related to the complexity and size of the files.
- I think it's important to save the raw data for potential reprocessing as software improves.
- Currently, the library is not actively involved in determining the nature of the data, and will take submitted data. Because of this, we do not have strict control over versioning, naming conventions, the dataset, etc. There is little control and knowledge about 3D so there is little knowledgeable support for these data. There is a risk of loss.
- Interoperability, bus [*sic*] factor for creators and owners.
- Long-term rendering for discovery and use. File format obsolescence.
- Lack of clear guidelines for digitization and preservation CH 3D models.

TABLE 3.A.39

Question 39: Preservation: What other preservation methods/standards might your repository use? (Select all that apply.) (N = 18)

Method/Standards	Count
Migration	15
As-is **bitstream preservation**	10
Emulation	6

TABLE 3.A.40

Question 40: Culling Considerations: What affects preservation priority (e.g., value of projection data vs. raw, or other derivatives)? (Select all that apply.) (N = 21)

Priority	Count
Institutional/departmental policy	15
Rarity of data	7
Disciplinary factors	6
Legal terms	6
Data are kept forever	5
Submitter's desire	4
Other	4
Funder	1

Responses:

- Raw data and high-resolution derivatives which get assigned a GUID are kept forever, derivatives for web consumption, etc. are kept as long as they are needed.
- Risk of loss.
- How often these data are used for teaching/learning.

TABLE 3.A.41

Question 41: Indexing: Is your repository open to indexing by external search engines?

Response Choice	Response Count
Yes	12
No (Please explain why below.)	9

No responses:

- Not public.
- Our private repository is not. I assume Sketchfab is indexed.
- There are three sides of storage: one is accessible by Picture library staff only, second is extended to curatorial colleagues [*sic*] and third is public-facing.

TABLE 3.A.42

Question 42: Data Preservation: What measures do you take to ensure your data are not lost? (Select all that apply.)

Measure	Count
Backups	14
Fixity checks	1
Virus scanning	2
Other	

Responses:

- Digital preservation system in progress.
- Multiple copies to protect against hard drive failure.
- Rely on virtual server provider for security measures.

TABLE 3.A.43

Question 43: [Creators] What challenges do you find in managing the preservation of 3D data? (N = 13)

Creator Challenges Themes Summarized	Count
Standards	1
Workflows	1
Size of data/storage space	9
Institutional support (infrastructure)	3
Just getting started	1
Data types (raw, derivative, etc.)	1
Formats (proprietary, etc.)	3
Metadata	2
Costs	2
Interoperability	1
Naming	1
Documentation accuracy	1
Access	1

Responses:

- Standards, workflows, size of data.
- We are not yet situated institutionally to preserve 3D data, but we are aware that we need to get situated ASAP.
- Just getting started so don't know all of the challenges.
- Deciding which raw, intermediate, and final assets to preserve. Formats. Disk space needed. Metadata challenges.
- Storage space, costs.

- Preservation is straightforward, convincing "non-digital" managers and obtaining the infrastructure is the biggest challenge, even if the costs are minimal (e.g., open-source software).
- File formats + metadata, interoperability, large files.
- Large datasets, proprietary software/file formats.
- There is no institutional support for affiliates in archiving data, and private cloud costs are prohibitive. Therefore, stuck doing multiple copies on dodgy external drives.
- Space. Naming iterations meaningfully. Multiple users have access to spreadsheet, and therefore there are multiple opportunities for user error, especially with inexperienced scanners.
- File size and access are difficult.
- Backing up/preserving large amounts of data (10s–100s of TB). There are no good or efficient solutions currently available to us.
- Storage space.

TABLE 3.A.44

Question 44: [Creator-Managers] What challenges do you find in managing the preservation of 3D data?

Both Challenges Themes Summarized	Count
Documentation	1
Associating parts of a project (data, models, records)	1
Annotation for models/collections	1
Formats	2
Volume/size/storage	3
Funding	1
Staffing	1
New to it	2
Lack of learning opportunities	1
User experience	1
Unknown standards (equipment, best practices)	2
Metadata	1

Responses:

- Fully documenting captures; associating captured data with models; associating captured data and models with collection records; annotation models and linking annotations back to collection records; this is all very complex, and not standardized. Yikes!
- Size of datasets and rapidly developing tech environment.
- There are no written standards so there is little buy-in institutionally to take 3D preservation seriously. We are waiting until someone else figures it out.
- Same as other imaging formats—volume and funding.
- We are ensuring that we maintain an Amazon S3 copy of our raw data and two separate offline hard drive copies. We would prefer to use versioned git repositories for this, but the large size of our original photogrammetry files means this isn't an option at the moment. We looked at GitHub LFS, but using https://datproject.org is more appealing in that we can self-host multiple versioned copies of these files.

- Staff resources for developing necessary infrastructure.
- New to this body of work …opportunities to learn from experienced institutions… improving the user experience …maintaining a viable user experience over time.
- We are just getting started, but storage space could be an issue.
- Equipment standards and proprietary file formats are often seen as the lay standard, meaning data is submitted in less-accessible format.
- Storage limits, meta and para data standards are lacking, I'm ignorant when it comes to best practices of 3D data storage.

TABLE 3.A.45

Question 45: [Managers] What challenges do you find in managing the preservation of 3D data?

Repository Challenges Themes Summarized	Count
Variety	2
Scale	1
Size/storage	4
Metadata	1
File formats	2
File organization	1
Standard for platforms	1
Rights management	1
Access (high res, original environment)	1
Cost	2

Responses:

- Variability, scale, and size.
- The metadata standard for VR objects is tricky; multiple files constitute a single object; file formats.
- (1) The size/complexity of the data, (2) the lack of best practices for existing platforms, e.g., Islandora, (3) rights management.
- Being able to provide access to a high-fidelity version of the file in its original (or near-original) environment. (We're looking forward to emulation projects, like EaaSI.)
- The size of files and multiple file formats.
- Cost.
- Time, staff, and storage resources for preservation are rarely included in projects/grant funding.

APPENDIX 3B
CS3DP Management: 3D Repositories Survey 2

Description

Q0 Purpose

The Community Standards for 3D Data Preservation Management and Storage Workgroup is collecting information about repositories that host and self-managing creators of 3D data of all kinds. This includes, but is not limited to: born-digital, CT scanning, laser/lidar/white-light/structured-light scanning, photogrammetry, etc.

Audience

The below survey is intended for repository managers as well as individual creators who perform their own personal data management.

Anonymity

This is an informational survey and no personal data need be submitted. Respondents have the option of submitting anonymously. The name of your repository/project/organization will be requested at the end of the survey. While this information is very much appreciated, it is not required. There will also be an option to include contact information if you are willing to answer clarification questions, should there be any.

Composition

The survey covers the following sections:

- Anonymous Basic Information
- Repository Infrastructure
- Usage and Permissions
- Risk Management
- Data Management
- Non-anonymous Basic Information

Anonymous Basic Information

Question 1

Q1 What kind of 3D data do you produce or host? (Select all that apply.)

- ☐ Photogrammetry
- ☐ 3D Laser
- ☐ Lidar scanning
- ☐ Structured Light

- ▢ CT Scanning
- ▢ Born-digital
- ▢ Other (please explain below):

Question 2

Q2 How do you define your role in relation to 3D data?

- o Creator
- o Repository manager
- o Both a creator and a repository manager
- o Other (please explain below):

Display Question #3

If How do you define your role in relation to 3D data? = Repository manager
Or How do you define your role in relation to 3D data? = Both a creator and a repository manager

Question 3

Q4 Repository Institution Type: What type of institution houses/maintains the repository? (Select all that apply.)

- ▢ Archive/library
- ▢ Commercial organization
- ▢ Governmental organization
- ▢ Museum
- ▢ Nonprofit
- ▢ University department/center
- ▢ Other (please explain below):

Question 4

Q6 Repository Information: In what year did you or your institution begin creating, managing, and/or archiving 3D data?

- o Before 1990
- o 1990–1994
- o 1995–1999
- o 2000–2004
- o 2005–2009
- o 2010–2014
- o 2015–2019

Display Question #5

If How do you define your role in relation to 3D data? = Repository manager
Or How do you define your role in relation to 3D data? = Both a creator and a repository manager

Question 5

Q7 Repository Information: Is your repository 3D-specific?

- o Yes
- o No
- o Other (please explain below):

Display Question #6

If How do you define your role in relation to 3D data? = Repository manager
Or How do you define your role in relation to 3D data? = Both a creator and a repository manager

Question 6

Q8 Repository Information: Is your repository subject-specific?

- o Yes. Please list the subjects.
- o No

Question 7

Q10 Repository Information: Approximately how many 3D datasets do you have in your collection? (Where one dataset corresponds to one metadata record)

- o 0
- o 1–5
- o 6–10
- o 11–25
- o 26–75
- o 76–150
- o 151+

Display Question #8

If How do you define your role in relation to 3D data? = Creator

Question 8

Q11 Access: Are these datasets shared online or through a repository?

- o Yes. Please explain below:

- o Some are. Please explain below:
- o No. Please explain below:

Question 9

Q12 User: In what areas does your 3D data meet user needs? (Select all that apply.)

- ☐ 3D Printing
- ☐ Art
- ☐ Augmented Reality
- ☐ Commercial
- ☐ Conservation/Preservation
- ☐ Education (Higher Ed.)
- ☐ Education (K–12)
- ☐ General Public
- ☐ Government/Policy
- ☐ Recreational/Hobbyist
- ☐ Research
- ☐ Technical Documentation (e.g., engineering, architectural drawings)
- ☐ Virtual Reality
- ☐ Other (please explain below):

Repository Infrastructure

Display Question #10

If How do you define your role in relation to 3D data? = Repository manager
Or How do you define your role in relation to 3D data? = Both a creator and a repository manager

Question 10

Q13 Cost Considerations: Does your repository charge to upload 3D data?

- o Yes
- o No
- o Sometimes

Display Question #11

If Cost Considerations: Does your repository charge to upload 3D data? = Yes
Or Cost Considerations: Does your repository charge to upload 3D data? = Sometimes

Question 11

Q14 Cost Considerations: What is your pricing model?

Question 12

Q15 Cost Considerations: How is your data archiving funded? (Select all that apply)

- ▢ Grants
- ▢ Institutional Support
- ▢ Personal Investment/Self-funded
- ▢ Other (please explain below):

Question 13

Q16 Cost Considerations: How is your data creation funded? (Select all that apply)

- ▢ Grants
- ▢ Institutional Support
- ▢ Personal Investment/Self-funded
- ▢ Other (please explain below):

Question 14

Q17 Financial Model: How do you plan for long-term costs? (Select all that apply)

- ▢ Institutional support
- ▢ Endowment
- ▢ Charge at-cost (data-contributor pays)
- ▢ Charge at-cost (user pays)
- ▢ Other (please explain below):

Display Question #15

If How do you define your role in relation to 3D data? = Repository manager
Or How do you define your role in relation to 3D data? = Both a creator and a repository manager

Question 15

Q18 Staffing: How many persons/FTE (full-time equivalent) do you employ to maintain the repository & its functions?

- o <1
- o 1–3
- o 4–9
- o 10+

Question 16

Q19 Staffing: How many persons create 3D data in your organization/department?

- o <1
- o 1–3

- o 4–9
- o 10+

Question 17

Q20 Preservation Platform: What system(s) are you using to store/manage your data?

- ▢ Hard drives
- ▢ Server setup
- ▢ Software package/s
- ▢ 3d party vendor (e.g.) Dataverse, MorphoSource, etc.)
- ▢ Open source option (Fedora/Samvera)
- ▢ Other (please explain below):

Display Question #18

If How do you define your role in relation to 3D data? = Repository manager
Or How do you define your role in relation to 3D data? = Both a creator and a repository manager

Question 18

Q22 Data Package Structure: What items are contained in the Delivery Information Packages (the downloadable files) offered by your repository? (select all that apply)

- o The raw data files
- o The derivatives or final versions
- o The derivatives, but keep raw files offline
- o Both raw and derivative files
- o Other (please explain below):

Display Question #19

If How do you define your role in relation to 3D data? = Creator

Question 19

Q23 [Creator] Saved Files: What files do you preserve? (select all that apply)

- ▢ The raw data files
- ▢ The derivatives or final versions
- ▢ The derivatives, but keep raw files offline
- ▢ Both raw and derivative files
- ▢ Other (please explain below):

Display Question #20

If How do you define your role in relation to 3D data? = Creator

Or How do you define your role in relation to 3D data? = Both a creator and a repository manager

Question 20

Q54 File Naming: What file naming conventions/systems do you use?

- ▢ Unique Identifier
- ▢ Date
- ▢ Location
- ▢ Text-based Title of Model
- ▢ Other (please describe below):
- ▢ Example:

Display Question #21

If How do you define your role in relation to 3D data? = Creator

Question 21

Q24 File Formats: What file formats do you use? (Select all that apply)

- ▢ 3DS
- ▢ 3MF
- ▢ BLEN
- ▢ DAE
- ▢ DXF
- ▢ geo TIFF
- ▢ gITF
- ▢ LWO
- ▢ OBJ
- ▢ OFF
- ▢ PLY
- ▢ PTS
- ▢ PTX
- ▢ SC1
- ▢ SCL
- ▢ SKP
- ▢ STL
- ▢ TRI
- ▢ V3D
- ▢ WRL - VRML
- ▢ X3D
- ▢ X3DV
- ▢ XSI

- ▢ ZTL
- ▢ XYZ
- ▢ Other

Display Question #22

If How do you define your role in relation to 3D data? = Repository manager
Or How do you define your role in relation to 3D data? = Both a creator and a repository manager

Question 22

Q25 File Formats: Do you have specific format requirements? (see examples here).

- o Yes
- o No

Display Question #23

If How do you define your role in relation to 3D data? = Repository manager
Or How do you define your role in relation to 3D data? = Both a creator and a repository manager
And File Formats: Do you have specific format requirements? (see examples here). = Yes

Question 23

Q78 File Formats: What file formats does your repository accept? (Select all that apply)

- ▢ 3DS
- ▢ 3MF
- ▢ BLEN
- ▢ DAE
- ▢ DXF
- ▢ geo TIFF
- ▢ gITF
- ▢ LWO
- ▢ OBJ
- ▢ OFF
- ▢ PLY
- ▢ PTS
- ▢ PTX
- ▢ SC1
- ▢ SCL
- ▢ SKP

- ▢ STL
- ▢ TRI
- ▢ V3D
- ▢ WRL - VRML
- ▢ X3D
- ▢ X3DV
- ▢ XSI
- ▢ ZTL
- ▢ XYZ
- ▢ Other
- ▢ All of the above

Question 24

Q27 File Formats: Describe your familiarity with ISO standard X3D file format?

- o Familiar, but do not use
- o Familiar, and use the file format
- o Unfamiliar, do not use

Display Question #25

If How do you define your role in relation to 3D data? = Repository manager
Or How do you define your role in relation to 3D data? = Both a creator and a repository manager

Question 25

Q28 File Sizes: Do you have a size limit for ingest of data packages?

- o Yes
- o No

Display Question #26

If File Sizes: Do you have a size limit for ingest of data packages? = Yes

Question 26

Q29 File Sizes: What are the size limits for ingest of data packages? (select all that apply)

- ▢ >2GB
- ▢ 2–5GB
- ▢ 5–10GB
- ▢ 10–20GB
- ▢ Other (please explain below):

Display Question #27

If File Sizes: What are the size limits for ingest of data packages? (select all that apply) = >2GB
Or File Sizes: What are the size limits for ingest of data packages? (select all that apply) = 2–5GB
Or File Sizes: What are the size limits for ingest of data packages? (select all that apply) = 5–10GB
Or File Sizes: What are the size limits for ingest of data packages? (select all that apply) = 10–20GB

Question 27

Q30 File Sizes: What is/are the limiting factor(s)?

- ☐ Software
- ☐ Browser
- ☐ Other (please explain below):

Display Question #28

If File Sizes: Do you have a size limit for ingest of data packages? = No

Question 28

Q31 File Sizes: What do you consider to be a large data file size in your current repository?

Display Question #29

If How do you define your role in relation to 3D data? = Creator
Or How do you define your role in relation to 3D data? = Repository manager
Or How do you define your role in relation to 3D data? = Both a creator and a repository manager

Question 29

Q32 Metadata Quality: Do you require or capture anything beyond basic metadata (e.g., Title, ID, Date Created, Creator, etc.)?

- o Yes. Please describe below:
- o No

Usage and Permissions

Display Question #30

If How do you define your role in relation to 3D data? = Repository manager

Or How do you define your role in relation to 3D data? = Both a creator and a repository manager

Question 30

Q38 Copyright/Licensing: Do you require depositors to affirm their ownership of the submitted data?

- o Yes
- o No

Display Question #31

If Copyright/Licensing: Do you require depositors to affirm their ownership of the submitted data? = Yes

Question 31

Q39 Copyright/Licensing: Are creators/submitters required to transfer copyright (where applicable) to the repository?

- o Yes
- o No

Display Question #32

If Copyright/Licensing: Are creators/submitters required to transfer copyright (where applicable) to… = Yes

Question 32

Q40 Copyright/Licensing: Do you ask them to choose a license (e.g., Creative Commons)?

- o Yes. If so, which licenses are available:
- o No

Display Question #33

If How do you define your role in relation to 3D data? = Repository manager
Or How do you define your role in relation to 3D data? = Both a creator and a repository manager

Question 33

Q41 Licensing/Usage Permitted: Are licenses (permitted uses) specified by your repository?

- o Yes. Please provide a link:

- o No
- o Other (please explain below):

Display Question #34

If Licensing/Usage Permitted: Are licenses (permitted uses) specified by your repository? = Yes. Please provide a link:

Question 34

Q42 Licensing: Do you require users downloading datasets to acknowledge these licenses?

- o Yes
- o No

Display Question #35

If How do you define your role in relation to 3D data? = Repository manager
Or How do you define your role in relation to 3D data? = Both a creator and a repository manager

Question 35

Q43 Policy: What is your policy for accessing data?

- o Open Access. Please link to your policy:
- o Closed Access
- o Varied Access. Please explain and/or provide a link:
- o No Policy

Risk Management

Display Question #36

If How do you define your role in relation to 3D data? = Repository manager
Or How do you define your role in relation to 3D data? = Both a creator and a repository manager

Question 36

Q47 Planning: What measures have you taken to ensure your 3D data's integrity? (Select all that apply)

- ▢ A risk management plan
- ▢ Catastrophic migration plan
- ▢ Documentation of repository infrastructure (backend architecture, staff, roles, etc.)

- ▢ Require/encourage checksums upon submission
- ▢ Complete checksums over time
- ▢ Utilize version histories
- ▢ Other (please explain below):

Display Question #37

If How do you define your role in relation to 3D data? = Repository manager
Or How do you define your role in relation to 3D data? = Both a creator and a repository manager

Question 37

Q49 Certification: What, if any trusted digital repository certifications has your repository pursued? (Select all that apply)

- ▢ TRAC (Trusted Repositories Audit & Certification)
- ▢ Core (CoreTrustSeal)
- ▢ Nestor (Network of Expertise in Long-Term Storage of Digital Resources)
- ▢ Other (please explain below):

Display Question #38

If How do you define your role in relation to 3D data? = Repository manager
Or How do you define your role in relation to 3D data? = Both a creator and a repository manager

Question 38

Q86 3D Planning: Do you see particular risks related to 3D files that need to be taken into account as part of risk management planning?

Data Management

Display Question #39

If How do you define your role in relation to 3D data? = Repository manager
Or How do you define your role in relation to 3D data? = Both a creator and a repository manager

Question 39

Q51 Preservation: What other preservation methods/standards might your repository use? (Select all that apply)

- ▢ Emulation
- ▢ Migration

- ☐ As-is Bitstream Preservation
- ☐ Other (please explain below):

Display Question #40

If How do you define your role in relation to 3D data? = Repository manager
Or How do you define your role in relation to 3D data? = Both a creator and a repository manager

Question 40

Q52 Culling Considerations: What affects preservation priority? (e.g., value of projection data vs raw, or other derivatives) (Select all that apply)

- ☐ Institutional/departmental policy
- ☐ Funder
- ☐ Legal terms
- ☐ Submitter's desire
- ☐ Disciplinary factors
- ☐ Rarity of data
- ☐ Data is kept forever
- ☐ Other (please explain below):

Display Question #41

If How do you define your role in relation to 3D data? = Repository manager
Or How do you define your role in relation to 3D data? = Both a creator and a repository manager

Question 41

Q53 Indexing: Is your repository open to indexing by external search engines?

- o Yes
- o No. Please explain why below:

Display Question #42

If How do you define your role in relation to 3D data? = Creator
Or How do you define your role in relation to 3D data? = Other (please explain below):

Question 42

Q55 Data Preservation: What measures do you take to ensure your data are not lost? (Select all that apply)

- ☐ Backups
- ☐ Fixity Checks

- ▢ Virus Scanning
- ▢ Other (please describe below):

Non-anonymous Basic Information

Question 43

Q82 What challenges do you find in managing the preservation of 3D data?

Question 44

Q56 What is the name of your organization or repository?

Question 45

Q57 Please provide your full name & email address:
End of Survey

Notes

1. Edward M. Corrado and Heather Moulaison Sandy, *Digital Preservation for Libraries, Archives, and Museums* (Lanham, MD: Rowman & Littlefield, 2014), xix.
2. Corrado and Sandy, *Digital Preservation*.
3. Consultative Committee for Space Data Systems, *Audit and Certification of Trustworthy Digital Repositories*, Recommended Practice, Issue 1, CCSDS 652.0-M-1, ISO 16363:2012 (Washington, DC: Consultative Committee for Space Data Systems, September 2011), https://public.ccsds.org/pubs/652x0m1.pdf.
4. Corrado and Sandy, *Digital Preservation*, 43–44.
5. CoreTrustSeal, "About," accessed May 2019, https://www.coretrustseal.org/about/; Nestor, "Nestor Seal for Trustworthy Digital Archives," accessed October 5, 2019, https://www.langzeitarchivierung.de/Webs/nestor/EN/Zertifizierung/nestor_Siegel/siegel.html.
6. Center for Research Libraries, "Digital Preservation Metrics," accessed 2019, https://www.crl.edu/archiving-preservation/digital-archives/metrics.
7. Digital Preservation Coalition, "Audit and Certification," *Digital Preservation Handbook*, 2nd ed. (Digital Preservation Coalition, 2015), https://www.dpconline.org/handbook/institutional-strategies/audit-and-certification.
8. Consultative Committee for Space Data Systems, *Reference Model for an Open Archival Information System (OAIS)*, Recommended Practice, Issue 2, CCSDS 650.0-M-2, Magenta Book (Washington DC: Consultative Committee for Space Data Systems, June 2012), https://public.ccsds.org/pubs/650x0m2.pdf.
9. Angie Payne, *Laser Scanning for Archaeology*, Guides to Good Practice (Archaeology Data Service and Digital Antiquity, 2011), https://guides.archaeologydataservice.ac.uk/g2gp/LaserScan_Toc.
10. Doug M. Boyer et al., "MorphoSource: Archiving And Sharing 3-D Digital Specimen Data," *Paleontological Society Papers* 22 (September 2016): 157–81, https://doi.org/10.1017/scs.2017.13; MorphoSource, "User Guide," accessed September 25, 2020, https://www.morphosource.org/docs/guide.
11. Katie Green, "Why Select and Appraise?" Guidance on the Selection of Material for Deposit and Archive, Archaeology Data Service, July 30, 2018, last modified July 2, 2019, https://archaeologydataservice.ac.uk/advice/selectionGuidance.xhtml#3.%20Why%20select%20and%20appraise.
12. Marc Levoy and Hector Garcia-Molina, *Creating Digital Archives of 3D Artworks*, white paper (Stanford, CA: Stanford University, December 4, 1999, last modified March 27, 2000), https://graphics.stanford.edu/projects/dli/white-paper/dli.html.

13. Martina Trognitz, Kieron Niven, and Valentijn Gilissen, "Section 4. Archiving Close-Range Photogrammetric Data," in *3D Models in Archaeology: A Guide to Good Practice*, Guides to Good Practice (Archaeology Data Service and Digital Antiquity, 2016), https://guides.archaeologydataservice.ac.uk/g2gp/Photogram_4-1.
14. Kenton McHenry and Peter Bajcsy, "An Overview of 3D Data Content, File Formats and Viewers," Technical Report isda08-002, National Center for Supercomputing Applications, 2008, https://www.archives.gov/files/applied-research/ncsa/8-an-overview-of-3d-data-content-file-formats-and-viewers.pdf.
15. McHenry and Bajcsy, "Overview of 3D Data Content."
16. McHenry and Bajcsy, "Overview of 3D Data Content."
17. Dibya Chakravorty, "8 Most Common 3D File Formats," All3DP, accessed June 25, 2019, https://all3dp.com/3d-file-format-3d-files-3d-printer-3d-cad-vrml-stl-obj (page content changed).
18. McHenry and Bajcsy, An Overview of 3D Data Content," 14.
19. Thomas Flynn, "Cultural Heritage User Survey 2019: Results, August 2019," Sketchfab, accessed September 2019, https://docs.google.com/presentation/d/1avExQtbfl6vDCFyo5VgSyBy638BjdDTAIhw79ClHT5A/edit?usp=sharing.
20. Martina Trognitz, Kieron Niven, and Valentijn Gilissen, "Section 3: Archiving 3D Data," in *3D Models in Archaeology: A Guide to Good Practice*, Guides to Good Practice (Archaeology Data Service and Digital Antiquity, 2016), https://guides.archaeologydataservice.ac.uk/g2gp/3d_3-1.
21. Library of Congress, "Recommended Formats Statement," "VIII. Design and 3D," "iii. Scanned 3D Objects (Output from Photogrammetry Scanning)," Preservation, accessed September 29, 2020, https://www.loc.gov/preservation/resources/rfs/design3D.html#scanned-3d.
22. Web3D Consortium, "What Is X3D?" accessed September 25, 2020, https://www.web3d.org/x3d/what-x3d.
23. Nicholas Polys, "Fulfilling the Mandate of Information Durability and Access" (presentation, Born to Be 3D: Digital Stewardship of Intrinsic 3D Data, Library of Congress, Washington, DC, November 2, 2018), https://www.loc.gov/preservation/digital/meetings/b2b3d/b2b3d2018.html; Martina Trognitz, Kieron Niven, and Valentijn Gilissen, "Section 2. Creating 3D Data," in *3D Models in Archaeology: A Guide to Good Practice*, Guides to Good Practice (Archaeology Data Service and Digital Antiquity, 2016), https://guides.archaeologydataservice.ac.uk/g2gp/3d_2-3.
24. Web3D Consortium, "1. 3D Modelling," "Develop and Deploy X3D," accessed September 25, 2020, https://www.web3d.org/develop-and-deploy-x3d.
25. National Digital Stewardship Alliance, "Levels of Digital Preservation," 2013, accessed June 25, 2019, https://ndsa.org//activities/levels-of-digital-preservation/.
26. Andy Klein, "Hard Drive Cost per Gigabyte," Backblaze, July 11, 2017, https://www.backblaze.com/blog/hard-drive-cost-per-gigabyte/.
27. Consultative Committee for Space Data Systems, "4: Detailed Models," in *Reference Model*, pp. 4-1–4-55.
28. Consultative Committee for Space Data Systems, *Reference Model*, p. 4-30.
29. Library of Congress, "Sustainability Factors," in "Sustainability of Digital Formats: Planning for Library of Congress Collections," accessed June 25, 2019, https://www.loc.gov/preservation/digital/formats/sustain/sustain.shtml.
30. Stefana Breitwieser, e-mail message to the authors, July 19, 2019.
31. Association of Research Libraries, *Code of Best Practices in Fair Use for Software Preservation* (Washington, DC: Association of Research Libraries, September 2018, last modified February 2019), https://www.arl.org/resources/code-of-best-practices-in-fair-use-for-software-preservation.
32. See Consultative Committee for Space Data Systems, *Reference Model*, table 4-1 on p. 4-32 for recommended Preservation Description Information for a software package.
33. M. D. McIlroy, E. N. Pinson, and B. A. Tague, "UNIX Time-Sharing System: Foreword," *Bell System Technical Journal* 57, no. 6 (July–August 1978): 1902–03; GNU Operating System, "Philosophy of the GNU Project," accessed July 1, 2019, https://www.gnu.org/philosophy/.
34. Global Biodiversity Information Facility, "IPT: The Integrated Publishing Toolkit," https://www.gbif.org/ipt.
35. Dryad, "Submission and Publication Process," Dryad, accessed November 9, 2019, https://datadryad.org/stash/submission_process.

36. Levoy and Garcia-Molina, *Creating Digital Archives*.
37. Duke University, "About MorphoSource," accessed 2019, https://web.archive.org/web/20190407020947/https://www.morphosource.org/About/home.
38. Boyer et al., "MorphoSource."
39. Mark Edward Phillips et al., " UNT Libraries' Archival Storage Replacement Fund" appendix O in *UNT Libraries: TRAC Conformance Document* (Denton: University of North Texas, October 2015),https://digital.library.unt.edu/ark:/67531/metadc1040526/.
40. Phillips et al., " UNT Libraries' Archival Storage Replacement Fund."
41. Martina Trognitz, Kieron Niven, and Valentijn Gilissen, table of contents, *3D Models in Archaeology: A Guide to Good Practice*, Guides to Good Practice (Archaeology Data Service and Digital Antiquity, 2016), https://guides.archaeologydataservice.ac.uk/g2gp/3d_Toc.
42. Lisa R. Johnston, *Curating Research Data* (Chicago: Association of College and Research Libraries, 2017), http://www.ala.org/acrl/sites/ala.org.acrl/files/content/publications/booksanddigitalresources/digital/9780838988633_crd_v2_OA.pdf.
43. William Kilbride, "3D4Ever: Why Is It So Hard to Talk about the Preservation of 3D Data?" *Journal of the Institute of Conservation* 40, no. 2 (2017): 183–89, https://doi.org/10.1080/19455224.2017.1317006.
44. Archaeology Data Service, "Charging Policy," 2016, https://archaeologydataservice.ac.uk/advice/PolicyDocuments.xhtml#CharPol.
45. Summer Decker and Jonathan Ford, "Management of 3D Image Data," chapter 13 in *Human Remains: Another Dimension* (Amsterdam, Netherlands: Elsevier, 2017), https://doi.org/10.1016/B978-0-12-804602-9.00014-X; Tony Austin and Jenny Mitcham, *Preservation and Management Strategies for Exceptionally Large Data Formats: "Big Data."* (Archaeology Data Service and English Heritage, 2007), https://archaeologydataservice.ac.uk/research/bigData.xhtml.
46. tDAR, "Search Results: Showing All Items Narrowed by: Object Type: '3D & Sensory Data,'" https://core.tdar.org/search/results?_tdar.searchType=advanced&groups%5B0%5D.operator=AND&groups%5B0%5D.fieldTypes%5B0%5D=ALL_FIELDS&groups%5B0%5D.allFields%5B0%5D=&objectTypes=SENSORY_DATA&groups%5B0%5D.latitudeLongitudeBoxes%5B0%5D.east=&groups%5B0%5D.latitudeLongitudeBoxes%5B0%5D.south=&groups%5B0%5D.latitudeLongitudeBoxes%5B0%5D.west=&groups%5B0%5D.latitudeLongitudeBoxes%5B0%5D.north=&sortField=RELEVANCE.
47. One existing model is that of Serge J. Goldstein and Mark Ratliff, "DataSpace: A Funding and Operational Model for Long-Term Preservation and Sharing of Research Data," DataSpace, 2010, accessed October 4, 2019, http://arks.princeton.edu/ark:/88435/dsp01w6634361k.
48. Kieron Niven, ed., "The Project Archive: Storage and Dissemination," Guides to Good Practice (Archaeology Data Service/Digital Antiquity, 2011), https://guides.archaeologydataservice.ac.uk/g2gp/ArchivalStrat_1-4.
49. Blue Ribbon Task Force on Sustainable Digital Preservation and Access, *Sustainable Economics for a Digital Planet* (Blue Ribbon Task Force on sustainable Digital Preservation and Access, February 2010), http://brtf.sdsc.edu/biblio/BRTF_Final_Report.pdf; Organisation for Economic Cooperation and Development, *Business Models for Sustainable Research Data Repositories*, OECD Science, Technology, and Innovation Policy Papers, no. 47 (Paris: OECD Publishing, December 2017), https://doi.org/10.1787/302b12bb-en; Nancy Maron, *A Guide to the Best Revenue Models and Funding Sources for Your Digital Resources*, research report (New York: Ithaka S+R, 2014), https://sr.ithaka.org/publications/a-guide-to-the-best-revenue-models-and-funding-sources-for-your-digital-resources/.
50. Debra L. Madsen and Jenny K. Oleen, "Staffing and Workflow of a Maturing Institutional Repository," *Journal of Librarianship and Scholarly Communication* 1, no. 3 (2013), eP1063, https://krex.k-state.edu/dspace/bitstream/handle/2097/15393/MadsenJofLibrarScholComm2013.pdf.

Bibliography

Archaeology Data Service. "Charging Policy." 2016. https://archaeologydataservice.ac.uk/advice/PolicyDocuments.xhtml#CharPol.

Association of Research Libraries. *Code of Best Practices in Fair Use for Software Preservation*. Washington, DC: Association of Research Libraries, September 2018, last modified February 2019. https://www.arl.org/resources/code-of-best-practices-in-fair-use-for-software-preservation.

Austin, Tony, and Jenny Mitcham. *Preservation and Management Strategies for Exceptionally Large Data Formats: "Big Data."* Archaeology Data Service and English Heritage. 2007. Accessed May 8, 2019. https://archaeologydataservice.ac.uk/research/bigData.xhtml.

Blue Ribbon Task Force on Sustainable Digital Preservation and Access. *Sustainable Economics for a Digital Planet: Ensuring Long-Term Access to Digital Information*. Blue Ribbon Task Force on Sustainable Digital Preservation and Access, 2010. http://brtf.sdsc.edu/biblio/BRTF_Final_Report.pdf.

Boyer, Doug M., Gregg F. Gunnell, Seth Kaufman, and Timothy M. McGeary. "Morphosource: Archiving and Sharing 3-D Digital Specimen Data." *Paleontological Society Papers* 22 (September 2016): 157–81. https://doi.org/10.1017/scs.2017.13.

Center for Research Libraries. "Digital Preservation Metrics." Accessed 2019. https://www.crl.edu/archiving-preservation/digital-archives/metrics.

Chakravorty, Dibya. "8 Most Common 3D File Formats." All3DP. Accessed June 25, 2019. https://all3dp.com/3d-file-format-3d-files-3d-printer-3d-cad-vrml-stl-obj (page content changed).

Consultative Committee for Space Data Systems. *Audit and Certification of Trustworthy Digital Repositories*. Recommended Practice, Issue1. CCSDS 652.0-M-1, ISO 16363:2012. Washington, DC: Consultative Committee for Space Data Systems, September 2011. https://public.ccsds.org/pubs/652x0m1.pdf.

———. *Reference Model for an Open Archival Information System (OAIS)*. Recommended Practice, Issue 2. CCSDS 650.0-M-2, Magenta Book. Washington DC: Consultative Committee for Space Data Systems, June 2012. https://public.ccsds.org/pubs/650x0m2.pdf.

CoreTrustSeal. "About." Accessed May 2019. https://www.coretrustseal.org/about/.

Corrado, Edward M., and Heather Moulaison Sandy. *Digital Preservation for Libraries, Archives, and Museums*. Lanham, MD: Rowman & Littlefield, 2014. ProQuest Ebook Central.

Digital Preservation Coalition. "Audit and Certification." *Digital Preservation Handbook*, 2nd ed. Digital Preservation Coalition, 2015. https://www.dpconline.org/handbook.

Decker, Summer, and Jonathan Ford. "Management of 3D Image Data." Chapter 13 in *Human Remains: Another Dimension: The Application of Imaging to the Study of Human Remains*. Amsterdam, Netherlands: Elsevier, 2017. https:// doi.org/10.1016/B978-0-12-804602-9.00014-X.

Dryad. "Submission and Publication Process." Accessed November 9, 2019. https://datadryad.org/stash/submission_process.

Duke University. "About MorphoSource."Accessed 2019. https://web.archive.org/web/20190407020947/https://www.morphosource.org/About/home.

Flynn, Thomas. "Cultural Heritage User Survey 2019: Results, August 2019." Sketchfab. Accessed September 2019. https://docs.google.com/presentation/d/1avExQtbfl6vDCFyo5VgSyBy638BjdDTAIhw79ClHT5A/edit?usp=sharing.

Global Biodiversity Information Facility. "IPT: The Integrated Publishing Toolkit." https://www.gbif.org/ipt.

GNU Operating System. "Philosophy of the GNU Project." Accessed July 1, 2019. https://www.gnu.org/philosophy/.

Goldstein, Serge J., and Mark Ratliff. "DataSpace: A Funding and Operational Model for Long-Term Preservation and Sharing of Research Data." DataSpace. 2010. Accessed October 4, 2019. http://arks.princeton.edu/ark:/88435/dsp01w6634361k.

Green, Katie. "Why Select and Appraise?" Guidance on the Selection of Material for Deposit and Archive. Archaeology Data Service, July 30, 2018, last modified July 2, 2019. https://archaeologydataservice.ac.uk/advice/selectionGuidance.xhtml#3.%20Why%20select%20and%20appraise.

Johnston, Lisa R. *Curating Research Data: A Handbook of Current Practice*. Chicago: Association of College and Research Libraries. 2017. http://www.ala.org/acrl/sites/ala.org.acrl/files/content/publications/booksanddigitalresources/digital/9780838988633_crd_v2_OA.pdf.

Kilbride, William. "3D4Ever: Why Is It So Hard to Talk about the Preservation of 3D Data?" *Journal of the Institute of Conservation* 40, no. 2 (2017): 183–89. https://doi.org/10.1080/19455224.2017.1317006.

Klein, Andy. "Hard Drive Cost Per Gigabyte." Backblaze, July 11, 2017. https://www.backblaze.com/blog/hard-drive-cost-per-gigabyte/.

Levoy, Marc, and Hector Garcia-Molina. "Creating Digital Archives of 3D Artworks." White paper. Stanford, CA: Stanford University, December 4, 1999, last modified March 27, 2000. https://graphics.stanford.edu/projects/dli/white-paper/dli.html.

Library of Congress. "Recommended Formats Statement." Preservation. Accessed September 29, 2020. https://www.loc.gov/preservation/resources/rfs/design3D.html#scanned-3d.

———. "Sustainability of Digital Formats: Planning for Library of Congress Collections." Accessed June 25, 2019. https://www.loc.gov/preservation/digital/formats/sustain/sustain.shtml.

Madsen, Debra L., and Jenny K. Oleen. "Staffing and Workflow of a Maturing Institutional Repository." *Journal of Librarianship and Scholarly Communication* 1, no. 3 (2013), eP1063. https://krex.k-state.edu/dspace/bitstream/handle/2097/15393/MadsenJofLibrarScholComm2013.pdf;sequence=1.

Maron, Nancy. *A Guide to the Best Revenue Models and Funding Sources for Your Digital Resources*. Research report. New York: Ithaka S+R, 2014. https://sr.ithaka.org/publications/a-guide-to-the-best-revenue-models-and-funding-sources-for-your-digital-resources/.

McHenry, Kenton, and Peter Bajcsy. "An Overview of 3D Data Content, File Formats and Viewers." Technical Report isda08-002. National Center for Supercomputing Applications, 2008. https://www.archives.gov/files/applied-research/ncsa/8-an-overview-of-3d-data-content-file-formats-and-viewers.pdf.

McIlroy, M. D., E. N. Pinson, and B. A. Tague. "UNIX Time-Sharing System: Foreword." *Bell System Technical Journal* 57, no. 6 (July–August 1978): 1899–1904.

MorphoSource. "User Guide." Accessed September 25, 2020. https://www.morphosource.org/docs/guide.

National Digital Stewardship Alliance. "Levels of Digital Preservation." 2013, accessed June 25, 2019. https://ndsa.org//activities/levels-of-digital-preservation/.

Nestor. "nestor Seal for Trustworthy Digital Archives." Accessed October 5, 2019. https://www.langzeitarchivierung.de/Webs/nestor/EN/Zertifizierung/nestor_Siegel/siegel.html.

Niven, Kieron, ed. "The Project Archive: Storage and Dissemination." Guides to Good Practice. Archaeology Data Service and Digital Antiquity, 2011. https://guides.archaeologydataservice.ac.uk/g2gp/ArchivalStrat_1-4.

Organisation for Economic Cooperation and Development. *Business Models for Sustainable Research Data Repositories*. OECD Science, Technology, and Innovation Policy Papers, no. 47. Paris: OECD Publishing, December 2017. https://doi.org/10.1787/302b12bb-en.

Payne, Angie. *Laser Scanning for Archaeology: A Guide to Good Practice*. Guides to Good Practice. Archaeology Data Service and Digital Antiquity, 2011. http://guides.archaeologydataservice.ac.uk/g2gp/LaserScan_Toc.

Phillips, Mark Edward, Ana Krahmer, Hannah Tarver, Daniel Alemneh, and Laura Waugh. "UNT Libraries' Archival Storage Replacement Fund." Appendix O in *UNT Libraries: TRAC Conformance Document*. Denton: University of North Texas, October 2015. https://digital.library.unt.edu/ark:/67531/metadc1040526/.

Polys, Nicholas. "Fulfilling the Mandate of Information Durability and Access." Presentation, Born to Be 3D: Digital Stewardship of Intrinsic 3D Data, Library of Congress, Washington, DC, November 2, 2018. https://www.loc.gov/preservation/digital/meetings/b2b3d/b2b3d2018.html.

Re3data.org (Registry of Research Data Repositories) home page. https://www.re3data.org/.

tDAR home page. https://www.tdar.org/.

Trognitz, Martina, Kieron Niven, and Valentijn Gilissen. *3D Models in Archaeology: A Guide to Good Practice*. Guides to Good Practice. Archaeology Data Service and Digital Antiquity, 2016. https://guides.archaeologydataservice.ac.uk/g2gp/Contents.

Web3D Consortium. "Develop and Deploy X3D." Accessed September 25, 2020. https://www.web3d.org/develop-and-deploy-x3d.

———. "What Is X3D?" Accessed September 25, 2020. https://www.web3d.org/x3d/what-x3d.

Chapter 4

Metadata Requirements for 3D Data

Jon Blundell, Jasmine L. Clark, Katherine E. DeVet, and Juliet L. Hardesty

ABSTRACT

This chapter provides recommendations for ***metadata*** *based on the five-stage* ***digital asset life cycle****. The section "Create" covers some of the principal ways 3D models are created and discusses what metadata can be captured during the creation process. It looks at not only what metadata could be captured during model creation, but also why capturing that information is important. The section "Manage" covers the metadata needs for organizing, verifying, and providing* ***access*** *to 3D data. Recommendations include grouping files together as much as possible (by 3D object, by collection of objects, and by project) in order to apply organizational metadata that can be used for access and reuse purposes. The section "Distribute and Publish" discusses the need for a variety of distribution platforms that support the broadly varying metadata needs of different disciplines. Examples include the need for more granular metadata to support reproducibility and privacy in certain fields, as well as concerns around metadata requirements for accessibility for disability more broadly. Though the circulation and access norms for 3D data are still evolving, the section "Access and Reuse" posits key metadata anticipated to be useful in the* ***discovery*** *and*

access of 3D data and models for research or reuse. The section "Archive" utilizes PREMIS as a basis for its recommendations. The rapid changes in the tools and platforms that support the creation and utilization of 3D data result in heavier emphasis on metadata that provide context to data that are often no longer supported by the latest technologies. Additional portions of PREMIS that may be of interest to readers are also specified. The chapter ends with an overall table of recommended metadata fields along with future work needed, naming annotation metadata and metadata for accessibility needs as top priorities for standardization and best practice recommendations.

Introduction

As discussed in previous chapters, one of the ongoing struggles of any technology-based medium, including 3D data and objects, is how it is stored, cataloged, and accessed for later reuse. The life span of 3D data can be greatly enhanced by the development of best practices for gathering and categorizing 3D metadata, including both information currently recognized by libraries and **archives** as well as creation data, termed **paradata** by the London Charter.[1] As mentioned in chapter 2, "Best Practices for 3D Data Preservation," the activity of metadata generation is a key intervention that should occur at many preservation intervention points (PIPs) throughout the digital asset life cycle. This chapter will make recommendations for metadata needs for 3D data and objects through the life cycle stages: create, manage, distribute and publish, access and reuse, and archive. The digital asset life cycle discussed here describes the stages used to manage digital files through their digital asset management (DAM) software and is a simplified version of the DCC Curation Lifecycle Model.[2] While not all steps include 3D-specific metadata, working within this framework helps to identify where 3D-specific activities and outputs occur and what metadata are needed to record and track within those stages of the life cycle. The recommendations include the types of metadata needed (names or dates, for example) but do not specify a particular metadata standard to use or controlled vocabularies to apply. This work is based on experience with metadata and 3D object creation and collection management from the CS3DP community. Examples from the community and feedback on the metadata recommendations will also be shared. While examples include a mix of metadata standards used for different types of information, future work is needed to gather consensus around standard metadata properties and controlled vocabularies for 3D models and collection management. Within the **Good/Better/Best** framework of CS3DP recommendations, this chapter offers something between Good and Better, with Best being common metadata standards adopted and used for all 3D object creation and collection management.

Methods

The recommendations made in this chapter are based upon data collected via two surveys along with feedback solicited via the second Community Standards for 3D Data Preservation (CS3DP) forum at the University of Michigan. The first survey asked stakeholders working with 3D data for their current metadata practices and models. These responses were then collated into a Google Sheets spreadsheet[3] and categorized by the types of data they described (project, model/data, preprocessing/processing, capture, and original item; see the next section, "Considerations, Decisions, and Scope," for more information). These results were then shared at the second CS3DP forum. To aid in our understanding of the information needed by different categories of stakeholders, we collected user profiles and user stories, much like the case studies seen in chapter 2, "Best Practices for 3D Preservation," taking care to consider individuals who may interact with 3D data at all points within the life cycle of those data.

Attendees of the second CS3DP forum brought a wide variety of perspectives on the creation and management of 3D data based on their role in the process (such as creators, publishers, or **repository managers**) and the **workflows** and technical lexicon of their particular community or institution. These diverse frames of reference led to confusion regarding the initial survey results among forum attendees when the survey's terminology and assumed workflows did not align with the forum attendees' practices. Based on this feedback, the decision was made to organize this chapter and its related metadata recommendations according to the digital asset life cycle. This was intended to ease data collection in our second survey, as respondents could focus on their respective areas of expertise, while also making the results easier for readers to navigate. The second survey (see Appendix) gathered data regarding the metadata fields collected during the first survey, while also acknowledging potential gaps, and asked respondents to identify additional metadata fields they thought were missing from the survey.

Considerations, Decisions, and Scope

Responses to the surveys were limited in number, and disciplinary representation was equally limited. There were ten responses to the first survey, with anthropology, archaeology, geology, and museums being the represented disciplines. There was a lack of data regarding utilized schemata, tools, phases of capture, workflows, and objects designed and modeled digitally (not scans of physical items). We received eight responses to the second survey and will discuss the results in each section of this chapter. There is also a need to acknowledge that the community of stakeholders producing large quantities of 3D data is limited in size and still evolving. The data captured within this chapter should be considered a snapshot of current 3D practices in the academic and cultural sectors.

Consideration was given to different methods of capture and creation along the lines of those mentioned in chapter 2, "Best Practices for 3D Data Preservation," but in attempting to list and address various methods, it was found that extensive knowledge of common or assumed workflows and technical requirements for each method would need to be articulated to prescribe metadata recommendations for each method (e.g., the metadata needs for CT data are distinctly different from those of photogrammetric data). This was determined to be out of scope for this chapter, and, as a result, this chapter will focus on examining the creation metadata elements common across all production methods.* The information available in this chapter pertains to commonalities and unique fields related to the general categories of reality-capture, sources-based, and artistic 3D data. While the creation metadata discussed in this chapter could be viewed as what the London Charter calls paradata, for the purpose of these recommendations, this information is considered to be another form of metadata that are able to be captured and cataloged.[4]

Digital Asset Life Cycle and 3D Metadata

Create

Metadata in the Create stage are associated with the process of collecting or capturing source data and the process of model construction (process inputs), the finished 3D model (process outputs), and, for models that are digitized versions of physical items, metadata associated with those physical items.

Metadata capture needs to begin as soon as a project involving creation of 3D models (or 3D data in general) is conceptualized. This is because the project's intent and the way the resulting models will be used strongly inform the type and level of documentation required. Central to understanding the metadata needs of a given 3D model is knowing the method used to create it. Creation documentation needs for a reality-capture model, created from 3D scans or **photogrammetry**, will significantly differ from those of a sources-based model, created manually using reference material. A purely artistic model's documentation needs vary further still. A model can be created using any mix of these approaches, and there are other approaches, such as procedural modeling, that, while outside of what is discussed in this chapter, would still need to be considered if applicable. Of those surveyed, every respondent documents the project for

* For examples of more granular, method-specific metadata, see responses to the first survey in this chapter's appendix.

which a model was created, as well as its method of creation or acquisition,[†] a practice which this chapter recommends.

Stakeholders

Model creation documentation is primarily the concern of those producing 3D models and, for models that represent real-world objects and environments, the holding entities and stewards of those subjects. Once created, this documentation is most relevant to those with expertise in digital 3D technologies and those who need a deep level of understanding of a 3D model, such as those with research, academic, or other technical interests in the material.

Survey

As indicated by the survey respondents, there are myriad ways to create 3D data. Each process requires specific approaches to capturing metadata that will be useful for the data's **preservation** and use or reuse. These 3D data creation and capture methodologies, as well as possible variations of metadata one should document as part of these processes, are outlined below.

General Creation Metadata

This chapter focuses on the three categories listed above: reality-capture, sources-based, and artistic. While each has specific metadata needs, common elements are seen across all three: the basic who, what, when, and how (with "where" depending on the subject of digitization). Of those surveyed, all reported recording a name, description, or identifier of what was created, actors involved, tools and software used, and the dates when actions happened.[‡] These are not 3D-specific terms, but broad terms that can describe most any act of creation. Thus, there is no need to use new or exotic metadata elements to capture this information. Commonly used **descriptive metadata** standards such as Dublin Core, Darwin Core, VRA Core, or MODS will cover this information, though a possible expansion of controlled vocabularies could be needed depending on the standard used. It should also be noted here that, if there is a real-world object or environment that a model is derived from, the necessity of metadata to describe that real-world object or environment will depend on whether or not the real-world object is described in a digital system elsewhere. If real-world description is needed, one of the commonly used descriptive metadata standards mentioned above or a relevant domain-specific metadata standard should be used. Depending on the data management approach, the system that stores 3D data and models may not need to store information about real-world counterparts if it exists in another system of record,

† Appendix, responses in Method Used and Project Identifying Information.

‡ Appendix, responses in Model Identifying Information, Linked Fields, Resources, and Processing Action.

other than the identifiers needed to link the two systems (preferably using a globally unique persistent and resolvable identifier, or GUPRI). This idea is supported by the survey, where all respondents used unique identifiers, when applicable, to identify the real-world object or environment on which a model was based.*

Metadata elements that track the finer-grained steps and decisions that go into a given creation (processing) action are used less than the basic who, what, when, and how metadata elements, as reported by the survey respondents.† This is reasonable as there are two major logistical challenges to this higher level of documentation. First, and maybe obviously, documentation becomes more time-consuming and burdensome as reporting requirements increase, decreasing the likelihood that the documentation will be undertaken. Second, the more granular creation metadata are, the more task- and tool-specific they need to be. An effective metadata reporting strategy and schema would go a long way toward facilitating more in-depth creation documentation, of which there are few, if any, accepted community standards.

Creation Transparency and Reuse

While acknowledging the burden of process documentation, such documentation is crucial for any models used in a published work or a scientific or academic setting. The source information for a model, as well as the manipulations and interpretations of the source data that contribute to a model's creation, is critical to understanding what a given model represents and, importantly, what it does not. Both reality-capture and source-based models originate from collected information, whether that information is scan data or reference observations and media. Understanding how this information is interpreted and manipulated to create a 3D model is key to understanding what elements of that model are representative of the source information and what elements are artifacts of processing or interpretation. As part of understanding this information, it is essential to recognize that any reality-capture or source-based model is modified by the lens of interpretation, whether that interpretation is coming from an algorithm in a piece of hardware or software or from a human making judgment calls while interacting with that hardware or software. The goal of tracking the steps that go into a model's creation is to give transparency to these interpretations, allowing those that interact with the produced model (for comparison, illustration, investigation, etc.) to fairly assess the model's trustworthiness for their needs.

Thus, it follows that any 3D models, created by reality capture or modeled from sources, that purport to represent a real-world object or scene should have as much creation documentation as is reasonable. Capturing creation metadata also allows for repeatability of the creation process, so the same or a similar enough model may be derived using the same source data and same processing methods. There is a balance

---

* Appendix, responses in Original Item.

† Appendix, responses in Processing Actions.

that must be struck between data transparency and workflow efficiency when creating and presenting 3D models as digital facsimiles of the physical world. How to arrive at that middle ground is something that creators and repositories must currently decide for themselves and should be a focus of community standards development in order to better facilitate the sharing and reuse of resources and data.

Metadata for Reality Capture Models

While reality capture and manually created source-based models share many core metadata features, each has its own needs and considerations. The source information for reality-capture models is the raw data from the capture device (laser scanner, CT scanner, camera, etc.). These raw data could be in the form of **meshes**, **point clouds**, or images. They could be output in open, accessible file types or closed proprietary file types. While the files that are produced from a scanning device may be termed raw data, it should be recognized that they have undergone some type of interpretation on the capture device to convert the electrical sensor readings to an intelligible format such as those listed above. While some of this interpretation is opaque to the end user, whether because it is a trade secret or simply not provided by the tool, capture equipment should have device parameters or user-definable settings that can be documented. As many tools and software do not encode parameters and settings used into the files they output in a user-accessible way, it should be standard practice to document any applicable parameters and settings, as well as basic information such as the make and model of the equipment used, as part of recording capture metadata.[‡] Along with documenting capture tools and their settings, it is also helpful, depending on the capture technique, to document the capture process and the intended strategy for data processing. Examples include, but are not limited to, if targets were used for aligning multiple capture types such as laser and photogrammetry, if a spherical or focal stack rig was employed, if color calibration targets were used, or if there are calibration or rig scaling datasets.

Metadata for Manually Created Sources-Based Models

The reference material for source-based, manually created models could come from a variety of resources. Examples could include measurements of existing or partial artifacts, sketches of hypothesized historical representations, written documentation from research sites, and photographs or other media sources depicting the object or environment (or even similar or related objects and environments). How these references are used and interpreted needs to be captured and cited, just as it would in academic writing. As with reality-capture models, it is critical to understand what sources were used during the creation of a model. For example, this information would allow someone viewing a historical reconstruction to understand what parts of the model are

‡ An example of a tool that does include parameters and settings in its output files via embedded EXIF metadata would be a digital camera.

based on documentation and what elements have been filled in from imagination or interpolated from relevant sources to complete the reconstruction. Citation possibilities are discussed later in this chapter in the section "Access and Reuse."

Model Processing Metadata

Once capture and source data documentation is complete, processing documentation needs to be considered. Of the survey respondents, roughly half reported recording processing decisions, action methods, and descriptions.* The relatively low number is understandable as the approaches and software that can be used to process a model are vast and varied. Also, the underlying algorithms and their inputs that are used to manipulate 3D data are opaque in most commercial software (an exception is that many photogrammetry software packages allow detailed reports, which include settings and other metrics, to be generated).† This makes process documentations complex and difficult to standardize. Documenting software and source datasets, as well as any large decisions or modifications such as noise reduction and hole filling, will go a long way toward model transparency. Citing a followed best practice, such as the Bureau of Land Management (BLM) or Cultural Heritage Imaging (CHI) photogrammetry workflow, can also provide significant insight into the processing workflow that created a model.[5] Another facet of processing documentation is whether down-sampled models are needed to facilitate use, such as derivative models made for web viewing or 3D printing. In this case, it is important to be able to document the source model for any derivative models created, as well as the methods of derivative model creation so that this information can be passed on to the model's consumers.

Metadata for Artistic Models

For artistic creations, the creation documentation needs will vary widely depending on the use case and the creator's intent. As works of art and artistic representations could be intended for any number of modes of consumption, both digital and physical, the information needed to codify the digital representation, and associated files could vary widely. For the most part, creation or capture data will be similar to those needed for capture-based or source-based models, depending on the creation modes and elements used. These models may also be a part of a time-based media experience, gallery or artistic installations, interactive experiences, and games. Taking this fact into account, relevant exhibition or other contextual information may also be necessary to note.

---

* Appendix, responses in Processing Actions.

† It should be noted that some 3D processing software can generate process and quality reports, and if available, these reports should be stored with the respective model.

Creation Metadata Example

A common 3D model creation method is digital photogrammetry. The source information for this creation method is a collection of digital images. Examples of creation metadata related to this source information would be

- the subject of digitization (e.g., descriptive information on the subject of capture, identifiers pertaining to the subject);
- equipment used (e.g., camera make and model, scale bar measurements), equipment settings (e.g., ISO speed, f-stop);
- capture approach (e.g., Was a camera array used? Was cross-polarization used?); and
- post-processing actions performed on the images before they are brought into a photogrammetry software package (e.g., Was chromatic aberration correction applied? Was image sharpening applied?).

Depending on the file types used, some of this information might be stored directly in the image files (e.g., EXIF metadata), some might be stored in standard digital photography **sidecar** files (e.g., XMP files), and some information will have to be recorded outside of standard digital photography tools and conventions.

Examples of creation metadata related to the model processing would include

- which images were used during model creation;
- how these images were used (e.g., Was an image used during alignment, surface reconstruction, or texture mapping?);
- what processing actions were taken (e.g., camera alignment, geometry smoothing), along with the software used and input parameters for those actions; and
- qualitative information characterizing the process (e.g., the project's average alignment or **reprojection error**).

While some software will produce limited reports on the model creation and editing processes, for the most part this process information does not have a standardized way of being exported from processing software or generally being documented. Much of the creation metadata will have to be recorded in an outside tool (e.g., spreadsheet, MorphoSource, the Digital Lab Notebook) in either a structured or narrative format.

At this point, the process creation metadata have been addressed. The remaining metadata of interest as they relate to a produced 3D model are mostly technical information (e.g., model size, model **resolution**, UV map types), as well as information about the subject the model represents. Without delving into the specific technical aspects of different 3D file formats, it is worth noting some formats lack the ability to fully describe themselves in relation to types of information they may contain, as well as the numerous methods by which a 3D model can be **rendered**. Because of this, it might be necessary to store additional **technical metadata** outside of the model files. For example, many 3D model file formats store **texture maps** in separate individual image files, and not all 3D file formats fully describe all possible texture map types. Thus, it might be necessary to record how a texture map should be used (e.g., diffuse color, normals, ambient occlusion) separately from the model files. It is also worth mentioning that few if any 3D model file formats store scale information (e.g., Are the model's units in millimeters, feet, etc.?), so this metadata would also need to be stored separately.

Implications

As with most metadata collections, the bottom line is that data creators need to take the time to record the relevant information in their workflow before it is lost. Creators should look to leverage tools to make their metadata collection tasks easier, whether that is creating "shoot sheets" or using a tracking program such as CHI's Digital Lab Notebook.[6] Much of the burden of creation documentation could be eased if tools and software self-reported by providing action logs in a nonproprietary format, though at the time of writing this is largely not the case. Lastly, there need to be realistic expectations on how much information can be reasonably tracked. One survey respondent reported the SHAPES (Sharing and Helping Academics Prepare for Educational Success) project, housed out of the Texas Tech University Libraries, started with over seventy metadata fields describing cataloged 3D models, but after realizing the massive task of reporting to this level, the schema was reworked to a more manageable eighteen fields.[7] In their words, when deciding on an appropriate schema, "find the fields that are the most important and focus on those, otherwise it'll take someone an hour to do one record, and that kind of time in the metadata processing denotes an inefficient schema" (see responses to the second survey in this chapter's appendix). While there is no prescribed amount of time necessary to create metadata, the point here is that context matters and considerations should be made based on expected use of the 3D model and available resources for metadata creation. Specific considerations are

- what metadata are needed to facilitate the expected downstream use of the 3D data and models;
- who is doing the work of metadata creation;
- how much time they have available for the task;
- how much work it will take; and
- what tools, if any, are available to ease the work.

Recommendations

The recommendations below are separated based on the level of effort required and rigor of intended reuse. Recommendations for Good practice assume basic access and use of datasets by casual users, addressing discovery, context, and citation needs. Recommendations for Better practice facilitate informed use of datasets and 3D models in a research or academic context where judgments on data quality and suitability are required. Recommendations for Best practice are aimed at full reproducibility and might require a high level of technical experience.

Metadata Recommendations: Create

Good

Document and use a folder structure and file naming convention to organize source data and 3D models for management purposes. For example, group data and models by digitization project, collection, etc.

Include the following information as structured data in README.txt files, CSV files, spreadsheet, or similar.

Project level information

- project name
- project identifier
- project date
- description/abstract
- project authors/creators
- stakeholders/contributors
- project rights information

Source data information, if a reality-capture or sources-based model

- method of creation
- creation date
- information identifying real-world object or environment
- creators

3D model information

- subjects
- creator
- geometry information (e.g., number of faces, bounding box size, etc.)
- textures (what types of UV maps are available)
- materials (if any material properties are applied to a model)

Better

Include the following additional information as structured data in README.txt files, CSV files, spreadsheet, or similar.

Source information for reality-capture or sources-based models

- persistent identifiers (preferably GUPRIs) for records of subject or sources
- data sources
- georeference information if applicable

Source data creation information

- capture device make and model
- capture event details

3D model processing information

- source data used
- software used

Documentation of capture and processing workflows

Best

Include the following additional information as structured data in README.txt files, CSV files, spreadsheet, or similar.

3D model processing information

- detailed steps and log outputs

Use standardized metadata properties to define project, file, and 3D model properties listed above.

Manage

The Manage stage within the digital 3D asset life cycle is meant to ensure that those charged with maintaining 3D models and data have access to those files and accompanying metadata. Metadata in the Manage stage are associated with reviewing, annotating, and approving activities, along with version control and the logistics of giving people access to view or create annotations on a digital asset. Digital asset management system software often provides roles with appropriate access levels that can be assigned to individuals or groups and can be useful for this purpose if other needs for storing and accessing 3D models are also met. Metadata at this stage are often associated with reviewing files to verify that what was expected has been received. Creating **checksums** (algorithmically generated sequences of numbers and letters that represent the data contents of files) before files are passed along to this stage can make this type of verification a fairly quick and objective task (see glossary for further definition and explanation of checksums). Received files can be verified using a checksum calculator to compare and ensure that nothing has changed about the received files during transmission. Metadata also help at this stage for activities such as determining intellectual property rights to establish how a 3D model can be accessed and reused, annotating 3D models with descriptive metadata, approving activities to verify any modifications, version control, and logistics of giving people access to view and comment on a digital asset. Allowing others, such as curators, catalogers, information specialists, or subject matter experts, to access digital assets at this point provides the ability to enrich the 3D model, add internal cataloging notes, or cite external resources.

Stakeholders

The main stakeholders with management concerns are collection owners and collection managers who will be working directly with digital objects to review, annotate, and approve that content is stored and available as expected. These stakeholders will also determine the appropriate level of access to set for public availability—what formats will be available and how widely.

Collection owners are likely to be in charge of multiple collections and can have a variety of concerns for managing collections online. Examples of collection owners include archivists, museum curators, archaeologists managing an excavation site, artists, game developers, or biologists with research specimens. Collection owners will be particularly concerned that access levels be set appropriately for 3D models and their accompanying files and metadata (if everything should be fully available or if restrictions should be applied). They will also want to ensure that all 3D models are present in the system, that they accurately represent any physical objects on which they might be based (for example, all files are present when compared to a manifest with checksums), that 3D models that are not reality-based are completely represented (for example, all 3D models that should be part of a virtual reality environment are included), and that everything is described accurately and appropriately.

Collection managers will not necessarily be stewards of a collection of 3D models but could be in charge of managing online access across various collections within a single digital asset management system. Examples of collection managers include digital collection service managers (system administrators or those who manage a **digital repository** or digital asset management system used by several collection owners), web developers, librarians, or others involved in information technology management. Collection managers will be concerned with ensuring all necessary files and required metadata are present and that a 3D model or collection of 3D models has enough information to be discoverable within a larger system or set of models and collections.

Survey

The initial use case survey identified metadata properties associated with managing 3D data. Based on this information, the follow-up survey asked questions about specific properties that can help identify a project or collection associated with a 3D object and provide high-level information (an overview) of the project or collection, identify people with various roles (such as creator, contributor, stakeholder), and supply copyright information that can help establish access rights for a project or collection of 3D objects.

Implications

The metadata to establish management capabilities need to identify the overall project for discovery purposes. This allows 3D models or sets to be grouped and organized for discoverability by project and, if possible, permissions set based on applicable rules (e.g., copyright, embargo, etc.) if they can be applied to these same sets as a whole. This in turn helps establish levels of access for an entire group of 3D models, which then helps establish access to individual 3D models.

Determining copyright and licensing inherent in an object is important to properly managing a 3D model or collection of models. The information provided in chapter 5, "Copyright and Legal Issues Surrounding 3D Data," addressing rights and licensing should be referenced to help make appropriate determinations regarding what content is copyrightable and how 3D models and their associated data can be accessed and reused. Determining that a 3D model is protected under copyright may mean the files and creation information that make up that 3D model are not available for duplication and reuse, but derivatives or portions of the model might still be viewable. Additionally, metadata describing the 3D model can still be used for discovery and aggregation purposes, making the 3D model discoverable even if it is not available for viewing or download.

Inherent in providing access to different files is knowing something about those files or about the data as a whole to understand what they are and how best to supply access. This requires not only appropriate capture metadata, but also effective validation workflows to verify the datasets are complete and uncorrupted. Validation is also necessary to ensure that appropriate technical requirements are in place to ensure access to end users. For example, original scans or point cloud files can be large enough that online delivery is not feasible even if the rights determination is that they should be openly available. Even when using a digital repository or digital asset management system, providing different levels of access for the different files involved in a single 3D model can prove complicated.

Tools like MeshLab and Blender are available for evaluating technical aspects particular to 3D models, such as number of faces, number of points, number of slices, and spacing between slices. Digital repository systems such as MorphoSource from Duke University or the systems used to manage collections at the Smithsonian Institution have used tools such as these as part of their workflows, but there are no set minimum standards for 3D data preservation yet. These tools have been used so far only during upload and ingest processes and not as part of verification for a 3D model or an entire set or collection of 3D models. This means the effectiveness of using these tools to verify technical information at scale for multiple models or an entire collection of 3D models as a preservation activity is unknown. Scripted options for evaluating batches of 3D models are possible (working with a command-line tool such as meshlabserver, for example), but a minimum standard for 3D models is needed for the Manage stage of the digital asset life cycle so that ingested models can be approved and accessed. Establishing internal minimum standards and using tools that can extract this kind of metadata will help to organize a collection of 3D models and determine how best to provide access to an individual 3D model.

Recommendations

The survey results show that there is some preference for certain types of project-level information and fields: project name; project identifier; project date, description/abstract, file/data types, and subjects; project authors/creators and stakeholders/contributors; and project copyright information. In the area of management, the fields identified are not particular or specialized for 3D models or data. This is a benefit to managing 3D collections in that the types of information needed and activities performed are similar to other digital collection management needs and processes and might be compatible with available content management system software. Because managing access by users or by groups to an entire collection or dataset is generally a necessary part of managing any type of digital collection, this sort of functionality is already provided in many content management systems, such as institutional repositories or digital repository systems.

If 3D models can be managed as sets or collections, copyright management could occur at the set or collection level, applying copyright status that then indicates appropriate access levels to set for a grouping of 3D models. This might not always be effective if individual 3D models have rights maintained separately or have rights connected to a physical object used to create a 3D model, so it is worth considering if groups of 3D models can have the same copyright applied and if the management software includes that capability. If different resolution derivatives of 3D models require different access levels (e.g., an access level derivative is openly available for viewing, but the files and creation information that make up that derivative are restricted to only authorized users), that might require a different management model that can store the openly accessible 3D model with information that provides a pointer to a more restricted location for the files and creation information behind it. Please refer to chapter 5,"Copyright and Legal Issues Surrounding 3D Data," for a complete discussion of concepts concerning rights for a complete 3D digital object versus rights for the parts that make up that object. Ensure that copyright statements are understandable to end users accessing these objects and in any place where these models are shared or aggregated outside of the content management system by using standardized rights (such as those at RightsStatements.org) and licensing information (such as Creative Commons).

Managing objects in a digital collection involves verifying that what was expected was received. This often includes checking technical metadata about the files themselves (i.e., filename, checksum, file format). The Library of Congress offers recommendations for metadata standards and protocols to use (although 3D is not included at this time) and annually updates a Recommended Formats Statement that should include 3D soon.[8] Open-source communities working on digital collection management also offer good resources. The Samvera Community offers recommendations for baseline technical metadata that should accompany any digital media file, including 3D models.[9] The Smithsonian Institution offers a 3D metadata model defining technical metadata fields in use there, as do the Archaeology Data Service/Digital Antiquity Guides to Good Practice.[10] Different types of digital objects will also have technical information specific to that object type that can be used to check the item being received. Although no preservation standards exist at this time, 3D models have similar technical features that can be used to verify that the minimum standards are met for that type of digital object (such as scale, number of faces, number of points, number of slices, spacing between slices).

Based on survey responses and additional research, recommendations for Good reflect file organization that does not require a digital asset management system and supplying accompanying metadata in associated text files. Recommendations for Better reflect use of a digital asset management system that supports user role definition and access features along with a way to define projects for grouping models and storing metadata about those models. Recommendations for Best reflect implementing standards for associated metadata when possible.

Metadata Recommendations: Manage

Good

Group 3D models together as project or collection for access and management purposes.

Include following project level data

- project name
- project identifier
- project date
- description/abstract
- file/data types
- subjects
- project authors/creators
- stakeholders/contributors
- project copyright information

Supply copyright information at highest level possible (project/collection/object/file).

Using available tools,* document the following properties from files in a structured file type such as a README.txt file.

File-specific

- filename
- checksum
- file format
- file size
- file creation date
- file modified date
- file version

3D model–specific

- scale
- number of vertices
- geometry type†
- number of faces

---

* Examples of available tools include DROID (https://www.nationalarchives.gov.uk/information-management/manage-information/preserving-digital-records/droid/) and ffprobe (https://ffmpeg.org/ffprobe.html), but others are available to also do this work.

† For example, polygonal quads, polygonal tris, polygonal ngons.

- number of points

CT data–specific

- number of slices
- spacing between slices

Better

Use digital repository/digital asset management system with access capabilities to define permissions and make project/collection discoverable.

Store project, file, and 3D model properties listed above with 3D models in digital repository/digital asset management system.

Best

Use RightsStatements.org statements and Creative Commons licenses to define standardized access and reuse levels.

Use standardized metadata properties to define project, file, and 3D model properties listed above.

Distribute and Publish

The distribution and publication stage encompasses the licensing, sharing, and dissemination of 3D data. The metadata associated with this stage are essential for discovery and, to some degree, will depend on the intended use. They should be human-readable and machine-actionable so that data can be identified and located via a **persistent identifier, preferably a GUPRI.** Recommendations are driven by distributor type, often related to the user profile (e.g., casual, academic, etc.).

Stakeholders

Distribution metadata are first and foremost aimed at end users. As a result, publishers, websites or **aggregators**, and repositories are all stakeholders as they attempt to support access and reuse. Additional stakeholders include creators looking to track the reuse of their work, those looking to determine the trustworthiness and source of 3D data, and those looking to identify additional work from relevant creators. Additionally, 3D data are used in a variety of industries and in various applications. The science, technology, engineering, and mathematics fields and related industries (medical, automotive, meteorological, etc.), as well as the arts and humanities (anthropology, art history, psychology, etc.), are among the many disciplines and professions utilizing 3D data. Each of these stakeholders has differing distribution and metadata requirements reflective of the norms and expectations around reuse within their particular discipline or profession. In response to this reality, there are variations in existing commercial, academic, professional, and recreational distribution platforms to support different types of metadata. However, there is plenty of room for further development, especially as 3D data formats and their uses continue to evolve.

Survey

Responses to our first survey, along with ensuing research, indicated that many of the metadata requirements for the distribution (or publication) phase of the 3D data life cycle are not 3D-specific. Metadata supporting fair use and discoverability, including copyright information and persistent identifiers, are already fairly standard for other digital formats like digitized photos and documents.* Our second survey looked to determine if any of these standard metadata were not applicable to stakeholders, while also providing the opportunity to share additional metadata that were not already represented in the survey. We asked stakeholders whether they used or would use DOI/PID/location, copyright, embargo, or citation information. None were listed as undesired. There were also no additional fields provided. However, it must be considered that our survey responses retain the biases of galleries, libraries, archives, and museums.

Implications

As previously stated, distribution of 3D data requires metadata that support copyright, controlled access, and the measurement of impact. These typically take the form of persistent identifiers, embargoes, and copyright statements, all commonly used in other forms of publishing. However, some additional needs that should be considered are distribution derivative and accessibility metadata. Derivatives for the purpose of general distribution should have metadata generated and applied at the point of creation. Whether generated by a digital asset management system or created manually, the derivative should have the same information recommended in this chapter for the originating dataset. However, the source information would point to the originating dataset and the originating collection, item, or material the derivative represents in order to provide context to the derivative's creation and purpose. Accessibility metadata should include fields for alt text, long descriptions, and information on the location of, or means of requesting, accessible alternative formats. This poses a challenge in systems that do not support these features and should be considered by stakeholders when deciding between creation, hosting, distribution, and preservation platforms, as well as metadata schemata.

When it comes to distribution platforms that support both the viewing of content and the more complex metadata demands of certain disciplines and professions, affordable options are limited. Distribution demands for casual, nonprofessional use will be different from the demands of industries and disciplines where dataset **accuracy** can

---

* Dublin Core (https://www.dublincore.org/specifications/dublin-core/), MODS (http://www.loc.gov/standards/mods/mods-overview.html), RDA (https://www.rdatoolkit.org/about), and other schemata have all been used with MARC (https://www.loc.gov/marc/) and other encoding standards to create metadata records for both digitized and born-digital audio, video, photographs, etc. Aggregators such as DPLA (https://dp.la/) and Europeana (https://www.europeana.eu/portal/en) have multiple examples of these types of records from various organizations.

be the determining factor in safety and crucial decision-making. 3D models, virtual reconstructions, aerial photogrammetric scans, and CAT scans are among the various formats that can be shared in digital repositories, commercial platforms, journals, and other distribution platforms. Each of these formats and its associated function will differ in the ways it handles copyright, access control, and tracking due to different legal requirements and desired reuse (see chapter 5, "Copyright and Legal Issues Surrounding 3D Data," for more information). However, while distribution metadata needs will vary from sector to sector, the limited availability of affordable hosting means that available platforms, like Sketchfab, often have a mixture of model types, yet limited metadata options, complicating metadata application and retention for data with more robust metadata needs. There are more specialized distribution platforms, such as the NIH 3D Print Exchange, which specializes in hosting medical 3D printing models for download, but these are meant to accommodate a very specific subset of 3D data for a very specific subset of use. Datasets utilized for academic, cultural heritage, or medical purposes that do not have printing as their sole purpose or at the core of their use will require very specific display, manipulation, and granular metadata support.

Recommendations

The recommendations in this section are based on the intended use of the dataset in question. Recommendations for Good practice will be based upon reuse of datasets for recreational virtual reality, 3D printing, or other casual, informal purposes. Recommendations for Better practice will be based upon datasets designed to support research by serving an illustrative or reference function, like historical recreations. Recommendations for Best practice will be based on datasets designed to be used as research data for fields and industries that depend upon accurate standards for reuse, such as meteorological and safety simulations and CAT scans in the medical field. All metadata from the previous sections should be carried into this one.

Metadata Recommendations: Distribute and Publish

Good

For casual use, users need to understand the format and intended use of the model. Intended use will usually dictate whether a model is suitable for printing, virtual reality, or other purposes. For an example, see Sketchfab.[11]

- system unique identifier
- authors/creators
- file type
- file size
- geometry information (quads, triangles, etc.)
- vertices
- textures/materials

- description
- accessibility information
  - alt text
  - long description
  - accessible format/information location (alternative accessible format or accessibility policy information)

Better

3D data meant to act as an illustration in scholarship should include the previous data in addition to the recommended fields below. For an example, see the NIH 3D Print Exchange.[12]

- link to original **source material** OR formal name and information of original item (scientific name, formal name, etc.)
- all processing done (textures added, artistic liberties taken)
- persistent identifier, preferably a GUPRI
- associated works
- citation information

Best

Datasets designed to be used as research data for fields and industries that depend upon accurate standards for reuse should include the previous data in addition to the following fields. For an example, see MorphoSource.[13]

- detailed capture metadata (see metadata recommendations in the Create section)
- project information
  - funding information
  - affiliated organizations
- technician information
- checksum

Access and Reuse

Metadata in the Access and Reuse stage are associated with the circulation, consumption, citation, and use of existing 3D data by end users and non-practitioners. The purpose of these data is twofold. In one case, these data are to facilitate ease of access, use, reuse, and interoperability (ideally) of systems once the data are identified. These data also allow for appropriate citation, underscoring the integrity, authenticity, and **provenance** of the 3D data. A user should be able to decipher, from these metadata, whether the 3D data are of use to them. The Access and Reuse stage of the digital asset life cycle is key in the continued vitality of any information source, including 3D data.

Stakeholders

Stakeholders in the Access and Reuse process come from all stages of the life cycle and have vastly different priorities when it comes to accessing and consuming 3D data. First and foremost, content creators of all disciplines—the sciences, cultural heritage preservation, the arts, or fields that are yet to be developed—are invested in how they can facilitate the reuse of their data, regardless of how they were created. Both to preserve research and for the benefit of those consuming data later, it is necessary to create metadata recording the creation of the work, as mentioned earlier in this chapter, in a way that is both archival and understandable. These metadata should be filtered and keyworded in ways that aid in the discoverability of a given dataset. The academic community is still formulating the research norms and preservation standards for 3D data using foundations such as the FAIR (findability, accessibility, interoperability, reuse) data principles.[14] While it may not seem immediately useful, given research patterns of other data, it is conceivable that researchers may find it useful to also be able to search larger groups of data by creation mode, source materials, related derivatives, or object size or scale in order to better serve their research. Similarly, holding entities of 3D data take a vested interest in the accessibility and discoverability of the items and data within their catalogs in order to better facilitate the access and reuse of that information.

On the path to publication, researchers in higher education and industry need reliable discoverability to find 3D data and models. They also need verifiable provenance and tools for recreation of both data and models to ensure the validity and accuracy of reproducible data. This may include a sort of intellectual chain of custody listing the researchers who have worked on an object and the methods used, as well as recommended software or viewers to access the item. In addition to this, the academic community will need a means to cite the 3D data being utilized as no sufficient standards for doing so currently exist, regardless of whether those data are open-access or rights-restricted. As these researchers may or may not be 3D practitioners themselves, consideration will need to be given to common language, dataset types, and creation formats, as well as standardized citation formats. In addition to general standardized language, this consideration should extend to discoverability for information specialists and research assistants aiding those researchers. Information specialists and librarians who are not 3D specialists will need to be able to assist patrons of all types in finding data through common web and database searches, as well as integrated library systems (ILSs) and discovery layers provided by a variety of vendors.

Beyond creators and researchers, a wide variety of consumers within the general public have interest in accessing 3D data. Some may be artists or enthusiasts, looking to use 3D data for personal enrichment or to augment personal projects. Teachers in primary and secondary education may look to 3D data to augment concepts taught through classroom instruction. Students in those classes may look to them to aid in class projects or to fuel further study in personal interests. Furthermore, with the advent of easily accessible resources, such as Thingiverse, MorphoSource, and the Smithsonian Digitization Program Office, and the increasing access to 3D printers, individuals from various backgrounds may have an interest in discovering 3D models for self-guided learning and 3D printing. In general, these consumers are less concerned about the ability to recreate the final project than the ease of finding the item using natural language or keyword searching.[15]

Survey

As universal standards involving the access, use, and resource sharing of 3D data and models have not yet been codified, survey responses regarding access to said items were limited to a record of downloads and use count as well as, potentially, fields to record who may have edited or processed the available files. While programs like SHAPES, housed out of the Texas Tech University Libraries, are working to establish resource sharing standards, adoption is slow. There has been more progress in open-access resources like MorphoSource and Thingiverse but access metadata for these databases are also currently limited to download counts.* As resource sharing of 3D data and models evolves, it is likely the related metadata needs will as well. With this in mind, current recommendations for Access and Reuse will focus on the metadata needed to locate and access a given 3D dataset or model, rather than that generated during the access process. As community standards for the access and sharing of these resources develop, necessary metadata beyond page views and download numbers may develop as well.

Implications

Data that cannot be found cannot be used and, therefore, have no longevity. In order for 3D data to continue being utilized, they need to be discoverable and accessible by both the academic community and the general public. As resource sharing for these formats is still developing, it remains to be seen what sort of circulation and access data should be preserved, but the topic should definitely be considered as the prevalence of 3D data increases. Though 3D data creation is far from new, access processes for 3D data are still developing. This will likely continue as the academic standards for previously nontraditional information formats are established. As the availability and predominance of 3D data increase, so too will the pathways and procedures needed to access them. In general, a record of who is accessing 3D data will not affect the long-term veracity of said data, but it may be helpful to content creators and archivists, along the lines of existing circulation statistics, download counts, and page views. These data may be collected in formats already in use within integrated library systems and website analytics. For future usage of 3D data, it may be worth considering whether different modes of access should be recorded and whether that record may differ from that gathered for more commonly sought digital objects, such as journal articles and images held in online databases.

---

* While MorphoSource and Thingiverse have different funding origins—MorphoSource is housed at Duke University, and Thingiverse is run by corporate entity MakerBot—both repositories serve as sources of 3D data and models that are readily and freely available to the general public.

Recommendations (3D-Specific and General)

In general, the recommendations listed below were conceived to help stakeholders at different levels of 3D creation and practice. As the potential consumers of 3D data come from a variety of areas of expertise and backgrounds, care should be taken throughout to utilize common and standardized terminology wherever possible, ensuring the usage of general keywords along with more specialized and technical creation data in order to make the end product more discoverable by both 3D specialists and the general public.

Good practices, in this case, are what will be the most useful for the casual user who will need to find 3D data or a 3D model that suits their needs. At bare minimum, there should be sufficient citation and keyword metadata for an individual not related to the original creation to be able to find the data using common search parameters in search engines or library discovery layers, as discussed in previous sections of this chapter. These may include but are not limited to original item data, method of creation, creation or publication date, author or creator information, and general subject terms including such information as area of study, general time period in history, reason for creation,[†] possible derivative chains, and where the data can be accessed. For example, Digital Morphology (DigiMorph) recommends the following citation format:

> Author/creator, [creation] year, "title of dataset," name of the holding collection. Access date and accessed URL.[‡]

As many casual users are currently seeking to consume 3D data through specific modes, file intent or purpose is also important at this level. This currently includes printability on a 3D printer (potentially including a necessary derivative or accommodation to afford this), usability within virtual reality, and access through a virtual viewer, either web-based or software-driven.

Better practices apply to most researchers, many of whom are not 3D practitioners themselves but are often familiar with common academic processes and information-gathering behaviors. To best assist these users, whenever possible, persistent identifiers such as digital object identifiers (DOIs) or Archival Resource Keys (ARKs) should be assigned to ensure a level of specificity in discovery on par with accession numbers like those utilized in library catalogs and large databases like WorldCat, a worldwide library database managed by the Online Computer Library Center (OCLC). These persistent identifiers allow consumers of all types to find a specific item with a minimal amount of information. Many of these users may also be looking for 3D data or models based on basic creation data or mode of creation—potentially photogrammetry, lidar capture, or CAD—but may not have the skills to create these data themselves or to parse more advanced capture data. Basic identifiers for available derivatives may also

† Some examples may include archival rendering, teaching or study aids, historical recreation based on available sources, etc.

‡ See examples of citation format at http://www.digimorph.org.

be useful to distinguish between versions, such as the articulated woolly mammoth created by model designer Teraoka Gensyou from the Smithsonian Digitization model of a woolly mammoth skeleton.[16]

Best practices for access and reuse of data may be the most useful to 3D specialists and information- or resource-sharing specialists assisting researchers in locating 3D data, as this level involves the usage of highly specific data not necessarily relevant to lower-intensity users. As it is not currently possible to stream high-quality 3D data over the web due to the significant file sizes, web-based viewers of 3D data compress those data out of necessity. Resolution and compression data, as well as adaptive and scalable considerations, may be key in how certain 3D specialists interact with a given dataset and verify that it will meet their needs and may give integral insight into how those data may be reused. Users at this level may find it useful to search by detailed collection data relevant to the raw data captured, including but not limited to information about photogrammetric sets, lidar settings, or poly count. Therefore, it is worth considering whether there should be multiple levels of persistent identifiers or whether there should be a single resolver for all iterations of a given set of 3D data. While a DOI for 2D data is used for accessibility, best practices for DOIs and access data for 3D data may need to go farther, potentially indicating compression rates, creation history, provenance, and more.

Metadata Recommendations: Access and Reuse

Good

Prioritize discoverability by non-practitioners, providing enough data to cite the model or dataset:

- authors/creators
- creation/publication year
- title of dataset/model
- name of the holding collection (if applicable)
- original item data (if generated from a scan of an object)
- source material
- file type
- file intent or purpose (printability vs. virtual reality vs. virtual view)

Better

All of the above data plus

- persistent identifier (preferably a GUPRI) such as DOI or ARK for master dataset
- basic creation data, mode of creation (photogrammetry, lidar, manual modeling, etc.)
- basic identifiers for available derivatives

Best

All of the above data plus

- detailed collection data relevant to raw data (see the section "Create" in this chapter)
- synthesis method to create master model

- delivery methods for available derivatives, including relevant resolution and compression data

Evaluate whether specific derivatives or delivery methods need separate persistent identifiers.

Archive

Metadata in the Archive stage are associated with the maintenance and preservation of 3D data. It is important to document the file formats, creation environments, and required software and hardware to ensure long-term access and reusability. Archival preservation of 3D data will often require consultation with a **digital preservation** specialist at the beginning stages of the data creation process.

Stakeholders

Archivists and repositories should take stock of these requirements and recommendations as they manage much of the transition between distribution and preservation. Creators donating or submitting their work should also consider the **preservation metadata** of their data. Finally, those looking to reuse datasets that are older or in unfamiliar formats will require much of the information recommended in this section to be able to gauge how, or whether or not, they will be able to access preserved data.

Survey

To address gaps in preservation metadata practices, this chapter uses the *PREMIS Data Dictionary* as a guide.[17] Many of the metadata specifications in PREMIS show how heavily dependent preservation metadata are on sufficient creation metadata. Based on our first survey, we found that practices, uses, and workflows for 3D data varied greatly, and we had no real responses indicating developed preservation metadata (as described later in this section's recommendations section). This complicates the development of digital preservation workflows, as each type of 3D data needed differing levels of information to support reuse. Based on this information, we focused on better understanding existing file formats and technical environment requirements (hardware and software needed to use 3D datasets). Regarding existing formats, respondents indicated a number of file formats used in various projects of varying sizes, scopes, and purposes. There were twenty different formats listed within our limited survey responses alone, with each of these formats experiencing varying levels of support, maintenance, and platform compatibility. Not all formats were 3D-specific; .pdf, .ppt, .txt, and .tif were among the non-3D file formats listed. Regarding environment, variation once again characterized the responses. Memory requirements and modes of storage vary with method of creation, project size, project type, and the creator's access to resources. Memory ranging from megabytes to multiple gigabytes and cloud storage versus local storage were present in our limited samples across the two surveys.

Implications

The presence of multiple file types makes sense when considering the state of digital preservation. The work of preserving digital materials is still developing and ever-changing. Aside from basic storage and maintenance questions like those around server space and data degradation, there are additional concerns regarding the rapid change of preferred, utilized formats. Formats from established companies, like Adobe and Microsoft, have the wide use required to ensure continued demand for access. They also have the support and funding required to meet that demand by supporting continued access to their older formats via conversion options. Their wide use also makes it profitable, or simply useful, for independent developers to invest in the development of tools that support the conversion and viewing of obsolete formats. There is also more precedent in preservation efforts for these older, established products (Word, Illustrator, PowerPoint, etc.). With that in mind, metadata become increasingly important because the data are removed from their original context. This means that the metadata will need to compensate for this loss of context, particularly in the case of datasets where reuse requires the ability to replicate creation. Particular attention will have to be paid to retaining detailed creation metadata.

Aside from file type and technical and environmental requirements, considerations around data derivatives for publication and sharing were discussed. These derivatives should be treated as unique datasets with their own creation metadata and additional fields linking them to, and describing their relationship with, the dataset upon which they are based.

Recommendations (3D-Specific and General)

Before any formal metadata are created, it is recommended that creators work with an archivist to make early decisions regarding the preservation needs of their work. This chapter recommends the utilization of PREMIS. The exact PREMIS entities and semantic units used will depend on the type of 3D data as well as their intended purpose, as PREMIS does allow a high degree of flexibility that would support the more granular metadata models utilized by institutions like the Smithsonian.[18] Special attention should be paid to the Special Topics portion of PREMIS. Metadata around environment and object characteristics are particularly important to 3D, especially if data have been compressed or are platform- or operating system–specific. The recommended metadata from the "Create" and "Manage" sections of this chapter should be carried over into this one. Recommendations in this section will be based upon the infrastructure and support available to stakeholders. Digital preservation typically requires a digital preservation specialist, server space, and a system that supports the metadata recommended in this chapter. As a result, recommendations under Good will be for those utilizing a proprietary content management system with limited metadata flexibility. Recommendations under Better will be based on access to content management systems that support 3D but do not specialize in it. Recommendations under Best will be based on access to a content management system designed to host 3D data.

Metadata Recommendations: Archive

Good

In order to capture and retain metadata in content management systems that do not support the more granular metadata required for reuse, it is recommended that unsupported essential metadata be kept in an accompanying easily supported file format such as .txt or .csv. It is also recommended that metadata fields be encoded into a commonly used markup language (like XML) when possible. This will make it easier to crosswalk and edit metadata should ingestion into a more supportive repository occur. At bare minimum, in addition to the original Create and Manage metadata, archivists should include

- repository ID
- *objectIdentifier (ISBN, DOI, etc.)
- *objectCategory
- *significantProperties (not required, but strongly recommended that the reader keep this unit in mind)
- *objectCharacteristics
- *formatDesignation
- *creatingApplicationName
- *creatingApplicationVersion
- *contentLocationValue
- related objects
- event information (metadata on compression, edits, or other actions taken)
- rights information
- accessibility information
    - alt text
    - long description
    - accessible format and information location (alternative accessible format or accessibility policy information)

Items marked with an asterisk are direct semantic units in PREMIS. Some of these have multiple parts. That is apparent when reading the PREMIS entity description.

Better/Best

The capabilities of the content management system being used will heavily determine the feasibility of incorporating additional metadata entities. Systems that can automatically capture metadata regarding event information and file format will be necessary to meet the more extensive requirements in PREMIS for 3D.

- more extensive agent metadata (see Agent Entity section of PREMIS)
- more extensive environment, as well as other special topics metadata (see Special Topics section of PREMIS)
- separate event records for tracking preservation activities

Gap Analysis/Future Work

As 3D creation, storage, and usage practices are still maturing, much of what is discussed in this chapter will continue to evolve as the practices solidify. Both software and hardware involved in the creation of 3D data are changing rapidly, which makes uniform suggestions difficult. The situation is further complicated because there is evolution on two fronts, the technology itself and its use. Both scholastic and industry approaches to the creation and use of 3D data are changing heavily. Initiatives like the Library of Congress's Born to Be 3D effort, LIB3DVR, and CS3DP are working to both contextualize and focus academic approaches for 3D data, while organizations like MorphoSource, Cultural Heritage Imaging, IIIF, and the Smithsonian are working to clarify both 3D creation standards and ease of access. While industry is not represented in the survey results for this chapter, it is important to note that there was industry participation at the second CS3DP forum, and industry efforts will impact how 3D data management practices progress.

3D models offer a powerful and information-rich way to document the physical condition and appearance of objects and environment as they existed in a moment in time. Aside from being a visual and metrological record, 3D models provide a scaffold for deeper documentation and compelling storytelling about the subjects they describe. 3D models can be annotated in a variety of ways, such as connecting text and other media to discrete points or regions of interest (either on the model's surface or volumetrically within the model's world space). When the primary goal of a 3D model is to document its subject, this information can be as important as the model itself. Citations and annotations of said data are still evolving and will continue to do so as scholars find more ways to take advantage of the opportunities this format presents. As the technology develops, so does the creativity in the community's use of it.

As 3D expands in its use, there are additional legal and ethical concerns that need to be addressed as well. One in particular is accessibility for disabled users, especially when 3D is being used in educational settings. While there are a myriad of issues that need to be addressed in order for 3D to be accessible, this chapter has shown the bare minimum metadata that should accompany 3D models when possible (alt text, long description, alternative format and accessibility information). However, these minimal metadata do not support equitable use and will need to be expanded upon and diversified.* As is the case with other categories of metadata, accessibility metadata needs

* Equitable use is a principle of universal design. In order to be considered as having provided equitable use, VR must (1a) provide the same means of use for all users: identical whenever possible, equivalent when not; (1b) avoid segregating or stigmatizing any users; (1c) make provisions for privacy, security, and safety equally available to all users; and (1d) make the design appealing to all users. (Centre for Excellence in Universal Design, "The 7 Principles: Principle 1: Equitable Use," accessed January 7, 2022, https://universaldesign.ie/what-is-universal-design/the-7-principles/the-7-principles.html#p1.)

will vary in different disciplines and industries, and metadata will need to convey both purpose and content. For example, functionality with screen readers will require more automated and nuanced orientation information. Having a version of 3D data that can be 3D printed, viewed, and manipulated in a browser; examined in layers so that textures and other features can be described; or examined and communicated as raw data are all potential approaches that will also require more effective metadata. This also means that there will need to be effective validation tools and audit workflows in place. Each field and discipline will have to work to determine what equitable use looks like for its disabled members and begin the process of determining what technology and metadata are needed. In order to truly support reuse for disabled users, avoid costly remediation, and foster the necessary cultural change required for buy-in, these issues should be addressed sooner rather than later.

Conclusion: Summary Recommendations

Table 4.1 on the next page summarizes the metadata practices recommended at PIPs through each stage of the digital asset life cycle. There is no single metadata standard that encompasses all recommended fields or fills all the specific needs for 3D model metadata, but metadata standards such as Dublin Core, PREMIS, and VRA Core can address portions of the recommended metadata fields. In terms of academic research data, metadata recommendations should adhere to the basic principles of FAIRness.[19] Additionally, domain-specific work is in progress within groups such as Audubon Core, which may further provide 3D-specific field options in the future. Further community-driven work is needed, however, to develop and refine a metadata standard that is truly cross-disciplinary, cross-modal, and widely interoperable. These recommendations identify necessary categories of metadata that are required in the emerging adaptable, flexible standard.

TABLE 4.1
Aggregated table of Good/Better/Best recommendations

<table>
<tr><th></th><th>Good</th><th>Better</th><th>Best</th></tr>
<tr><td>Create</td><td>Document and use a folder structure and file naming convention to organize source data and 3D models for management purposes. For example, group data and models by digitization project, collection, etc.

Include the following information as structured data in README.txt files, CSV files, spreadsheet or similar.

Project level information
project name
project identifier
project date
description/abstract
project authors/creators
stakeholders/contributors
project rights information
Source data information, if a reality-capture or sources-based model
method of creation
creation date
information identifying real-world object or environment
creators
3D model information:
subjects
creator
geometry information
textures (what types of UV maps are available)
materials (If any material properties are applied to a model)</td><td>Include the following additional information as structured data in README.txt files, CSV files, spreadsheet, or similar.

Source information for reality-capture or sources-based models
persistent identifiers (preferably globally unique persistent and resolvable identifiers, or GUPRIs) for records of subject or sources
data source
georeference information if applicable
Source data creation information
capture device make and model
capture event details
3D model processing information
source data used
software used
Documentation of capture and processing workflows</td><td>Include the following additional information in README.txt files, CSV files, spreadsheet, or similar.

3D model processing information
detailed steps and log outputs
Use standardized metadata properties to define project, file, and 3D model properties listed under Good and Better.</td></tr>
</table>

TABLE 4.1
Aggregated table of Good/Better/Best recommendations

	Good	Better	Best
Manage	Group 3D models together as project or collection for access and management purposes. Include following project level data: • project name • project identifier • project date • description/abstract • file/data types • subjects • project authors/creators • stakeholders/contributors • project copyright information Supply copyright information at highest level possible (project/collection/object/file). Using available tools,[a] document the following properties from files in a structured file type such as a README.txt file. File-specific • filename • checksum • file format • file size • file creation date • file modified date • file version 3D model–specific • scale • number of vertices • geometry type[b] • number of faces • number of points CT data-specific • number of slices • spacing between slices	Use digital repository/digital asset management system with access capabilities to define permissions and make project/collection discoverable. Store project, file, and 3D model properties from Good column with 3D models in digital repository/digital asset management system.	Use RightsStatements.org statements and Creative Commons licenses to define standardized access and reuse levels. Use standardized metadata properties to define project, file, and 3D model properties listed in Good column.

TABLE 4.1
Aggregated table of Good/Better/Best recommendations

	Good	**Better**	**Best**
Distribute and Publish	For casual use, users need to understand the format and intended use of the model. Intended use will usually dictate whether a model is suitable for printing, virtual reality, or other purposes. For an example, see Sketchfab.[c] • system unique identifier • authors/creators • file type • file size • geometry information (quads, triangles, etc.) • vertices • textures/materials • description • accessibility information o alt text o long description o accessible format/information location (alternative accessible format or accessibility policy information)	3D data meant to act as an illustration in scholarship should include the data listed in the Good column in addition to the following recommended fields. For an example see the NIH 3D Print Exchange.[d] • link to original source material OR formal name and information of original item (scientific name, formal name, etc.) • all processing done (textures added, artistic liberties taken) • DOI/PID • associated works • citation information	Datasets designed to be used as research data for fields and industries that depend upon accurate standards for reuse should include the data listed in the Good and Better columns in addition to the following fields. For an example, see MorphoSource.[e] • detailed capture metadata (see recommendations in Create section) • project information o funding information o affiliated organizations • technician information • checksum
Access and Reuse	Prioritize discoverability by non-practitioners, providing enough data to cite the model or dataset: • authors/creators • creation/publication year • title of dataset/model • name of the holding collection (if applicable) • original item data (if generated from a scan of an object) • source material • file type • file intent or purpose (printability vs. virtual reality vs. virtual view)	All of the data listed in the Good column plus • persistent identifier (preferably a GUPRI) such as DOI or ARK for master dataset • basic creation data, mode of creation (photogrammetry, lidar, manual modeling, etc.) • basic identifiers for available derivatives	All of the data listed in the Good and Better columns plus • detailed collection data relevant to raw data (see recommendations in the Create section) • synthesis method to create master model • delivery methods for available derivatives, including relevant resolution and compression data Evaluate whether specific derivatives and/or delivery methods need separate persistent identifiers.

TABLE 4.1
Aggregated table of Good/Better/Best recommendations

	Good	Better	Best

Archive	In order to capture and retain metadata in content management systems that do not support the more granular metadata required for reuse, it is recommended that unsupported essential metadata be kept in an accompanying easily supported file format such as .txt or .csv. It is also recommended that metadata fields be encoded into a commonly used markup language (like XML) when possible. This will make it easier to crosswalk and edit metadata should ingestion into a more supportive repository occur. At bare minimum, in addition to the original Create and Manage metadata, archivists should include[f]
	• repository ID
	• *objectIdentifier (ISBN, DOI, etc.)
	• *objectCategory
	• *significantProperties (not required, but strongly recommended that the reader keep this unit in mind)
	• *objectCharacteristics
	• *formatDesignation
	• *creatingApplicationName
	• *creatingApplicationVersion
	• *contentLocationValue
	• related objects
	• event information (metadata on compression, edits, or other actions taken)
	• rights information
	• accessibility information
	  o alt text
	  o long description
	  o accessible format and information location (alternative accessible format or accessibility policy information)
		The capabilities of the content management system being used will heavily determine the feasibility of incorporating additional metadata entities. Systems that can automatically capture metadata regarding event information and file format will be necessary to meet the more extensive requirements in PREMIS for 3D.
		• more extensive agent metadata (see Agent Entity section of PREMIS)
		• more extensive environment, as well as other special topics metadata (see Special Topics section of PREMIS
		• separate event records for tracking preservation activities
			See Better column

a. Examples of available tools include DROID (https://www.nationalarchives.gov.uk/information-management/manage-information/preserving-digital-records/droid/) and FFProbe (https://ffmpeg.org/ffprobe.html) but there are others available to also do this work.

b. For example, polygonal quads, polygonal tris, polygonal ngons.

c. Sketchfab, https://sketchfab.com/.

d. NIH 3D Print Exchange, https://3dprint.nih.gov/.

e. MorphoSource, https://www.morphosource.org/.

f. Items marked with an asterisk are direct semantic units in PREMIS. Some of these have multiple parts. That is apparent when reading the PREMIS entity description.

APPENDIX
Survey 2 Results

CS3DP Metadata Working Group Feedback Form

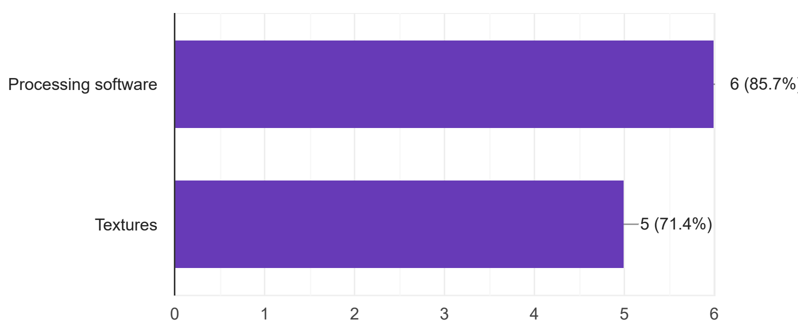
10 responses

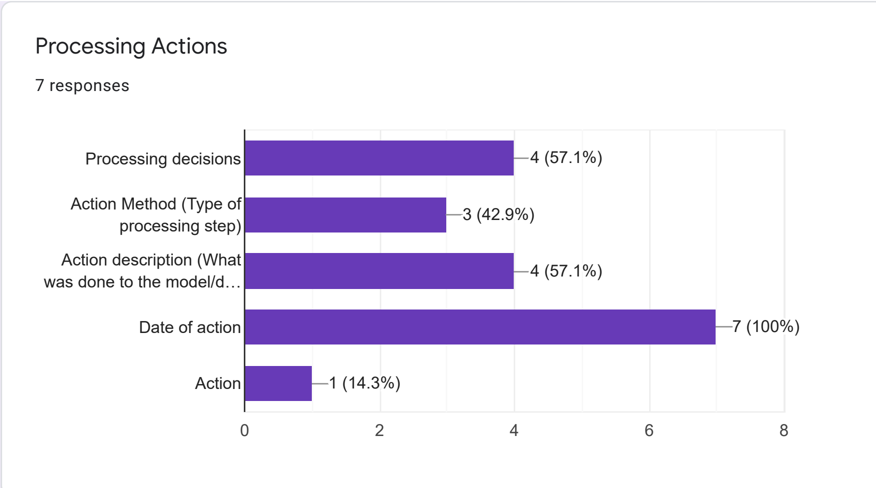
Add your User Stories!

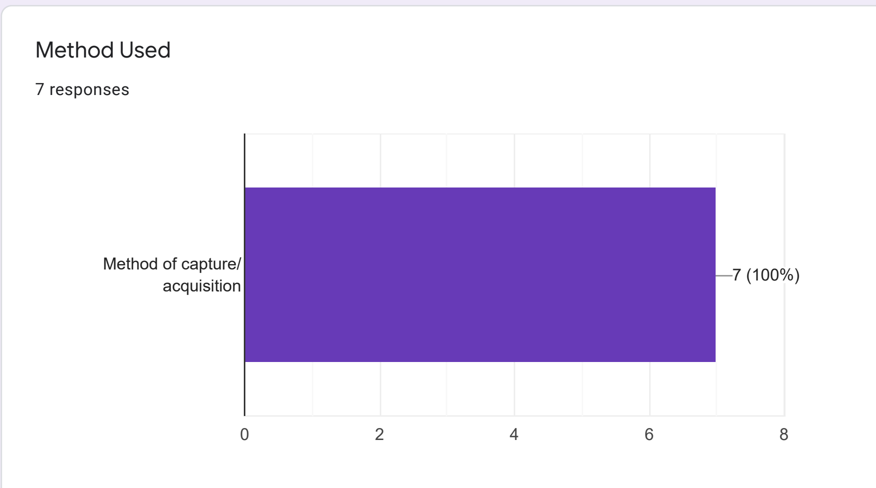
Create - Paradata (Pre-processing/Processing)

Resources

7 responses

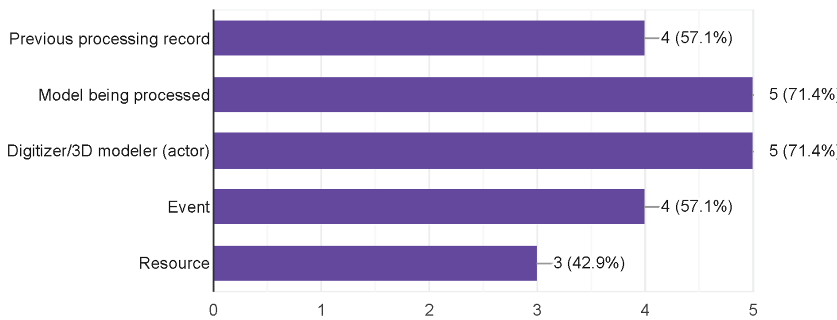



Processing Actions
7 responses
Processing decisions
Action Method (Type of processing step)
Action description (What was done to the model/d…
Date of action
Action
4 (57.1%)
3 (42.9%)
4 (57.1%)
7 (100%)
1 (14.3%)
0
2
4
6
8


Method Used
7 responses
Method of capture/ acquisition
7 (100%)
0
2
4
6
8

Linked Fields

7 responses



What other fields do you use or would you use within the Create - Paradata portion of the Digital Asset Lifecycle?

2 responses

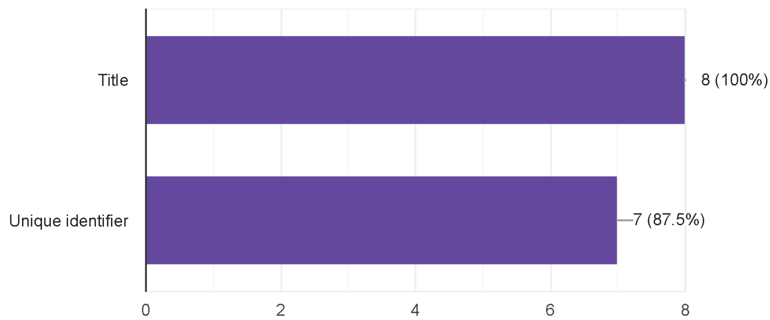
Material Type

Processing hardware

Create - Model/Data

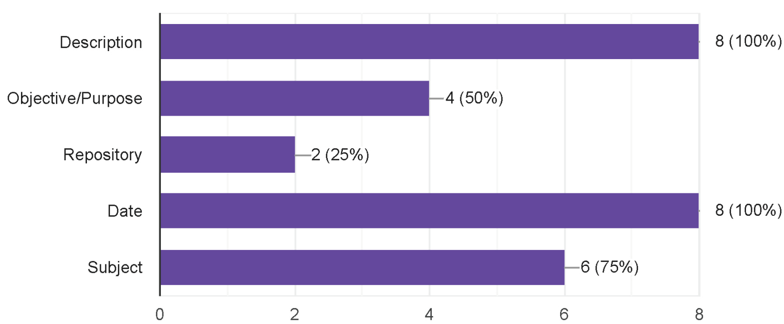
Model Identifying Information

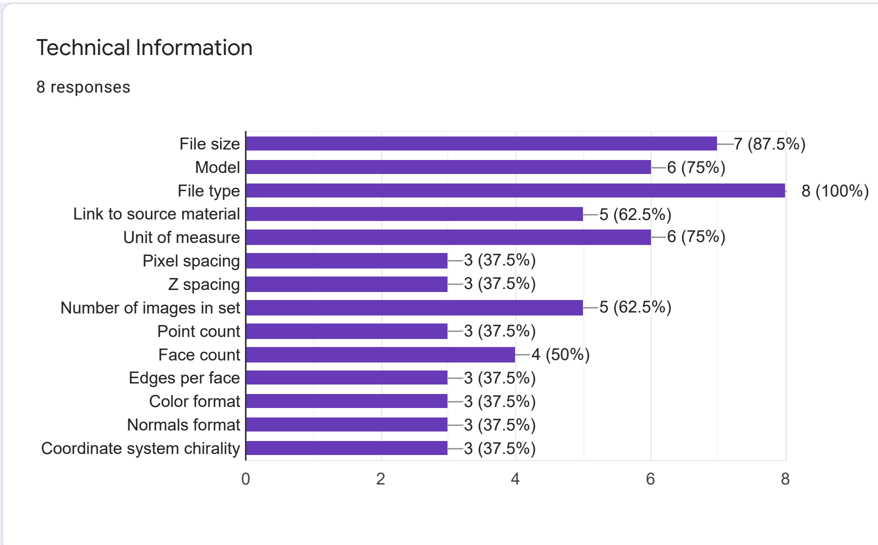
8 responses

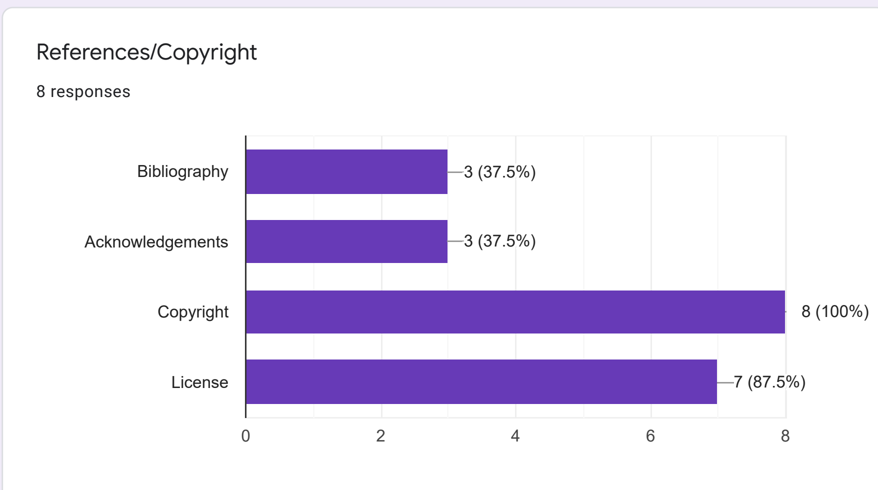


Overview Information

8 responses




Technical Information
8 responses
File size
Model
File type
Link to source material
Unit of measure
Pixel spacing
Z spacing
Number of images in set
Point count
Face count
Edges per face
Color format
Normals format
Coordinate system chirality
7 (87.5%)
6 (75%)
8 (100%)
5 (62.5%)
6 (75%)
3 (37.5%)
3 (37.5%)
5 (62.5%)
3 (37.5%)
4 (50%)
3 (37.5%)
3 (37.5%)
3 (37.5%)
3 (37.5%)
0
2
4
6
8


References/Copyright
8 responses
Bibliography
Acknowledgements
Copyright
License
3 (37.5%)
3 (37.5%)
8 (100%)
7 (87.5%)
0
2
4
6
8

What other fields do you use or would you use within the Create - Model/Data portion of the Digital Asset Lifecycle?

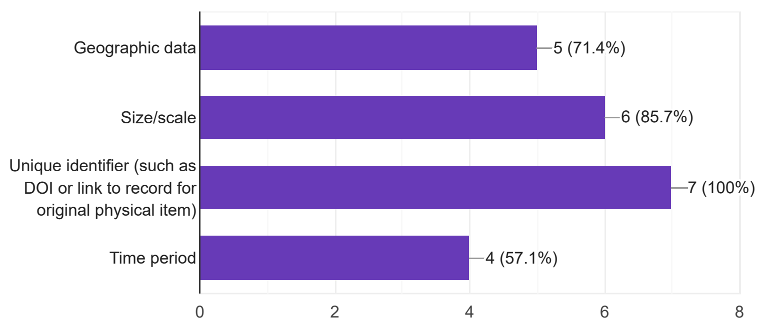
2 responses

Scale, Amount of Material Used, Cost of Model

We have uploaded VR files created by students for a course so we also included information on the students (program of study, year in school) and the course (course name/number, faculty teaching).

Create - Original Item

7 responses

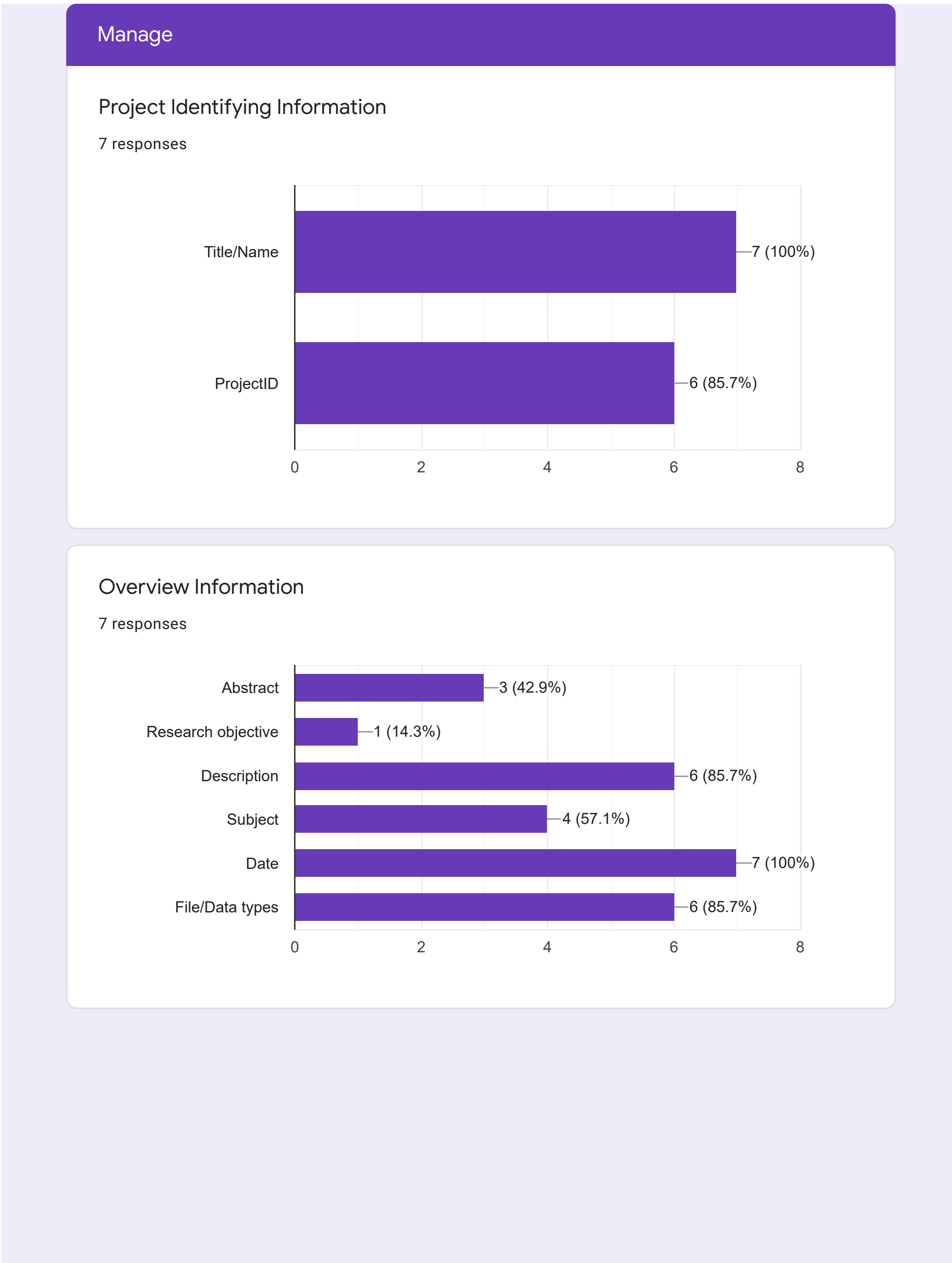


What other fields do you use or would you use within the Create - Original Item portion of the Digital Asset Lifecycle?

2 responses

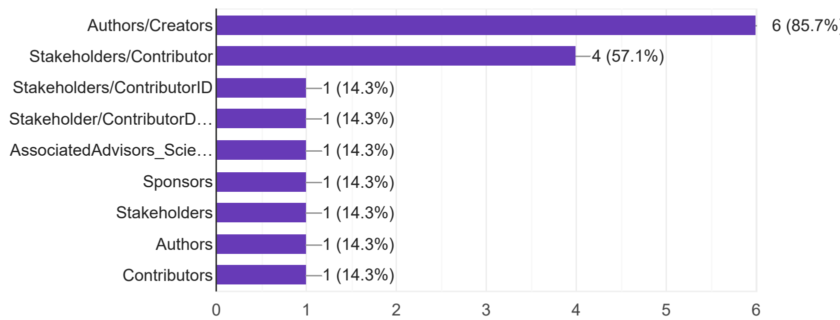
Acquisition conditions

Description - A general description which can include geographic data or a time period, but can also include other types of information about the physical item(s).


Manage
Project Identifying Information
7 responses
Title/Name
7 (100%)
ProjectID
6 (85.7%)
0
2
4
6
8
Overview Information
7 responses
Abstract
3 (42.9%)
Research objective
1 (14.3%)
Description
6 (85.7%)
Subject
4 (57.1%)
Date
7 (100%)
File/Data types
6 (85.7%)
0
2
4
6
8

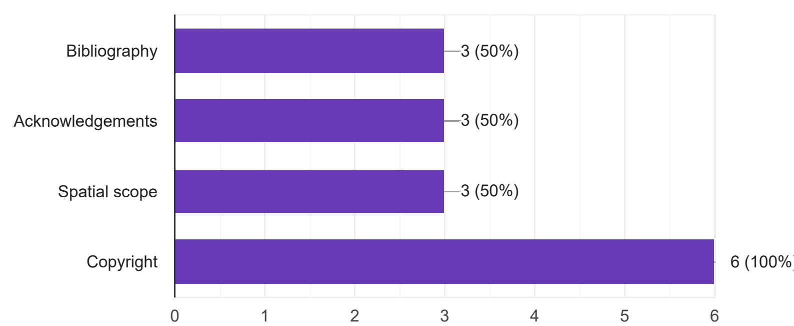
Stakeholder/Creator/Contributor Information

7 responses



References

6 responses



What fields or information do you store for annotations on 3D models or 3D datasets?

1 response

Short descriptions

What other fields do you use or would you use within the Manage portion of the Digital Asset Lifecycle?

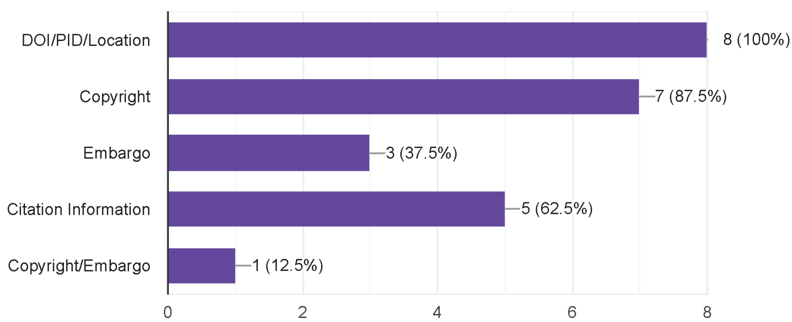
0 responses

No responses yet for this question.

Distribute

General

8 responses



Please share any other, or expand upon fields you use for distributing 3D models/data sets.

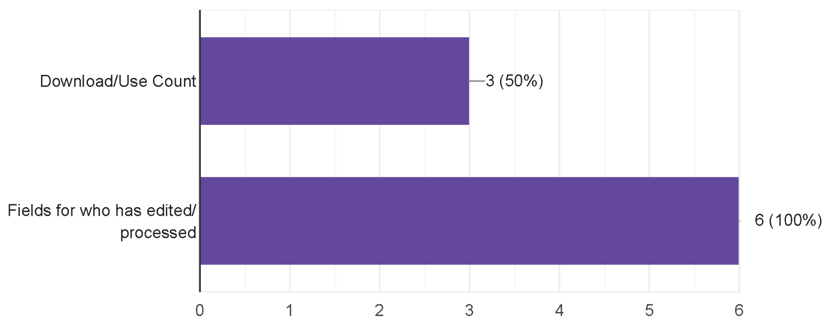
0 responses

No responses yet for this question.

Retrieve/Re-use

General

6 responses



Please share any other, or expand upon fields you use for retrieving/re-using 3D models/data sets.

0 responses

No responses yet for this question.

Archive

What file formats are you using for long-term storage?

7 responses

- stl, obj, jpg, png, psd, psx, mp4, mov
- ASCII text (point clouds), TIF (images), OBJ (meshes, objects), e57 (point clouds)
- OBJ, VRML
- Stl
- e57,pts,obj,txt
- .zip, .dll, .exe, .avi, .pdf, .ppt
- obj

What specific technical/environment requirements (memory allocation, etc) are needed?

3 responses

- 1-200MB for 3d model, much larger space for image data sets for photogrammetry (100MB - multiple GB)
- Cloud Storage
- not positive, so far maybe 1gb total

What other fields do you use or would you use within the Archive portion of the Digital Asset Lifecycle?

0 responses

No responses yet for this question.

Additional Feedback

Please provide any authorities you use for creating identifiers (for project, model/data, or other types of identifiers).

2 responses

DOI

EZID, DOI

Please provide any additional feedback regarding the findings of the CS3DP Metadata Working Group.

1 response

Be very realistic with how many fields you want to fill out. Our project, SHAPES, had over 70 fields in its first iteration which described everything very thoroughly. But when it was time to fill in the blanks for the records, it was a massive task. The entire schema was redone many times to get to a smaller 18 fields, which allows for a more concise description of the 3D model data. Find the fields that are the most important and focus on those, otherwise it'll take someone an hour to do one record, and that kind of time in the metadata processing denotes an inefficient schema.

Acknowledgments

The CS3DP Metadata working group would like to thank all survey participants and acknowledge the contributions of Andrea Thomer in helping to conduct Survey 1 and gather responses from participants.

Notes

1. London Charter, "Glossary," accessed August 28, 2020, http://www.londoncharter.org/glossary.html.
2. Filecamp, "Digital Asset Management: What Is It and Why Do You Need It?" accessed August 28, 2020, https://filecamp.com/blog/what-is-digital-asset-management/; Digital Curation Centre, "The DCC Curation Lifecycle Model," accessed October 2, 2020, https://www.dcc.ac.uk/sites/default/files/documents/publications/DCCLifecycle.pdf.
3. Survey 1 results; see https://dx.doi.org/10.34944/dspace/7467
4. London Charter, "Glossary."
5. N. A. Matthews, *Aerial and Close-Range Photogrammetric Technology*, Technical Note 428 (Denver, CO: US Department of the Interior, Bureau of Land Management, National Operations Center, 2008); Cultural Heritage Imaging, "Photogrammetry: Practical, Scientific Use of Photogrammetry," Vimeo, accessed August 28, 2020, https://vimeo.com/channels/practicalphotogrammetry.
6. Cultural Heritage Imaging, "Digital Lab Notebook," accessed August 28, 2020, http://culturalheritageimaging.org/Technologies/Digital_Lab_Notebook/.
7. Texas Tech University Libraries, SHAPES home page, accessed August 28, 2020, https://shapes.lib.ttu.edu/.
8. Library of Congress, "Standards," accessed August 28, 2020, https://www.loc.gov/librarians/standards; Library of Congress, "Recommended Formats Statement," accessed August 28, 2020, https://www.loc.gov/preservation/resources/rfs/TOC.html.
9. Samvera home page, accessed August 28, 2020, https://samvera.org; Samvera, "Technical Metadata Application Profile," accessed August 28, 2020, https://wiki.duraspace.org/display/samvera/Technical+Metadata+Application+Profile (requires login).
10. Smithsonian Digitization Program Office, "Smithsonian 3D Metadata Model," *SI Digi Blog*, November 1, 2018, https://dpo.si.edu/blog/smithsonian-3d-metadata-model; Martina Trognitz, Kieron Niven, and Valentijn Gilissen, "Section 3. Archiving 3D Data," in 3D Models in Archaeology: A Guide to Good Practice, Guides to Good Practice (Archaeology Data Service and Digital Antiquity, 2009), https://guides.archaeologydataservice.ac.uk/g2gp/3d_3-3.
11. Sketchfab, home page, accessed October 7, 2020, https://sketchfab.com/.
12. NIH 3D Print Exchange home page, accessed October 7, 2020, https://3dprint.nih.gov/.
13. MorphoSource, "Information for Contributors," accessed January 7, 2022, https://www.morphosource.org/docs/contributors.
14. Force 11, "The FAIR Data Principles," https://www.force11.org/group/fairgroup/fairprinciples.
15. Don R. Swanson, "Searching Natural Language Text by Computer," *Science*, new series 132, no. 3434 (1960): 1099–104, https://www.jstor.org/stable/1706747.
16. Smithsonian 3D Digitization, "Articulated Wooly Mammoth," accessed February 7, 2019, https://3d.si.edu/explorer/articulated-woolly-mammoth (page discontinued); Teraoka Gensyou, "Articulated Woolly Mammoth Manga," *SI Digi Blog*, May 16, 2018, https://dpo.si.edu/blog/articulated-woolly-mammoth-manga.
17. PREMIS Editorial Committee, *PREMIS Data Dictionary for Preservation Metadata*, version 3.0, (Washington, DC: Library of Congress, June 2015, last modified November 2015), https://www.loc.gov/standards/premis/v3/.

18. Smithsonian Digitization Program Office, "Smithsonian 3D Metadata Model."
19. Force 11, "FAIR Data Principles."

Bibliography

Cultural Heritage Imaging. "Digital Lab Notebook." Accessed August 28, 2020. http://culturalheritageimaging.org/Technologies/Digital_Lab_Notebook/.

———. "Photogrammetry: Practical, Scientific Use of Photogrammetry." Vimeo. Accessed August 28, 2020. https://vimeo.com/channels/practicalphotogrammetry.

Digital Curation Centre. "The DCC Curation Lifecycle Model." Accessed October 2, 2020. https://www.dcc.ac.uk/sites/default/files/documents/publications/DCCLifecycle.pdf.

Digital Public Library of America home page. Accessed October 7, 2020. https://dp.la/.

Dublin Core Metadata Initiative. "Dublin Core." Accessed October 7, 2020. https://www.dublincore.org/specifications/dublin-core/.

Europeana home page. Accessed October 7, 2020. https://www.europeana.eu/en.

FFmpeg. "ffprobe Documentation." Accessed October 7, 2020. https://ffmpeg.org/ffprobe.html.

Filecamp. "Digital Asset Management: What Is It and Why Do You Need It?" Accessed August 28, 2020. https://filecamp.com/blog/what-is-digital-asset-management/.

Force 11. "The FAIR Data Principles." https://www.force11.org/group/fairgroup/fairprinciples.

Gensyou, Teraoka. "Articulated Woolly Mammoth Manga." *SI Digi Blog*, May 16, 2018. https://dpo.si.edu/blog/articulated-woolly-mammoth-manga.

Library of Congress. "MARC Standards." Accessed October 7, 2020. https://www.loc.gov/marc/.

———. "MODS: Uses and Features." Accessed October 7, 2020. https://www.loc.gov/standards/mods/mods-overview.html.

——— "Recommended Formats Statement." Accessed August 28, 2020. https://www.loc.gov/preservation/resources/rfs/TOC.html.

———. "Standards." Accessed August 28, 2020. https://www.loc.gov/librarians/standards.

London Charter. "Glossary." Accessed August 28, 2020. http://www.londoncharter.org/glossary.html.

Matthews, N. A. *Aerial and Close-Range Photogrammetric Technology: Providing Resource Documentation, Interpretation, and Preservation*. Technical Note 428 (Denver, CO: US Department of the Interior, Bureau of Land Management, National Operations Center, 2008).

MorphoSource. "Information for Contributors." Accessed January 7, 2022. https://www.morphosource.org/docs/contributors.

National Archives. "Download DROID: File Format Identification Tool." Accessed October 7, 2020. https://www.nationalarchives.gov.uk/information-management/manage-information/preserving-digital-records/droid/.

NIH 3D Print Exchange home page. Accessed October 7, 2020. https://3dprint.nih.gov/.

PREMIS Editorial Committee. *PREMIS Data Dictionary for Preservation Metadata*, version 3.0. Washington, DC: Library of Congress, June 2015, last modified November 2015. https://www.loc.gov/standards/premis/v3/.

RDA: Resource Description and Access Toolkit. "About." Accessed October 7, 2020. https://www.rdatoolkit.org/about.

Samvera home page. Accessed August 28, 2020. https://samvera.org.

———. "Technical Metadata Application Profile." Accessed August 28, 2020. https://wiki.duraspace.org/display/samvera/Technical+Metadata+Application+Profile (requires login).

Sketchfab home page. Accessed October 7, 2020. https://sketchfab.com/.

Smithsonian 3D Digitization. "Articulated Wooly Mammoth." Accessed February 7, 2019. https://3d.si.edu/explorer/articulated-woolly-mammoth (page discontinued).

Smithsonian Digitization Program Office. "Smithsonian 3D Metadata Model." *SI Digi Blog*, November 1, 2018. https://dpo.si.edu/blog/smithsonian-3d-metadata-model.

Swanson, Don R. "Searching Natural Language Text by Computer." *Science*, new series 132, no. 3434 (1960): 1099–104. https://www.jstor.org/stable/1706747.

Texas Tech University Libraries. SHAPES home page. Accessed August 28, 2020. https://shapes.lib.ttu.edu/.

Trognitz, Martina, Kieron Niven, and Valentijn Gilissen. "Section 3. Archiving 3D Data." In *3D Models in Archaeology: A Guide to Good Practice*. Guides to Good Practice. Archaeology Data Service and Digital Antiquity, 2009. https://guides.archaeologydataservice.ac.uk/g2gp/3d_3-3.

US Access Board. "Principles of Universal Design." Accessed October 7, 2020. https://www.access-board.gov/guidelines-and-standards/communications-and-it/26-255-guidelines/825-principles-of-universal-design (page discontinued).

Chapter 5

Copyright and Legal Issues Surrounding 3D Data

Andrea D'Andrea, Michael Conyers, Kyle K. Courtney, Emily Finch, Melissa Levine, Nicole Meyer, Adam Rountrey, Hannah Scates Kettler, Kate Webbink, and Ann Whiteside

ABSTRACT

An overview of essential legal concepts and strategies, this chapter synthesizes the ideas, questions, and legal issues that arise in relation to 3D data. Case studies provide scenarios based on real-world situations that will help readers recognize legal and policy issues. Readers will have a framework for thoughtful decision-making that is consistent with their particular mission.

We begin with a general overview of US copyright law and then focus on case law that is relevant for understanding the legal status of 3D models. Case studies focus on creation or acquisition methodologies including institutional ***photogrammetry*** *of an object, Indigenous community and nonprofit organization partnership to digitally document and preserve cultural artifacts,*

transferring ownership of 3D data to an institutional repository, and a complex researcher-developed 3D model. These case studies are used to provide relevant illustrations of practices and situations that may prompt legal questions, but we also recommend considering more complex ethical issues early on. These case studies will help readers recognize legal and policy issues that may be relevant to their current practices in 3D creation and dissemination, and review will emphasize expectations under both open and restricted ***access*** *scenarios, including contracts and licensing. In certain case studies, expansions are included to highlight additional domain-specific questions.*

Introduction

The Copyright/Ownership working group examined legal questions related to generating and preserving 3D data. We surveyed basic copyright as currently defined by US law, then mapped out elements in a generalized 3D-data **workflow** as a way to explore the legal and ethical issues. We developed four scenarios that might be associated with data from 3D representations of cultural objects to help readers identify legal issues in their own work. Each scenario includes corollary questions for issue-spotting for (1) **preservation**, (2) sharing of or access to data, and (3) use and reuse of data.

We extensively discussed questions of ethics and practice within particular disciplines (e.g., paleontology, anthropology, archaeology, etc.). Questions of ethics and practice are critical. They influence—but are distinct from—formal legal frameworks. Thus, we tried to focus the scenarios on legal questions because they are more generalizable than the ethics and disciplinary constructs. To emphasize, the ethical questions were not less important than the legal and business questions. However, ethical questions are subject to greater variety and range depending on the purpose of the data and the discipline in which they will be used.

At this stage, our recommendations for professional practice are notional, and we urge practitioners to continue to think about frameworks for their needs and interests. Our hope is that the questions investigated here support that ongoing work.

We start with an overview of essential legal concepts.

Foundations: Copyright and the "Bundle of Rights"

Under current United States copyright law (Title 17 of the United States Code or "U.S.C."), authors or creators of original works hold copyrights in their works automatically. The law, under 17 U.S.C. § 106, grants authors a set of exclusive rights often referred to as a "bundle of rights." These rights include the right to reproduce, distribute, make derivatives, publicly perform, and publicly display their works. Categories of copyrightable works include literary works; musical works; dramatic works; pantomimes and choreographic works; pictorial, graphic, and sculptural works; motion pictures and other audiovisual works; sound recordings; and architectural works

Fixed and Creative

In order to be eligible for these exclusive rights, the work must meet a threshold test: copyright protects only "original works of authorship fixed in any tangible medium of expression."[1] This means that the work must have been created *independently and contain a sufficient amount of creativity*. It is relatively easy to meet these requirements for some creative works like fictional novels and new music. However, some kinds of works are not eligible for copyright and therefore are excluded from protection. For example, copyright does not protect things such as ideas, methods, titles, recipes, data, or facts.

This is where 3D data and copyright intersect. As a general rule, data are not eligible for copyright protection as a type of fact, and facts cannot be copyrighted. However, some expression of data may be sufficiently original to qualify for copyright or, further, may be arranged or compiled in a way that is eligible for copyright. Copyright can protect creative *expressions* of fact or data.

For example, the statement "The Sun is, on average, 93,000,000 miles away from Earth" is a fact. Therefore, it is not copyrightable. However, what if we took this sentence, and its data, and expressed them in the drawing in figure 5.1?

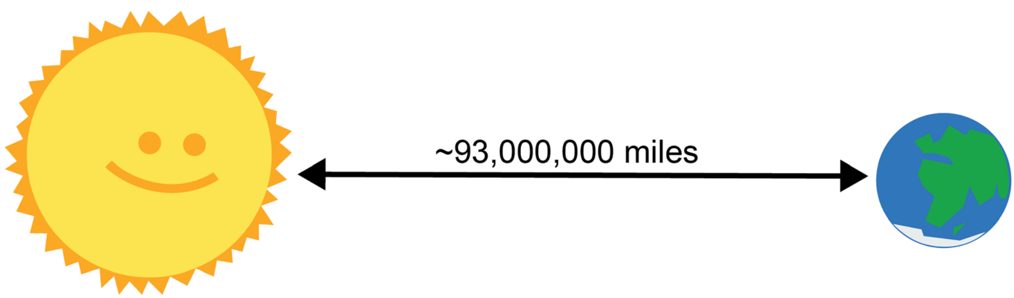


Figure 5.1
A potentially copyrightable expression of facts

Now this *drawing* could, potentially, be copyrightable. It is a creative expression of the facts (the Sun and Earth) and the data (93 million miles).

Thinking Globally

This chapter focuses on US law, which treats databases differently from the EU approach. In the EU, Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases governs the organization of facts in databases. This directive was reconsidered by the EU in 2018. In our discussions about data from scans of 3D objects, we observed that many projects are implicitly or explicitly global in nature. Thus it is important to develop a general awareness beyond US law for many projects. One of the key recommendations that emerged from our discussions: it is vital to design projects with the end result in mind. If the goal is to have an open-access result, that needs to be designed into the structure of the project. If a more closed approach is desired, that too needs to be designed into the project. (Think of situations where there are privacy or ethical concerns.) We recommend being clear and explicit about your assumptions and goals from the outset. It is a worthwhile investment of time and expertise.

Foundations: Case Law

In thinking about how data from scans of 3D objects might be treated, it is helpful to consider case law. Case law consists of decisions written by judges to address disputes brought to court. Copyright cases help provide factual examples of how judges interpret the copyright laws passed by the US Congress. In doing so, they give us guidance on how a given situation is akin to a given set of facts and thus could be treated analogously.

In the US court system, and in common law in general (that is, law based on precedent and opinions written by judges rather than statutory law), courts typically look to earlier decisions to determine the outcome of the law's interaction with new technology. This was true of famous cases involving player pianos, photography, VCRs, MP3s, and more recently BitTorrent file sharing. Where a statute like the US Copyright Act is involved, court decisions are used to interpret these statutes, the broad categories of content they may protect, and the facts of the case.

The application of the law to 3D data is no different—we look at precedential cases and make an analysis. Sometimes even the simplest things can be the focus of important copyright case law. Take, for example, telephone books.

What Is "Original" for the Purpose of Copyright? Feist Publications v. Rural Telephone Service Co., *499 U.S. 340 (1991)*

The first modern case that draws upon the current copyright law is *Feist Publications v. Rural Telephone Service Co.* In *Feist*, a company called Rural Telephone Service published a phone book with an alphabetical list of the names of its subscribers in the white pages. Feist Publications asked for a license to use and publish the white page listings that Rural had collected. Rural refused the license, so Feist extracted the listings it needed from Rural's directory without permission. Feist then published a similar phone book using the same list of subscribers from Rural's white pages. Rural sued Feist for copyright infringement of its directory.[2]

The US Supreme Court examined a former principle known as "sweat of the brow" doctrine, which stated that acquiring copyright was a function of hard work or effort. The Court even heard evidence that Feist and Rural were discussing, during their failed negotiations, ideas surrounding the time and effort it might take to make most accurate white pages possible. The Court discussed the "sweat of the brow" doctrine, focusing on the time, money, and effort to hire staff to go door-to-door to confirm that the white page information was correct. Ultimately, though, the Court rejected these principles as a means of acquiring copyright.[3]

This is a seminal case because it illustrates the modern legal interpretation of copyrightability, moving from the effort-based "sweat of the brow" test to the more statute-centered "originality" test. As another court summarized in 1985, "[i]n 14 hours Mozart could write a piano concerto, J. S. Bach a cantata, or Dickens a week's installment of *Bleak House*. The Laffer Curve, an economic graph prominent in political debates, appeared on the back of a napkin after dinner, the work of a minute. All of these are copyrightable."[4] Effort and time do not always equate to copyright's originality test (see figure 5.2).

Feist, however, ultimately centered on two well-established principles in United States copyright law. First, *facts* are not copyrightable. Second, while facts are not copyrightable, *compilations of facts can be copyrightable if they possess the requisite level of originality/creativity.*[5]

Figure 5.2
A picture of white pages showing no originality and yellow pages showing some possible originality. Yellow pages image by "How Can I Recycle This?" (https://www.recyclethis.co.uk) is licensed under CC BY 2.0 (https://creativecommons.org/licenses/by/2.0/).

The Supreme Court determined that white page listings lacked the minimal degree of creativity necessary for copyright protections. The Court observed that "[a]s a constitutional matter, copyright protects only those constituent elements of a work that possess more than a de minimis quantum of creativity."[6] Further, it found that there can be no copyright in a work in which "the creative spark is utterly lacking or so trivial as to be virtually nonexistent."[7] Arranging a list of data alphabetically is not enough to establish copyright protection in a work. The court concluded by stating that "this decision should not be construed as demeaning Rural's efforts in compiling its directory, *but rather as making clear that copyright rewards originality, not effort*" (authors' emphasis).[8]

Building upon *Feist*, US courts have continually found that facts and data are rarely copyrightable. The *Feist* test can be broken down into two parts. First, the work must have been independently created by the author, not copied from another work. Second, the work must possess sufficient creativity. Only a "modicum" of creativity is necessary. The Supreme Court has ruled that some works fail to meet even this low threshold. As simple lists arranged in alphabetical order, the white pages telephone books in *Feist* are a common example of creations that are insufficiently original to garner copyright protection.

What Are Regulations?

In addition to cases, we have a variety of other laws called ***regulations*** that are put forth by the US Copyright Office. These appear in a publication called the ***Code of Federal Regulations*** (CFR). The regulations implement a more detailed requirement of originality, combining interpretation of the Copyright Act and the ***Feist*** decision. For example, 37 C.F.R. § 202.1(a) prohibits registration of "[w]ords and short phrases such as names, titles, slogans; familiar symbols or designs; [and] mere variations of typographic ornamentation, lettering, or coloring." Later, § 202.10(a) states "to be acceptable as a pictorial, graphic, or sculptural work, the work must embody some creative authorship in its delineation or form." Regulations often provide very practical guidance about the application of the law.

Some combinations of common or standard design elements may contain sufficient creativity with respect to how they are juxtaposed or arranged to support a copyright. Nevertheless, not every combination or arrangement will be sufficient to meet this test. A determination of copyrightability in the combination of standard design elements depends on whether the selection, coordination, or arrangement is done in such a way as to result in copyrightable authorship.

Another Example: Bridgeman Art Library v. Corel, *36 F. Supp. 2d 191 (S.D.N.Y. 1999)*

In 1999, eight years after the Supreme Court ruled on *Feist Publications v. Rural Telephone Service Co., (1991)*, the United States District Court for the Southern District of New York ruled on *Bridgeman Art Library v. Corel (1999)*. While the decision's precedent is persuasive rather than binding, the case is frequently cited in debates on originality requirements for copyright protection, especially with photographs depicting ancient *two-dimensional* works.[9]

The Bridgeman Art Library held a large collection of photographs and digital images of famous public domain artwork. It sued Corel Corporation for copyright infringement, asserting that Corel had illegally obtained and distributed copies of these images. While the copyright on the original art had long since expired (see figure 5.3), the Bridgeman Art Library claimed to have copyright ownership of these new reproductions.[10]

Figure 5.3
Laughing Cavalier, a seventeenth-century painting by Frans Hals, is currently in the public domain.

In the first judgment, the court ascertained whether or not the photographs were copyrightable considering the public domain status of the originals. The judge found that there could be no infringement in the photographs because they were images of art that was already in the public domain (they were no longer eligible for copyright). A subsequent judgment dismissed the case. Judge Lewis Kaplan said that *exact photographic reproductions of two-dimensional objects (public domain art, for example) were "slavish reproductions," and as such could not receive copyright protection due to their lack of originality*. The *Bridgeman* case added a new term to the art and museum worlds: "slavish reproduction." This doctrine helped settle the law that two-dimensional reproductions of two-dimensional material in the public domain could not qualify for copyright protection.

With the concepts from the *Feist* and *Bridgeman* cases in mind, we turn to the question of whether and how scans of three-dimensional objects or physical sites are sufficiently original to warrant copyright protection where they are produced with tools like computed tomography (CT) scan or photogrammetry. What aspect of the scanning data might be eligible for protection automatically? What can you do to enhance or reduce the likelihood of copyright depending on your project design goals?

3D Scans in the Courts: Meshwerks, Inc. v. Toyota Motor Sales U.S.A., Inc., *528 F.3d 1258 (10th Cir. 2008)*

One of the few cases we have involving 3D data emerges in a 2008 case from the Tenth Circuit. The *Meshwerks* case relies on the standards laid out in the *Feist* decision. Toyota hired an advertising firm, which in turn hired Meshwerks into create digital 3D models of Toyota vehicles to be used in advertising materials. Meshwerks personnel took numerous measurements of each vehicle and then used software to generate a digital image resembling a wireframe model of the vehicle based on the measurements taken (think of measurements as facts). From there, Meshwerks personnel digitally tweaked each model in order for it to more closely resemble the Toyota vehicle it was trying to recreate (think of this kind of skill and effort to reproduce something as closely as possible as "slavish" reproduction). Meshwerks personnel spent approximately 80 to 100 hours per vehicle, working on details with each digital model. The results were two-dimensional wireframe depictions of Toyota vehicles that appeared three-dimensional on screen (see figure 5.4). They were exact replicas of the cars; no new features or additions were made to the digital models.

Toyota used these models in various print, online, and television advertisements. However, according to Meshwerks, the agreement was for only a single use of the models. And since Meshwerks had registered the work with the US Copyright Office, a prerequisite to bringing a lawsuit, it claimed that the additional advertising uses were copyright infringement. Meshwerks sued Toyota, alleging that Toyota infringed on the copyrights Meshwerks held in the digital models it created.

Copyright Registration

The Copyright Office uses the *Compendium of U.S. Copyright Office Practices*, 3rd ed., its governing administrative manual, to make copyright registration determinations. Ultimately, registration decisions are assigned to copyright examiners who use the Compendium to review the application. The copyright examiners make copyright registration determinations that are unique to each examiner. As a result, there are registration decisions that result in different determinations depending on the examiner.

While Meshwerks argued specific points about the time, effort, and skills necessary to produce the models, the court's focus was on the test for creativity and originality. The court focused on whether Meshwerks' models qualify as independent creations, as opposed to exact copies of Toyota's vehicles or designs. The court relied on the decision made in *Feist*, which stated a work must "possess at least some minimal degree of creativity" to qualify for copyright protection.

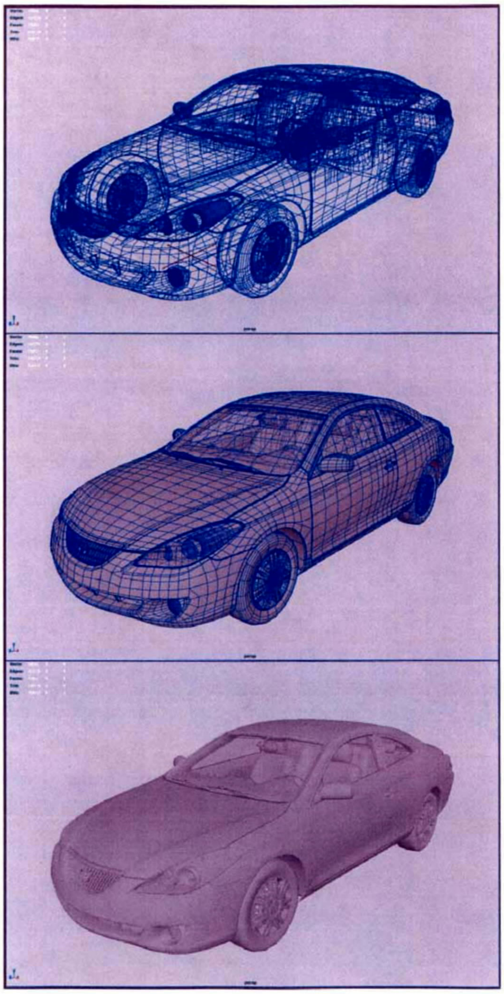


Figure 5.4
Meshwerks' 3D wireframe of a Toyota from APPENDIX A, *Meshwerks, Inc. v. Toyota Motor Sales U.S.A., Inc.*, 528 F.3d 1258, 1271 (10th Cir. 2008).

The court concluded by saying that Meshwerks did not have a valid copyright in the digitized models it created. The models were created using designs that the automobile manufacturer (Toyota) produced, so they were not original works of art that could be copyrighted, and because there was no valid copyright, there could be no infringement. The wireframes were nothing but very good copies of the original cars, adding no new original expression. The court stated "[o]riginality is the [necessary component] of copyright. The designs of the vehicles, however, [are not original]. [They] owe their origins to Toyota, not to Meshwerks, and so we are unable to reward Meshwerks' digital wire-frame models, no doubt the product of significant labor, skill, and judgment, with copyright protection."

Photographs and Photogrammetry: How Copyright Cases Intersect with 3D Scans or Models

Photogrammetry is the science of making measurements from photographs, and the basic idea is almost as old as photography. Photogrammetry was traditionally used for land surveying, but with developments in structure from motion algorithms and GPU (Graphics Processing Unit) computing, it has become a common methodology for obtaining 3D digital models of physical objects. The input of photogrammetry is the photographs, and the output is typically a 3D model or measurement set. The creation of a 3D model using photogrammetry (see figure 5.5) relies on algorithms that identify matching points in photos and calculate lens distortion and camera position in order to create an output (3D model, map, etc.).

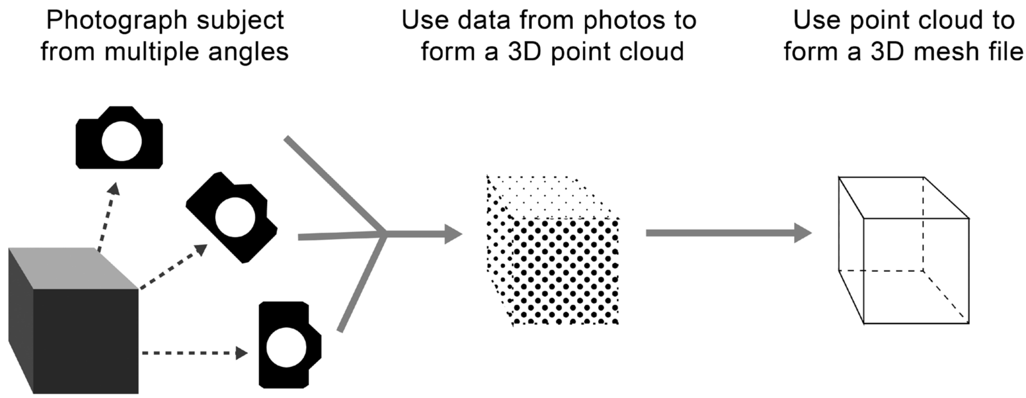


Figure 5.5
Generation of a 3D file using photogrammetry.

It is well established that photographs are typically protected by copyright law, provided they are sufficiently original. Photographs were first added as a category of protected works to the United States Copyright Act in 1865 and continue to receive the same protection today. Photographs, like all other copyrightable works, must be "original works of authorship fixed in any tangible medium of expression" to qualify for copyright protection. In other words, photographs must possess a minimal amount of creativity to be copyrightable (in contrast to the facts in the *Bridgeman* case, where the photographs of art in the public domain were intentionally as unoriginal as possible).

Early in the history of photography as a new technology, some argued that taking a photograph was merely a mechanical process accomplished by the camera rather than a creative expression. But in 1884, the US Supreme Court determined in *Burrow-Giles Lithographic Co. v. Sarony* that photography is a creative expression (art) rather than just a mechanical reproduction of an object. The Court came to this conclusion *because of the creative choices* a photographer must make. The photographer must *choose the lighting, camera angle, how the subject will be posed, distance from the subject, as well as other artistic features.*

To further explain the differences between photographer intentions: If an art photographer shoots art in museum galleries of 2D works that are in the public domain and intends to make what the law calls a "slavish reproduction" of the work, regardless of the lighting, distance, or other factors, the photographer does not have a copyright over their photo of that public domain work. However, if the art photographer shoots art in museum galleries of 2D works *still under copyright*, or 3D works, choices about camera angle, lighting, distance, and other factors may be enough creative expression for copyright of the photos. The photogrammetrist's goal is to produce 3D data from the photos for the object to be replicated. However, the photogrammetrist's photos *might* have the creativity necessary to be copyrightable.

The concept of "slavish reproduction" in copyright applies only to reproductions of 2D works in the public domain. The concept is that the photographer can't acquire a new copyright in their photo because it lacks the creativity necessary because it is a near-perfect reproduction of that 2D work that is in the public domain. However, if you are taking a photo of a 3D work, even a 3D work that is in the public domain, and even if you take hundreds of these photos, there is a presumption in the law that the lighting, angle, and other aspects of the photo have the minimum spark of creativity necessary for copyright to exist in the photos. The law does not look to the "intent" of the photos. The fact that photogrammetry rigs take hundreds of pictures from varying angles certainly makes a case for slavish reproduction, but there is no case or statute that states these photos are clearly not copyrightable.

With all this in mind, photogrammetry photographs could qualify for copyright protection because one must make these kinds of decisions. A photogrammetrist has to make choices about many elements of the photograph such as angle of camera, lighting, photo overlap, and distance from which to take the photos. The ultimate goal of photogrammetry is to extract the data or measurements from the photos in order to create something else (3D model, map, etc.). Although a work as a whole can be protected by copyright, there are occasions when individual parts of that work may not be protected by copyright. For example, an author of a textbook would most certainly receive copyright protection for the book. However, if the author used a public domain photograph in that book, the photo would not be protected by copyright and another person could come along and extract and use that particular photo. Another example is recipes used in a cookbook. A cookbook as a whole can be protected by copyright, but the author cannot receive copyright protection in the individual recipes. If the recipes are simply lists of ingredients with directions for combining them, they are outside the scope of copyright protection.

By legal analogy, the same scenario is true for photogrammetry. The photographs taken by the individual would be copyrightable, but the data or measurements extracted from those photos would not qualify for copyright protection. For example, the data extracted from the photos in the photogrammetry process might consist of factual measurements of distances between points. As discussed, data and facts are not afforded copyright protection. Even manipulation of the extracted data in pursuit of **accuracy** (such as choice of the correct photos from which to extract data or manual corrections to make the data more precise) would fall under the same "facts" category. These results fall outside of the categories of types of work that are generally protected by copyright.

Foundations: Licensing and Contracts

In addition to having an essential understanding of copyright concepts, it is critical to have a basic understanding about how contracts work as a legal and practical matter. In addition to being legally binding documents, contracts are excellent tools for communication so that the parties to a relationship understand what is mutually expected. Whether they are called contracts, agreements, terms of use, memoranda of understanding, or something else, these are important tools for designing your project. Also keep in mind that agreements with funders are contractual commitments. We interact with contracts on a daily basis, so they seem familiar. It is important to review some basic information in this chapter since understanding contracts and their roles in 3D data preservation are critical.

What Is a Contract? What Is a License?

It is important that you understand the basics of offer, acceptance, and the value of expressing things in writing. At its simplest, a contract is an agreement between two or more people or entities (like a corporation, a university, a state) to do or not do something in return for some kind of valuable benefit. That benefit is called "consideration." For a contract to exist, there are seven key requirements. To form a contract, you need

1. an offer
2. acceptance of that offer
3. a promise to perform
4. valuable consideration (for example, a promise or a payment)
5. a time or event at which the performance must be accomplished
6. any terms and conditions for the performance

The seventh requirement is actual performance if the contract is "unilateral." That is, there is a promise to pay (or give the valuable consideration) in return for actual performance. ("You agree to take photos of my rock for $100 by Friday." The performance is the taking of the photos.) By contrast, a "bilateral" contract is one in which a promise is exchanged for a promise. ("You promise to take photos of my rock by Friday, and I promise to pay you $100 on Friday.")

As a legal matter, many kinds of contracts can be written or oral. *We strongly recommend the practice of expressing your contracts in writing* for a few reasons—both legal and practical. Written contracts are easier to prove and are more likely to articulate the intention of the parties than a contract entered into by verbal exchanges or implied by circumstances. Contracts for illegal purposes are not enforceable. Secondarily, courts often look to "the four corners of the contract" to find the intention of the parties, and that which was not written within the "four corners" may not be held as a valid part of that contract. One other advantage of a written contract is that the drafting process tends to flesh out places of disagreement, unexamined assumptions, or misunderstanding between the parties that can typically be worked out in negotiation discussions.

A license is a kind of contract that typically gives someone permission to do or use something. Think of it as a kind of authorization to do or use something by someone who owns something or has the authority to give the authorization. (This is distinct from, say, a license to practice law or medicine or a driver's license.) In the area of intellectual property like copyright, a license can be used to permit copying and thus excuse the licensee from being accused of copyright infringement. Intellectual property licenses typically describe the subject of the license along with a term (duration), sometimes the territory in which the license applies (a state, region, or country, for example Mexico/the United States/Canada but not Japan), renewal provisions (say, a three-year term with automatic renewal until one party gives the other notice of termination), and other conditions as appropriate. You see these every day in the click-through licenses you routinely agree to.

One of the key benefits to written contracts is that they allow for clarity. If you have ever been involved in a contract negotiation, you have had the experience of exchanging draft documents. Each side might strike out particular language that it cannot or will not agree to. They might provide new suggested language. Reading the document and passing it back and forth can improve the likelihood that the contract will be a guide for the relationship, in addition to being legally binding. There are also situations where you do not have the option to negotiate (or do not want to offer the option), typically for administrative reasons. It is impractical to renegotiate every contract for social media products, e-book or music licenses, terms of use on any website. Neither approach is particularly good nor bad, better nor worse. Each tool has a different purpose. Our hope is that you will start to think about contracts as tools that you can use to express your intentions and design your projects—or understand how to participate in someone else's project. See Appendix B for examples.

You will see in the case studies that there are no firm rules or explicit practices that can be dictated. There is unavoidable complexity that can be addressed through careful thought and articulation of expectations. With that in mind, we turn to the case studies.

Case Studies

We developed several case studies for current scenarios to identify legal issues that might be associated with data from 3D scans of cultural objects—and the corollary questions that might be associated with (1) preservation of data, (2) sharing of or access to those data, and (3) use and reuse of those data. We also discussed questions of ethics and practices within particular disciplines (for example, paleontology, anthropology, archaeology, or other fields). The latter was complex and difficult to define in the hypothetical. Questions of ethics and practice are critical—but distinct from formal legal frameworks. Thus, we tried to focus on the legal questions as somewhat less theoretical than the ethics and disciplinary constructs. To emphasize, the latter were not less important to the authors than the legal and business questions. Indeed, there was more discussion of these concerns than of the formal legal questions. In part, this is because they are indeed important to practitioners—and they are still somewhat subjective and more familiar than some of the formal legal concepts.

The structure of each of the case studies varies somewhat to address the particulars of each situation. Despite the variations, each one considers common issues. Critically, you need to know what you want to accomplish to structure a new project in a way that the action of law supports—or at least does not hinder—your desired outcomes. Data composed of facts are not eligible for copyright protection. However, collections of data as compilations may be eligible for purposes of copyright protection. These are commonly controlled through contracts (licenses) that confine use and reuse of the collection or its component parts. There is a general preference reflected in our conversations for openness, for minimizing controls and simplifying use and reuse. In our discussions, some participants wanted to generate totally open data, available to anyone for any reason. There were some situations where some level of "closure" was legitimately desirable:

- A scan is a reproduction of an object or site that has some cultural sensitivity (ethics and sometimes law).
- A scan is a reproduction of an object that might have some explicit quality that requires security (making it inappropriate for open access) for ethical reasons or as a matter of law, such as medical privacy (for example, consider the relevance of laws like HIPAA—the Health Insurance Portability and Accountability Act of 1996).
- The creator of a scan, or a repository, has a business model that requires users to pay a fee for access to data.

You could have scenarios where one or more of the above are relevant. We discussed the legal implications primarily under US law—but we also spent time on EU perspectives on copyright, privacy, and database directives. Because libraries, **archives**, and museums now work locally and globally, practitioners should become familiar with work by the Research Data Alliance (https://www.rd-alliance.org/), the Digital Curation Centre (https://www.dcc.ac.uk/), and others to think about how to preserve data in an optimal manner from a global perspective. There are some essential questions that arise in the design of a preservation approach for the scenarios considered:

- *What result do you want?* Freely open? Controlled/permissioned access and use?
- *What result is dictated by the inherent nature of the object?* Is the object to be scanned imbued with any copyright intrinsically? Does a scan create a "new" set of facts (measurements of the object) that are ineligible for copyright—or a derivative work tied to the rightsholder of the original work (if applicable)?
- *What is the practice in a given discipline engaged in the work?* The field of practice is critical to ethics and behavior vis-à-vis preservation, access, use, and reuse of scan data and in some cases any resulting model (whether software or 3D print). An art historian might make different choices than an anthropologist …a geologist would make different choices than an architectural scholar.
- *How should you respond in this evolving arena?* Our discussions raised some important questions for which there is no immediate answer. For example, are import/export laws applicable to datasets of 3D scans in the same way as the

physical objects they represent? (If you are not permitted to take a specimen over a border, is the reproduction treated the same way—or no?) Are *reproductions* of culturally sensitive items to be treated in the same way as the original? Is the reproduction vested with some meaning, like the underlying object? Or is the reproduction or dataset free for use and cross-border movement? Is this something that requires case-by-case consideration?

The case studies provide some perspectives on real situations and ways that this working group imagined handling these scenarios as examples for discussion.

Case Study 1: Natural History Specimen, Institutional Project

Background

This case study is intended to represent a situation in which a natural history object, such as a rock, fossil, bone, or other nonhuman biological specimen is subject to photogrammetry with the goal of producing a three-dimensional surface **mesh** with color information. Instances like this would be commonly found in museum and university settings. The context is general, but the location at which the specimen was collected and the university itself are assumed to be in the United States.

Object

The object being digitized is a fossil mastodon skull that was accessioned by a university museum. At the time of acquisition, the donor signed a deed of gift asserting that they were the true and legal owner of the specimen and that they had the right to convey it. The donor irrevocably and unconditionally transferred ownership and all rights, title, and interest in the specimen to the university through the deed of gift.

Process

The 3D mesh is produced using photographs and photogrammetry software. A series of approximately 200 photos, taken in several rings around the object, are imported into a commercial photogrammetry software package. The photographer is an employee of the university and is doing the work as part of normal responsibilities. Lighting is designed to be diffuse and even, limiting sharp shadows on the specimen. Camera settings are chosen to optimize for depth of field and low noise. The goal is to produce photos that can be used to create a detailed photorealistic digital replica of the specimen.

Output

The photogrammetry software produces a scaled 3D mesh (e.g., PLY file) with color information (either as vertex colors or UV-mapping of texture images). There may be multiple versions of the mesh at different sizes and **resolutions**. The photographs used in the photogrammetry process will be archived along with the 3D mesh.

Usage

The 3D model will be distributed through an online institutional repository. Users may view and download the mesh file. The university's intent is that the model be used for research and educational purposes and that appropriate credit will be given when it is used. Commercial usage will require payment.

Issues

In this case study, a natural history object is the subject of digital reproduction by photogrammetry. As an unadorned work of nature, the mastodon skull itself is not subject to copyright protection. The *Compendium of U.S. Copyright Office Practices* explicitly excludes works "produced by nature, animals, or plants" from registration.[11] However, the lack of protection under copyright law does not indicate the object is free from all restrictions. In some cases, prior contractual agreements may limit how the object can be used. For example, a cast of a specimen acquired from another museum may be encumbered by a contractual agreement stating that no additional duplicates are to be produced, or if the skull was collected on public land, the National Park Service or other government agency may place limitations on the production of copies. Here, the university museum acquired the skull from a donor who agreed to transfer the skull and all associated rights, title, and interest to the museum, and there are no contracts that would limit the use of the skull.

As part of the process of creating the 3D mesh, a series of several hundred photographs will be produced by a museum employee as part of their normal duties. Nearly all photographs of 3D objects are subject to copyright protection because of the creative choices made by the photographer in the process (see discussion earlier in the chapter regarding *Burrow-Giles Lithographic Co. v. Sarony*), and the photographs of this mastodon skull, produced by the museum, are probably protected by copyright. Here, as the employee produced the photographs as part of their normal duties, they would be considered "work for hire,"[12] and the copyright in the photos would be held by the museum as the employer. (The work-for-hire doctrine is an exception to the general rule that authors own copyright in their own creations.)

The photogrammetry software extracts data, such as feature positions and depth maps, from the set of photographs to produce a 3D mesh with color information. This mesh can be considered as a "copy" of the physical object; the intent is to create an accurate, three-dimensional, digital representation of the physical object. Copies or slavish reproductions are not original works of authorship (see earlier discussion of *Meshwerks, Inc. v. Toyota Motor Sales U.S.A.*), and *new* rights are not created in copies. Because the skull itself is not subject to copyright protection and no new rights are created in the copy, the 3D mesh produced in this project is not subject to copyright.

The university museum intends to share the 3D mesh publicly for research and educational purposes, but it wants to be credited appropriately when the mesh is used, and it wants to require payment for commercial usage. While at first glance, a Creative Commons license like CC-BY-NC might seem appropriate, Creative Commons licenses assume that copyright protections exist for the original work. Because the model is not subject to copyright and a CC license would imply copyright protection, a different means of meeting the university museum's requirements is needed. A contractual agreement could be used in this case. For example, users might need to agree to specific terms of use that require attribution and limits commercial use.

Other Thoughts

In this example, photographs were produced by a single individual with the goal of producing a 3D digital replica of the object. However, there may be cases in which an institution would seek to use photographs taken by many individuals. For example, an institution might seek to crowdsource photos of a popular tourist destination for use in the creation of a 3D model. In the US, the copyright for each photo would generally be held by the photographer. Yet, the use of hundreds of such photos for 3D reconstruction would likely be considered a fair use (see the extended fair use discussion in Case Study 4 below). Is there copyright associated with the 3D model produced by this process? We would argue that there is not likely to be copyright in such a 3D model unless there is some additional creative expression. Some organizations may consider using a click-through agreement or similar simple form that helps the user grant whatever rights are needed in order to use their work product as you design your project.

A 3D model produced in the example above might be used in ways that involve new, creative expression, and those works may be subject to copyright. A **rendering** (a two-dimensional image of the model in virtual space) or an animation of the model would involve choices similar to those made by a photographer and might thus be eligible for copyright. Is this your intended outcome? This choice is ideally documented in writing prior to the commencement of the production or creation of the new creative work.

Case Study 2: Native American Artifact, Tax-Exempt Organization, Chief's Regalia

Background

This case study considers a scenario where an organization partners with a tribal authority to digitize an object for archival preservation. This scenario also applies where a tribal authority requests the digitization of an object. For example, a tribe might seek collaboration or consultation with an organization with technical or subject matter expertise for digitizing an artifact or collection.

Here, a US-based tax-exempt organization (the Partner) is working with a Native American tribal authority (the Tribal Authority) to digitally render and preserve an artifact from the tribal museum's collection in a 3D format. The artifact consists of several pieces of clothing known as the "Chief's Regalia," created in 1890. Certainly, if the object was a more modern creation, where the "life of the author plus seventy" copyright term of protection was used, the analysis below might be different. Additionally, a "utilitarian item," defined loosely as "something that people use," is generally not protectable under copyright. As a result, clothing designs, for example, are not always protected under copyright law. There are, however, exceptions to this utilitarian doctrine. For example, a person that creates fabric can rely on copyright to protect designs imprinted on the fabric if the design features the sufficient amount of creative expression to be protectable.

The Partner's mission is "to help preserve Native American Culture by partnering with tribes to provide training, support and/or services to promote **digital preservation** of artifacts." The organization's rules and ethics are defined and publicly shared via their Native American Collection Policy. The Partner confirms that the artifact is associated with this particular, single Tribal Authority through research and consultation. If the artifact were associated with multiple communities, the Partner would have consulted with those communities as possible before the digitization project moved forward.

The Tribal Authority possesses the Chief's Regalia and seeks to obtain a 3D reproduction through digitization for archival preservation on its own servers and for a term on a third-party depository provided by the Partner. The Partner and the Tribal Authority enter a written Memorandum of Agreement (the Agreement) that outlines their plans and conditions for the project. Here is an overview of some of the key provisions of the Agreement between the Partner and the Tribal Authority for this project.

Object

The object represented by the 3D model (Chief's Regalia) is in the possession of the Tribal Authority and is on display in its cultural heritage museum with seasonal viewing restrictions. The object consists of a feathered headdress and clothing worn on the body arranged on a clothing form. The object is the material that will be reproduced in 3D scans in this project.

- Two-dimensional images owned and copyrighted by the tribe were previously being seasonally displayed on the tribal museum website with a CC license displayed.
- The parties agree to include the existing local community's cultural protocols for access and use to the 3D model; these will be provided by the Tribal Authority in detail.

Copyright

The Tribal Authority takes the position that it is the sole owner of the Chief's Regalia and any reproductions. The Partner does not believe that as a legal matter there is any copyright in the artifact, because of the 1890 date of the creation, but agrees to the Tribal Authority's open license requirements as they do not seem inconsistent with the goal of the project. Further, the license requirements align with the Partner's Native American Collection Policy and ethics, which call for culturally affiliated communities to assert control over their own cultural heritage.

As the source images are taken by a single volunteer and member of the tribe, the Agreement could require that the volunteer transfer copyright and control of the source images to the tribe via contract. This transfer is necessary since the volunteer would, at least initially, own copyright in the photos they created. The overall goal of this transfer is so that the output is managed by the Tribal Authority rather than the single individual. (This is also distinct from cases like *Meshwerks v. Toyota* because this is not work for hire.)

The following licenses and descriptions will be applied to the 3D data (service copies and future reproductions) and the source images. *(Note that each of the following will be further explained in the discussion.)*

- A Basic License—The Tribal Authority wants to assert copyright over the outputs (source images and 3D data) and wishes to also apply the Creative Commons license Attribution-NonCommercial- NoDerivatives (CC BY-NC-ND) to the 3D data and images. However, this license is appropriate only for the source images, not the 3D data. The copyright holder of the source images may use a CC license to allow public uses subject to the CC license. However, there is no copyright in the 3D data. The Tribal Authority should use a basic Terms of Use or other license to meet its distribution goals for the 3D data.
- Traditional Knowledge Labels ("TK Labels")—The 3D data and photos will be assigned TK Labels to incorporate protocols for tribal cultural practices related to the artifact into the **metadata** and for public display. TK Labels are informational or educational tags or badges that help identify and clarify materials that have community-specific restrictions regarding access and use. They help non-Indigenous people better understand and respect different cultural perspectives and concerns about access and use of heritage that derive from the local contexts where the material was made and continues to have meaning.[13]
  - The parties agree to embed and display the "TK Verified" (figure 5.6a) label to let users know that the digital content was created and is responsibly represented with tribal approval. The "TK Seasonal" (figure 5.6b) label will be used, as the artifact has seasonal conditions of access and use.
  - Dates of Display—Consistent with the TK Seasonal label, the Agreement states what times of the year the digital content may be displayed as the Tribal Authority has defined expectations about proper handling and viewing conditions.

Figure 5.6a
Traditional Knowledge Label for TK V (Verified). (https://localcontexts.org/tk-labels/).

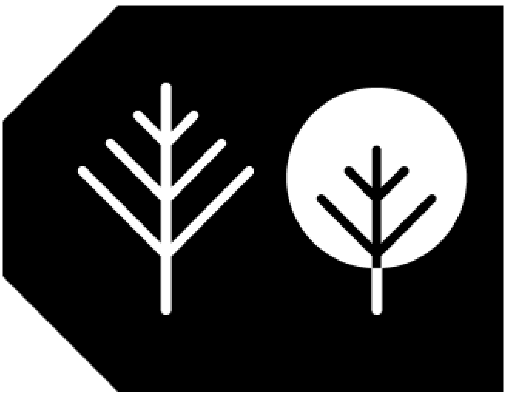


Figure 5.6b
Traditional Knowledge Label for TK S (Seasonal). (https://localcontexts.org/tk-labels/).

PRESERVATION OF MASTER/SOURCE IMAGES

- The Partner will preserve the digitized 3D data master file and source images in long-term digital storage.
- The Partner will use Open Archival Information System (OAIS) standards for a minimum of twenty years in three repositories in different geographic locations with **fixity** checking every six months.
- The same access and preservation protocols provided by the Tribal Authority will apply to all produced 3D master and service copies, all source images, and any future reproductions.
- The Partner will retain a single master copy of the 3D data and source images for long-term preservation as well as a service copy for general permitted uses according to agreed-upon access Terms of Use and protocols as defined in the Agreement.
- The Tribal Authority will retain and possess master and service copies of the 3D data and source images.
- Embed license and protocol information in the source images and all the 3D file metadata.

- Master and service copies of textual descriptive information will be retained with the files in .pdf format. Any oral descriptions provided by the Tribal Authority will be preserved in .wav format.
- Terms of Use or appropriate licenses and Traditional Knowledge (TK) Labels will be used with all files associated with the artifact including but not limited to the 3D data and source images.

Process and Output

3D MODEL AND SOURCE IMAGES

- The 3D model will be comprised of 900 source images captured via a photogrammetry-based process shot in RAW format (used at the Partner's request).
- The source images will be taken by a volunteer member of the Tribal Authority after online training and after the test shots are uploaded and analyzed.
- Partner will review, approve, and discard the test images. Then, a new set of production photos will be uploaded to the Partner by the tribal volunteer photographer.
- Upon submission to the partner, images will be initially be saved in two locations as working and preservation copies. All copies will include the rights, licensing, and TK Label information as described in this agreement.
- The 3D processing team then will create the digital 3D data from the working file source images.

TREATMENT OF SOURCE IMAGES

- Source images will be uploaded in RAW format to the Partner via its submissions web page. Corresponding metadata question forms will be completed by the submitter.
- Available fields for source images should include photographer information, affiliated community, and image copyright restrictions as well as artifact descriptive information provided by the Tribal Authority.
- The photographer's name (though not associated with the rights statement in favor of the community name) will be embedded in the metadata for reference purposes.
- If desired, an audio file can be uploaded for expanded oral description of the object with appropriate copyright ownership of the creator, licensing, and TK Labels.

ARCHIVAL PROCESS

- Upon completion of the 3D data production, administrative, descriptive, and reference metadata will be embedded as described in this agreement.

- The Partner will archive in a **digital repository** a preservation copy of the data and the specific source images in accordance with the Agreement.

Usage

ACCESS

- The Partner may display the 3D data; the parties acknowledge and agree that all metadata, website description, and depictions will be reviewed for proper context in consultation with the Tribal Authority. The Tribal Authority will verify or amend as necessary.
- The Agreement is signed by the authorized representatives of the Partner and the Tribal Authority. Each party receives a copy of the signed agreement.

Issues/Discussion

Copyright is a form of protection to authors and creators of "original works of authorship." A work is automatically protected by copyright when it is created, that is, "fixed" in a copy or the first time. Copyright protects original "pictorial, graphic, and sculptural works," which include two- and three-dimensional works of fine, graphic, and applied art. Neither registration in the Copyright Office nor publication is required for copyright protection. As a result, the Chief's Regalia was certainly protected by copyright when it was first created.

However, the copyright protection that exists in the Chief's Regalia does have a limit. Generally, copyright protection ceases after a certain period, which is defined by the applicable law at the time of creation. Duration of copyright has expanded over time, from an initial fourteen years in the first copyright act, to seventy years after the creator's death under modern copyright law. Once the term has ended, works are no longer copyright-protected but have dropped into the public domain. Anyone can use a public domain work, for any use, without obtaining permission from the original creator. Here, the Chief's Regalia, with an estimated creation date of 1890, is in the public domain as a matter of copyright law, although the tribe philosophically and culturally rejects the concept of public domain.[14]

Yet, despite the lack of copyright and the object's public domain status, the Tribal Authority does still hold possession of the work and can limit access to reproductions through a basic license that meets the goals of public access. As noted above, while the Tribal Authority wants to assert copyright over the object and wishes to also apply the Creative Commons license Attribution-NonCommercial-NoDerivatives (CC BY-NC-ND) to the 3D data, this is not an appropriate license for this type of use. The CC license is for the copyright creator or owners to allow uses subject to the CC license. Here, however, there is no copyright in the original object, because of the public domain, or in its 3D data. The Tribal Authority could use a basic Terms of Use or other out-of-copyright license to meet its distribution goals, allowing certain types of usage, access, and attribution, but without any of the copyright language.

Additionally, there may be some copyrighted materials as part of the project. As part of the photogrammetry process, a series of several hundred photographs will be taken by a member of the Tribal Authority after online training and after the test shots are uploaded and analyzed. It is common for photographs of 3D objects to have copyright protection (see discussion earlier in the chapter regarding *Burrow-Giles Lithographic Co. v. Sarony*), so the *photographs* of the Chief's Regalia are likely copyrightable. To be sure, the individual volunteer, a member of the tribe, may transfer rights in their photos to the Tribal Authority. Or the Tribal Authority, after training the volunteer, may have the volunteer sign a "work for hire" contract.[15]

Case Study 3: Natural History Specimen, Institutional Repository, Preserved Fish Specimen

Background

This case study is intended to represent a situation in which an individual submits a constructed 3D model representing a natural history object (in this case, a preserved fish specimen) to an institutional repository. This scenario is very similar to that of Case Study 1 but differs critically in terms of who generates the 3D data and who is interested in keeping a copy of them. Instances like this would be commonly found in museum and university settings. The location where the specimen was collected and the institution itself are assumed to be in the United States. The institution owns the specimen, but the individual who created the 3D model of the specimen is not a staff member or affiliated with the institution.

3D Model

The individual constructed the 3D model (e.g., PLY file) from photographs or other scan data (e.g., CT image stacks or lidar data). There may be multiple versions of the 3D model at different sizes and resolutions. The photographs or scan data used in the model construction process will be archived along with the constructed 3D model. Metadata documenting data capture and model construction methods will also be archived alongside the 3D model.

Object

The institution owns the preserved fish specimen represented by the 3D model. At the time of acquisition, the donor signed a deed of gift asserting that they (the donor) were the true and legal owner of the specimen and that they had the right to convey it.

The donor irrevocably and unconditionally transferred ownership and all rights, title, and interest in the specimen to the institution through the deed of gift.

Specimen Loan/Access Agreement

When the individual borrowed the specimen from the institution, they agreed to provide the institution with a copy of specimen images and data that they created as part of the formal loan agreement. The loan was in accordance with museum policies for this use.

Media Submission Process

The institution's goal is to archive a 3D model and associated raw media in perpetuity and to make it available for research or education purposes (noncommercial use). The individual who constructed the 3D model shares the following files with the institutional repository:

- 3D model (e.g., PLY file)
- photographs and raw scan data (e.g., CT image stack, lidar data)
- documentation on who created the media and how
- a signed release form transferring copyright to the institutional repository
    - The individual may request an embargo period of up to three years.
    - If the model or media involve other conditions (e.g., cultural sensitivity), access restrictions can also be requested or otherwise applied.

Archiving Process

The model and media are stored in the institution's digital asset management system ("DAMS"), where it can be associated with related specimen data. Storage and management systems follow institutional policies for redundancy, data validation, and security.

Public Access

The institution aims to provide access to the model (and appropriate credit information) for research and educational (noncommercial) purposes. Institutional data norms are as follows:

- Published collections data are public domain.
- Published collections media are copyrighted by the institution and licensed "CC-BY-NC."

A preview of the 3D model is published online alongside related specimen data following any embargo or other restrictions as indicated in the release form. The model and raw media files themselves are provided to users upon request. (Due to file size limits and manual review of usage requests, the model and media files are not immediately downloadable.)

Commercial usage (if approved) requires documented permission and payment.

Issues

In this case study, a natural history object, a preserved fish specimen is the subject of digital reproduction from photographs and CT image stacks. As an unadorned work of nature, the object itself is not subject to copyright protection. The *Compendium of U.S. Copyright Office Practices* explicitly excludes works "produced by nature, animals, or plants" from registration.[16] However, the lack of protection under copyright law does not indicate that the object is free from all restrictions. In some cases, prior contractual agreements may limit how the object can be used. For example, a cast of a specimen acquired from another museum may be encumbered by a contractual agreement stating that no additional duplicates are to be produced, or if the fish specimen was collected on public land, the National Park Service or other government agency may place limitations on the production of copies. Here, the institution owns the specimen, but the individual who created the 3D model of the specimen is not employed by or affiliated with the institution.

As part of the process of creating the 3D mesh, a series of several hundred photographs and CT scan data will be produced by a third-party scanning firm. While the raw data are not protectable under copyright (see earlier discussion of *Meshwerks, Inc. v. Toyota Motor Sales U.S.A.*), most photographs of 3D objects are subject to copyright protection as a result of the creative judgments made by the photographer in the process such as angle, lightning, exposure, and other factors (see discussion earlier in the chapter regarding *Burrow-Giles Lithographic Co. v. Sarony*). Therefore the photographs of this preserved fish specimen, produced by the individual employee of the scanning company, are probably protected by copyright. So, unless the contract indicated otherwise, the photos, at least initially, would belong to the scanning company taking the photos until they are legally transferred to the museum in some capacity, typically via a copyright licensing agreement (see the section on contracts and licensing above).

The photogrammetry software extracts data, such as feature positions and depth maps, from the set of photographs to produce a 3D mesh with color information. This mesh can be considered as a "copy" of the physical object; the intent is to create an accurate, three-dimensional, digital representation of the physical object. Copies or slavish reproductions are not original works of authorship (see earlier discussion of *Meshwerks, Inc. v. Toyota Motor Sales U.S.A.*), and *new* rights are not created in copies. Because the preserved fish specimen itself is not subject to copyright protection and no new rights are created in the copy, the 3D mesh produced in this project is not subject to copyright.

The museum intends to share the 3D mesh publicly for research and educational purposes, but the scanning company (and the donor) might want to be credited appropriately when the mesh is used. While at first glance, a Creative Commons license like CC-BY-NC might seem appropriate, Creative Commons licenses assume that copyright protections exist for the work. Because the 3D model is not subject to copyright and a CC license would imply copyright protection, a different means of meeting the museum's requirements is needed. A contractual agreement could be used in this case. For example, users might need to agree to a Terms of Use statement that requires attribution and limits commercial use.

For any of the photographs taken during the scanning company's process in building a 3D model of the preserved fish specimen, it more than likely has a copyright in those photos. Certainly, if the museum would want to utilize some of those photos, the scanning company could transfer copyright in the photos in a basic copyright transfer agreement. However, if the museum had, in its initial contract with the scanning company, a "work for hire" provision, then all the copyrighted work, including the photos, would be owned by the museum. Work for hire is defined in Section 101 of the Copyright Act (title 17 of the U.S. Code) in two parts: (1) a work prepared by an employee within the scope of his or her employment or (2) a work specially ordered or commissioned for use. The scanning company would most likely fall under the second section, and that would require that both parties, the museum and the scanning company, expressly agree to a work-for-hire clause in a written contract.

Other Thoughts

In this example, photographs and CT scans were produced by the scanning company with the goal of producing a 3D digital replica of the object. At the end of the process, we would argue that there is not likely to be copyright in such a 3D model unless there is some additional creative expression. Some organizations may consider using a click-through agreement or similar simple form that helps the user grant whatever rights are needed in order to use their work product as you design your project.

Case Study 4: Researcher-Developed 3D Model

Background

This situation represents the creation of a 3D model of an ancient Egyptian complex for research and teaching purposes. The resulting 3D model is based on a combination of bibliographic research such as drawings, sketches, photography, and plans, as well as on-site 3D scanning and 3D hypothesis based on previous and new research. The goal of the 3D model is to provide a visual representation of a no-longer-available/disappeared/destroyed/ephemeral cultural artifact or environment. The scanning was generated by a US institution; the research is based on bibliographic resources from around the world; the current creator of the 3D object is US-based.

Object

The object was created using measurements taken during an archaeological investigation in the 1950s funded by the University of Caltexico. The specific object in this case is a 3D reconstruction of mortuary complex of a pharaoh in the Valley of the Kings. The object created in a representation of a cultural artifact inasmuch as a new scholarly monograph or performance is a new cultural artifact. It is an amalgamation of scholarly practice and creative research. The 3D model reconstruction of the mortuary complex is intended for publication as a whole and complete 3D reconstruction to represent the total hypothesis as well as segmented derivatives that highlight specific areas for scholarly provocation. These segmented 3D objects that are derived from the 3D reconstruction are new data that do not exist physically. The resulting 3D reconstruction is an amalgamation of 3D scanned materials, scholarly research, and projections in the 3D environment based on the researcher's hypothesis as well as other scholars' 3D reconstructions to create the whole.

Process

The measurements taken during the investigation were incomplete and did not allow the recreation of a complete elevation of the mortuary complex. Therefore, a combination comparison to other like sites in the area, 3D scans of relevant artifacts held at the University of Caltexico, and historical research were conducted by an interdisciplinary team of faculty and students at the US-based public Xavo University to fill out missing details.

Output

The output is an annotated 3D model reconstruction published online as an educational resource. There is a copy in the Xavo University institutional repository with associated creation metadata listing authors (the faculty and students of Xavo), and the bibliographic records used to create the model, i.e., University of Caltexico's archaeological reports, the other sources used to create the model, and any new measurements created to fill in any missing data to complete the mortuary complex elevation. The output includes an .obj file of the completed model, spreadsheets that represent the measurements and dimensions used to create the model, and a .txt or .csv that represents the bibliography, including links to other 3D scans imported to the model as well as a .json file for 3D annotations.

Usage

The 3D object will be published by Fordstan University Press. Xavo University holds a copy of the 3D output in the institutional repository, but because of the upcoming publication of the 3D reconstruction, the content will be embargoed for a specified amount of time. After the embargo, the 3D object will be made fully publicly accessible free of charge via Xavo University institutional repository.

Discussion

Here we have an amalgam of copyright and non-copyright works mixed together to form a virtual architectural 3D model. Certainly, there could be, as with the three cases above, many questions as to the copyrightability of the data, including measurements, notes, photos, and the scholarly guesswork that filled in any missing information. Arguably, some could have copyright protection, and some could not. However, in this case, we will focus on the transformative nature of the use of all these works to create the virtual architectural 3D model. It is more than likely that this is considered a fair use of the copyrighted works.

At its core, fair use ensures that there are some kinds of uses that do not require permission or payment and provides flexibility for users and new creators. There are, however, no easy rules for fair use. The source of fair use law is statutory: Section 107 of the Copyright Act provides that fair use of a work "for purposes such as criticism, comment, news reporting, teaching . . . scholarship, or research" is not copyright infringement. This list is not exhaustive; other uses of copyrighted work without permission may also be fair. Section 107 further provides:

> In determining whether the use made of a work in any particular case is a fair use the factors to be considered shall include—
>
> (1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;
>
> (2) the nature of the copyrighted work;
>
> (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and
>
> (4) the effect of the use upon the potential market for or value of the copyrighted work.[17]

Since the doctrine is an equitable rule of reason, no generally applicable definition is possible, and each case raising the question must be decided on its own facts. This examination of the four factors determines whether the use is fair or constitutes copyright infringement. Courts weigh each factor and make a decision based on the overview of all four factors. In this 3D model creation context, this four-factor test is used as a form of risk mitigation. By reviewing the four factors as a court might, a 3D data creator can determine whether or not the action she is taking might risk infringement or fall squarely within the realm of fair use.

In recent years, US courts have focused increasingly on whether an alleged fair use is "transformative." A work is transformative if, in the words of the Supreme Court, it "adds something new, with a further purpose or different character, altering the first with new expression, meaning or message."[18]

There are various ways that copyrighted third-party material—such as images, text, videos, and sketches—can be used transformatively. The key is the repurposing of the copyrighted material to advance a point made in the presentation, lecture, chapter, or other work. Imagine, for example, taking a college music course on jazz. A jazz album may be written for a specific purpose: to share artistically created melodies for entertainment. However, if you are in a jazz class in college, the use of jazz songs is not for entertainment. The class may analyze, comment on, criticize, and compare the music. Commentary on styles, historical development, instruments, and clips from various famous jazz composers would be part of the curriculum. This class on jazz, and the use of clips and songs in lectures, would be repurposing the original material that was for entertainment for a new and different purpose: the scholarly study of a unique form of American music.

This repurposing comes in many forms. As illustrated, the work could be the subject of the instructor's analysis. In that case, the material is necessary because the instructor is analyzing, critiquing, or explaining it. Or the material could illustrate the instructor's point or help to make it more comprehensible. These examples are not exhaustive. The key is that the material is being repurposed to significantly advance the instructor's own point.

This could be readily adaptable to projects that make 3D reconstructions made up of multiple sources of copyrightable materials, some 3D data, the integration of other 3D models, and non-copyrightable materials from other 3D facsimiles. Materials that are in the public domain or unprotectable under copyright can easily be used. For materials that are still under copyright but are necessary for inclusion in the project, their fair use is all dependent on the type of use and the potential for repurposing those third-party copyrighted materials to serve this new use: building a complete 3D reconstruction.

In these scenarios the copyright judgment is indeed more complex. A best practice for determining copyright, or other related rights, is to consult the institution's intellectual property rights (IPR) of the physical objects. Other rights to consider as part of this reconstruction include the IPR of the 3D reconstructions integrated into the completed reconstruction.

Only a good metadata schema could keep track of all possible IPRs. Metadata give use information about authorship, creation date, publication, acquisition, and other information that aids the user in making a judgment as to the potential copyright. And it certainly is a matter of good scholarly record to cite to any materials used—whether it's footnotes in an article or in the creation of a 3D reconstruction work. A single metadata record should be associated with each part of the digital reconstruction. A final record should register the IPR of the complete reconstruction with appending IPR for the reconstruction's subparts to better identify the digital objects' **provenance**.

For a similar example including US and non-US sources, see appendix A (Case Study 5).

3D Architectural Models: Further Discussion

3D models produced during the design of new buildings present different copyright and ownership concerns compared to 3D models of non-copyrighted historical structures, archaeological sites, or scientific and cultural objects. Models in building design are often developed as parts of a web of referenced and linked files from many sources, which complicates copyright considerations.

Architectural records, including 3D models, fall under the class of cultural works afforded protection by copyright law. In most countries, the buildings themselves are also protected under copyright. In the US, the protection afforded to architectural documents was not extended to architectural works themselves until 1990, so only buildings designed or built after this date are protected. Copyright law allows the creator of an architectural work, like the creator of other cultural objects, the exclusive right to reproduce the work (including to build the work), create derivatives, and display the work publicly. In some countries, including the US, a Freedom of Panorama provision of the copyright law allows photography, video recording, or otherwise capturing and reproducing buildings located in a public place without the permission of the copyright holder. Keep in mind that some countries, such as Argentina, Belgium, France, and Greece, have a much more limited Freedom of Panorama right, which prevents commercial or other uses. Or, in the case of Italy, there is a highly restrictive law against publishing any pictures of artworks that are in public space without explicit permission.

In the context of the contemporary design process, 3D architectural models are fully copyrighted materials produced in the course of an architectural firm's business, and as such are embedded in complex layers of rights, ownership, and licensing issues that govern their access and use. These terms are defined in contractual agreements between a client and an architect, between an architect and a contractor, and between an architect and their sub-consultants. Typically, architects maintain both copyright and ownership of the 3D models and other documents they produce when designing a building; these documents are licensed to their clients as "instruments of service" toward the creation of a work of architecture. Sub-consultants or subcontracted engineers will in turn license their own documents to the architect in a similar manner. Careful control of access and use of design documents helps protect from loss of fees, damage to reputation, or liability for building safety that an architect might incur if the documents were used to create a building outside of their supervision or permission.

Today's digital design processes introduce additional copyright concerns. 3D architectural models may be authored in proprietary and even custom-built software, with important data that might be lost in translation to open formats. Advanced methods like building information modeling (BIM) can result in complex documents with multiple authors and contributing disciplines that might have copyright over the material. Architects often develop their 3D model on top of site models and spatial data provided by the client or sourced elsewhere; similarly, vendors of building components are increasingly producing pre-rendered 3D models of their wares for architects to incorporate directly into their 3D model.

Similar copyright and contractual issues govern non-digital architectural material, but since architectural works on paper are typically accessed in the reading room as originals, rights issues often don't arise until a work is selected for traveling exhibition or publication. In contrast, digital architectural practice and its digital archival practice require archivists to address copyright and ownership at each step of acquisition, preservation, and access.

Conclusion

Interaction between law and modern 3D technologies is in its infancy. Additionally, the laws that are on the books, including copyright, contract, licensing, and others, were developed in the past, prior to the advancement of our modern 3D processes and tools. There is very little legislation and very few cases that deal directly with the specific factual scenarios we outline. However, in common law systems like the US, we have the ability to adapt and use the law we have to interpret our work.

This chapter, hopefully, lays the legal and procedural groundwork for conversations surrounding the cutting-edge 3D work being done in our cultural institutions. We can't do our work unless we know what the law actually says. And we can also use this knowledge to myth-bust any long-held beliefs and assumptions about the law that could jeopardize our institution's mission. Lastly, we then learn how to harness that law to carry out our organizational mission.

Here, we reviewed some of the most relevant laws and cases that cultural institutions can utilize to make their 3D works available to the world. Some are not 3D-related cases or laws, but they still have enormous value to our dialogue. With the fundamental copyright policies that eliminate most prohibitions on sharing facts, data, and other information, combined with an understanding of where protections can still exist, a cultural institution can continue to play its role in being a balancing mechanism between rights and access. With a greater understanding of licenses and contracts, we can still preserve some of the values associated with our collections. Additionally, we can provide both donors and users with the proper information they need, including any potential concerns surrounding questions about rights, future uses, and long-term access. Determining and documenting rights associated with 3D data is a critical component of long-term preservation.

Being able to identify the key legal issues is the first step toward structuring an approach. As you can see from the case studies, the legal and substantive questions are often intertwined. We sought to keep scenarios as simple as possible in order to identify essential common elements that would help articulate community practices. This was remarkably difficult. However, with these case studies, much like the case law, we want them to be used to benchmark some of the other 3D projects happening in cultural institutions. Read them, examine the law, distinguish the outcomes, and use them as a template or ground floor for the discussions that often surround new and challenging 3D data work in our workplaces.

APPENDIX 5A

Case Study 5: Integrating Different 3D Technologies That Imply Several Different IPRs

We have included this scenario as an appendix because of the extended detail.

Background

This case study represents a scenario characterized by the integration of different 3D techniques and processing with different 3D digital objects, such as the virtual reconstruction of the monument and the digital replicas of existing structures and statutes located in the museum. In this context, different IPRs (intellectual property rights) can be identified since different digital provenances can be isolated and recognized.

Object

The case study is the virtual reconstruction of the so-called Basilica in Herculaneum. The monument is not visible with the exception of the southern side, which is only partially preserved. The building was explored in eighteenth century by surveyors, and, 250 years after its discovery, it is still largely unexplained. The identification and function of this structure have been disputed since its discovery.

3D Modeling and Data Acquisition

The 3D model has been built on the base of the final architectural drawing and incorporates reconstructions proposed by some scholars. The CAD model has been integrated by the 3D model of the quadrifrons, still visible in the southeastern part of the building, carried out by close-range photogrammetry. Thanks to common 3D modeling tools (extrusion, loft, sweep, Boolean operations, etc.), it has been possible to generate surfaces and solids from the lines for both the general structure of the building and the several more detailed decorative elements such as capital with acanthus leaves, column bases, and altar cornices. 2D CAD plans, sections, and elevations have also been used to support the creation of the building with its correct dimensional, formal, and geometric characteristics. Simultaneously the 3D replicas of the quadrifrons and of four statues, originally placed inside the Basilica, have been carried out with close-range photogrammetry.

Virtual Reconstruction

All the 3D objects have been imported into Rhinoceros for the integration of the CAD model with the 3D replicas of the statues and quadrifrons. The integration has been based on the alignment of the models on the basis of common elements, such as corners and edges of the two altars on the southeastern side. Because of lack of information about the surface, only two materials, white marble and white plaster, have been applied in order to have a realistic textured model of the building, while, according to the literature information, the frescoes have been placed into the two apses by texturing surfaces. The one element not included in the original drawings is the tiled roof of the peristyle, which has been added in order to have a complete representation of the building.

Output

The 3D objects and the final reconstruction have been carried out in the framework of the European project 3D ICONS. The reconstruction aims at highlighting some controversial parts of the monument. Metadata associated with the digital replica describe the physical object and register all phases from data acquisition to data visualization. The metadata record deals with the 3D process, which includes 3D data capture, post-processing, and publication.[19]

Issue

In order to identity possible and multiple IPRs with related issues, it is necessary to clarify how the virtual reconstruction has been carried out and what the sources, physical and digital, are.

Short History of the So-Called Basilica

The monument is known in the literature with different names: Porticus, Forum, Basilica; the latest research identifies it as a building linked to the imperial cult, precisely the Augusteum, even if there is no epigraphic evidence that testifies to this function. The building has been investigated only thanks to the well-known system of eighteenth-century tunnels; therefore, it is still largely buried under a thick layer of volcanic deposits.

In 1744 Bardet de Villeneuve (who directed the excavations between 1741 and 1744) drew up three plans of the building. The monument had been excavated for the first time a few years earlier in 1739 by de Alcubierre; likely Bardet's plans referred to maps drafted at that time. This hypothesis is supported by a number of significant inaccuracies. The three drawings represent the structure in its entirety; two of them reproduced the monument in a broader context of public buildings, including the theater. At first glance, the misplacement of the buildings is apparent: the theater (rotated 90 degrees), the Porticus, the front of Collegio degli Augustali, and the so-called Basilica Noniana are too distant from each other and, consequently, from the Decumano Massimo. Even if the general map is incorrect, by contrast the Augusteum drawing is detailed and highly accurate; it also includes a front view of the west wall of the building.

Between 1750 and 1751, J.-C. Bellicard and C.-N. Cochin visited Herculaneum. In their publication in 1754, there is a much more accurate plan of the northwest area of the excavations. The buildings were put into a more correct spatial relationship. It is likely that the two gentlemen had access to the plans of the original excavation carried out by Alcubierre, which not long afterward disappeared, together with the general plan of the excavations provided in 1759 by Weber. By comparing the plans of Bardet and Bellicard it is possible to suppose that the two surveyors used different maps created by the early excavators.

An evocative representation of the building is in an engraving made by F. Morghen in 1835. The work is a bird's-eye view of the great porticoed building, with, inside, some statues in place, including the two equestrian statues of M. Nonius Balbus and son. The two side bases on which the artist placed the statues (actually coming from the public area of the city) are imaginary. By checking the detailed map provided by Bardet's survey, instead of the bases, there are two small platforms placed against the western and eastern walls of the porticos and accessible by two steps. It is probable that these platforms were tribunalia rather than bases for statues.

In the engraving (which represents the building completely dug up), there are some inaccuracies also on the long walls. The Bardet map shows a detailed description of the western wall with five large curved niches open to the ground alternated with groups of two or three small arched and rectangular niches. The large niches are surrounded by pilasters with Corinthian capitals and have on the top of the arc a rectangular space, which is filled, in the second niche, with the inscription of dedication of the Augustales discovered in 1741 (not in situ).

The structure has been revisited by different scholars, who provided an axonometric plan of the monument and the first complete digital reconstruction.

This reconstruction focused mainly on the comparison of the eighteenth-century plans with the extant remains and therefore on the accuracy of the previous maps. The correctness of the geometry of the eighteenth-century drawings has been checked, and some mistakes have been highlighted in the reconstruction of the southern part of the monument. Thanks to the new survey, seven openings have been identified in the southern facade instead of five as drawn by the eighteenth-century surveyors. The published 3D digital reconstruction is very essential and without any decoration.

On the basis of this reconstruction, recently a wooden model of the building in 1:50 scale has been provided. For the first time the model is enriched by adding frescoes, statues, and other decorative elements.

The correct assignment and positioning of the statues and frescoes are still debated. The eighteenth-century surveyors dug many tunnels simultaneously in several parts of ancient Herculaneum, and they often exploited these tunnels many times during the works. Therefore, they did not always register correctly the provenance of the objects they found.

As regards the Basilica, only a few data are surely correct. Inside the central exedra (a semicircular room or portico) of the north wall of the building, the excavators found a group of imperial statues in marble, two seated (Augustus and Claudius), and a third one *loricata* (Titus in armor). Two other statutes, representing Augustus and Claudius, were probably on the bases set in front of the niches. The frescoes (readily detached by the eighteenth-century surveyors) also have debatable provenance and few data are available for precise location. For example, four large pictures, whose surface is slightly concave, can surely be placed in the two niches in the bottom of the north wall of the porticos. All statues and frescoes are currently stored in the Archaeological Museum of Naples.[20]

Discussion: Possible IPRs

The historical plans are stored in public archives, the statues and frescoes belong to the Museo Archeologico Nazionale di Napoli, and the visible structures (quadriphon and the south facade) of the so-called Basilica are under the protection of the Parco Archeologico di Ercolano.

Surely as regards the physical objects, the IPR belong to the public institutions (archives and museums). Under Italian law, public institutions can authorize private or public institutions to carry out 3D digital replication of any object under their control. So these institutions maintain their rights in the 3D digital asset.

Furthermore, in this reconstruction, other rights can be identified:

- the rights of the scholars who provided the first digital reconstruction (available only as figures in their publication) and the CAD model based on historical maps. (The CAD model has kindly been provided by the architect.)
- the rights for the individual parts of the reconstruction: the 3D modeling from the CAD and the data acquisition of the quadriphon and the statutes
- the IPR for the complete virtual reconstruction

This case study avoids assigning IPRs on derivative models (for instance a decimated model) as this is an automatic procedure included in the software. This case study also does not deal with the photos shot for the close-range photogrammetry.

Only a good metadata schema could keep track of all possible IPRs. A single metadata record should be associated with each part of the digital reconstruction. A final record should register the IPR of the complete reconstruction. A hierarchical parent-child system should be chosen to better identify the workflow of the digital provenance.

APPENDIX 5B

Field Museum Media Creation and Assignment Agreement

The Field Museum's media agreement to document permission for the museum to archive media, including 3D data, from non–Field Museum researchers. (Field Museum of Natural History, "Final Media Creation and Assignment Agreement," August 26, 2019, https://mm.fieldmuseum.org/b3198478-7efc-4896-b0ae-ca1a9431263f.)

FIELD MUSEUM OF NATURAL HISTORY

MEDIA CREATION AND ASSIGNMENT AGREEMENT

The Field Museum of Natural History ("Museum") makes its collections available to qualified researchers for academic purposes; it also makes its collections data and media publicly available under its Data Norms. As a condition of accessing or borrowing Collections Items, you are agreeing to assign the rights in Media created from or of these items so that the Museum may, at its discretion, archive the Media and distribute it to the public under its Data Norms, i.e., under a CC-BY-NC License.

For further information, see frequently asked questions (https://tinyurl.com/y2qhwj9x).

As a condition of borrowing or accessing Collections Items, you, the undersigned, agree to the following terms and conditions.

A. Definitions.

Access Agreement refers to a Loan Agreement, Research Access Agreement, or similar document specifying the Collections Items you will be working with, the purpose of your use, and other details of your loan or research visit.

CC-BY-NC License refers to the Creative Commons Attribution-Non-Commercial 4.0 International License, full text available at https://creativecommons.org/licenses/by-nc/4.0/legalcode . The Museum makes multimedia served publicly from its authorized Collections Management System available under this license, which allows reuse without further authorization provided the use is non-commercial and credits the Museum as required in the Data Norms. This is a broad, flexible license that allows for most common academic uses of the Media.

Collections Items refers to specimens and artifacts accessioned or deposited into the Field Museum of Natural History's scientific collections.

Data Norms refers to the Museum's Conditions and Suggested Norms for Use of Collections Data and Images, available at https://www.fieldmuseum.org/field-museum-natural-history-conditions-and-suggested-norms-use-collections-data-and-images .

Media refers to all media created from or of Collections Items and derivatives of that media. This includes, but is not limited to, (a) representations, such as photographs, audio/visual recordings, scans and raw scanning data created by various means, or other recordings or images created by technology now known or hereafter developed and (b) derivative assets, whether physical or digital, such as models, images, molds, casts, three-dimensional printing files, etc. Media does not include facts (e.g., measurements, observations, or genetic sequences) or academic research, writing, or other work product other than the representations and derivative assets as described above.

Researcher refers to any individual borrowing or accessing Collections Items for non-commercial, academic purposes or, in some circumstances, cultural groups accessing Collections Items for purposes related to cultural traditions or pursuant to a repatriation request. A Researcher's qualification for access to Collections Items will be assessed in accordance with the Museum's Collections Management Policy.

B. What You Are Agreeing To

You hereby assign all of your rights and interest to the Museum in any Media created of or from the Museum's Collections Items. This includes assigning all copyrights or other intellectual property rights that may exist and waiving any moral rights you may have in the Media.

You will ensure that any person named on your Access Agreement and, in addition, any students, co-researchers, subgrantees, subcontractors, and other agents who work with the Collections Item or create Media of or from the Collections Items assign their rights and interests in the Media to the Museum and abide by the terms of this Agreement.

You will provide Media requested by the Museum in a mutually agreed upon format and timeframe. Media must be provided within two years of the submission of your final grant or project report; in the absence of such an end date, the Media must be provided within two years of the date of the Media's creation, unless agreed otherwise.

You will use the Media only as allowed by the CC-BY-NC License and Data Norms unless your Access Agreement specifies different terms of use, in which case those terms apply; any other use requires the Museum's prior written permission.

You will, in all uses of the Media attach, the correct citation as required by the Data Norms.

You may use and distribute the Media under a CC-BY-NC license unless your Access Agreement states otherwise, regardless of whether the Museum requests the Media from you or makes the Media available to the public.

C. The Museum's Responsibilities

The Museum will credit you as the creator of the Media.

Preferred Credit Line: ____

The Museum will treat any Media that it incorporates into its Collections Management System with the same care and under the same policies and procedures it uses for the Media it creates.

Page 2 of 4

Last updated 8.23.19

The Museum does not guarantee that it will archive or distribute any of the Media it requests or accepts from you or that it will archive or distribute it in perpetuity. The Museum strives to adhere to the field's common standards and best practices related to accuracy, transparency, and open access, but makes no guarantee or warranty that the Media, as archived or distributed, will be free of errors or technical defects or will be fit for any particular purpose. You are allowed to create and maintain backup copies of the Media.

D. Outreach and Publicity Opportunities

If you or your institution plan to engage in a formal outreach or publicity campaign ("PR activities") (e.g., press releases, publications by news media, social media campaigns) using the Media or concerning the Collections Items, you must notify the Museum and allow the Museum to approve the use, such approval not to be unreasonably withheld. If the Museum does not specifically object to your proposed PR activities within five business days, the use is deemed approved. The Museum may, at its own expense, choose to coordinate with such PR activities. Please communicate your plans to the Museum with as much lead time as possible, but at least seven days in advance. For other, informal public communications (e.g., isolated social media posts, incidental references in a blog post or public program), you need only cite the Museum as required by the Data Norms (https://www.fieldmuseum.org/field-museum-natural-history-conditions-and-suggested-norms-use-collections-data-and-images).

Requests should be sent to Museum staff who authorized the loan and the PR contact listed on the Access Agreement.

E. Special Circumstances

1. *Embargos*

Upon request, the Museum will embargo distribution of the Media for three years after Media creation or until the date the Media is published, whichever is sooner. After the initial three-year term you may, for reasonable cause, request up to three additional one-year extensions.

______ Embargo requested. Expires ___________________________

2. *Additional Restrictions*

The use of some Collections Items may be further restricted due to cultural sensitivity issues (e.g., human remains, sacred objects) or the Museum's contractual or legal obligations (e.g., conditions on accession or collecting permits, third party intellectual property rights, Nagoya compliance requirements). Nothing in this Agreement should be construed as superseding or obviating any requirements of your Access Agreement or in any other previous or subsequent agreements between you and the Museum concerning any restrictions placed on your use of the Collections Items or the creation or use of Media derived from such Collections Items.

Such restrictions will be defined in your Access Agreement. If a term in this Agreement conflicts with a term in your Access Agreement, the terms of your Access Agreement control.

I have read and understand the above and agree to these terms and conditions.

For Researcher/Borrower:

____
Individual Name

____
Institution Name

____
Date

For Field Museum:

____
Responsible Employee

____
Title

____
Date

Contact Information:
PR Contact: media@fieldmuseum.org
Collections Contact: [Refer to Access Agreement for Museum staff who authorized the loan/access.]

Last updated 8.23.19

APPENDIX 5C

MorphoSource Permissive Usage Agreement

One of nineteen usage agreement options available on MorphoSource.

MorphoSource Permissive Usage Agreement

I agree to comply with each of the following basic requirements (1, 2, and 3) below whenever I download or access from MorphoSource any file(s), image(s), or model(s) (each, individually and all collectively, as applicable, referred to as "data") for use in any release, research paper, educational course, presentation, analysis project, website, blog post, social media post, or other public communication in the form of measurements, analyses, discussion, or 2D and 3D figures. I also agree that if I create digital or physical derivatives of the data (as distinguished from "data" or "original data"), I will comply with standards 1 and 2 immediately below, as well as the additional citation requirements and restrictions listed further below specifically referencing derivatives of data.

Basic Requirements

1. Any instructions for data citation and funding acknowledgment found on the relevant MorphoSource page for the files.

2. Any restrictions and requirements of any copyright licenses or usage agreements that may be attached to these files

3. Additional citation requirements and restrictions of use that are listed below.

Additional Citation Requirements and Restrictions of Use:

a. **Acknowledgement of Source**. Any release, research paper, educational course, presentation, analysis project, website, blog post, social media post, or other public communication referencing the data and/or derivatives of the data must clearly and explicitly identify MorphoSource, the collection organization (i.e., the specific collection, department, or laboratory), and the collection institution (i.e., the museum or university) as the original source of the data or derivative. This same acknowledgement must be made for data and/or derivatives of data originally sourced for MorphoSource that have been re-archived in third-party repositories (see below for any further restrictions on re-archival of this data or derivatives, if any exist).

b. **Citation**. Any release, research paper, educational course, presentation, analysis project, website, blog post, social media post, or other public communication referencing the data and/or digital or physical derivatives of the data will label, cite, or otherwise be captioned both with the institutional catalog number of the physical object or specimen that the data or derivative represents (preferably using the "Darwin Core triplet" of [Institution Code]:[Collection Code]:[Catalog Number]; for example, MCZ:Herp:A-12345), and the MorphoSource media identifier(s) of the downloaded file(s). Specifically, the MorphoSource *Digital Object Identifier* (DOI) should be included when available. If a DOI is not available for the file object in question, then a MorphoSource *Archival Resource Key* (ARK) identifier must be included. See our guide to best practices for citation. These requirements for citation also apply to data and/or derivatives of data originally sourced on MorphoSource that have been re-archived in third-party.

APPENDIX 5D
MorphoSource Standard Usage Agreement

One of nineteen usage agreement options available on MorphoSource.

MorphoSource Standard Usage Agreement

I agree to comply with each of the following basic requirements (1, 2, and 3) below whenever I download or access from MorphoSource any file(s), image(s), or model(s) (each, individually and all collectively, as applicable, referred to as "data") for use in any release, research paper, educational course, presentation, analysis project, website, blog post, social media post, or other public communication in the form of measurements, analyses, discussion, or 2D and 3D figures. I also agree that if I create digital or physical derivatives of the data (as distinguished from "data" or "original data"), I will comply with standards 1 and 2 immediately below, as well as the additional citation requirements and restrictions listed further below specifically referencing derivatives of data.

Basic Requirements

1. Any instructions for data citation and funding acknowledgment found on the relevant MorphoSource page for the files.

2. Any restrictions and requirements of any copyright licenses or usage agreements that may be attached to these files

3. Additional citation requirements and restrictions of use that are listed below.

Additional Citation Requirements and Restrictions of Use:

a. **Acknowledgement of Source**. Any release, research paper, educational course, presentation, analysis project, website, blog post, social media post, or other public communication referencing the data and/or derivatives of the data must clearly and explicitly identify MorphoSource, the collection organization (i.e., the specific collection, department, or laboratory), and the collection institution (i.e., the museum or university) as the original source of the data or derivative. This same acknowledgement must be made for data and/or derivatives of data originally sourced for MorphoSource that have been re-archived in third-party repositories (see below for any further restrictions on re-archival of this data or derivatives, if any exist).

b. **Citation**. Any release, research paper, educational course, presentation, analysis project, website, blog post, social media post, or other public communication referencing the data and/or digital or physical derivatives of the data will label, cite, or otherwise be captioned both with the institutional catalog number of the physical object or specimen that the data or derivative represents (preferably using the "Darwin Core triplet" of [Institution Code]:[Collection Code]:[Catalog Number]; for example, MCZ:Herp:A-12345), and the MorphoSource media identifier(s) of the downloaded file(s). Specifically, the MorphoSource *Digital Object Identifier* (DOI) should be included when available. If a DOI is not available for the file object in question, then a MorphoSource *Archival Resource Key* (ARK) identifier must be included. See our guide to best practices for citation. These requirements for citation also apply to data and/or derivatives of data originally sourced on MorphoSource that have been re-archived in third-party.

c. **Limits on informal or private sharing**. Informal or private sharing includes both any direct person-to-person transfer of files and the act of making files available to a select limited group of individuals whether using a digital system or by sharing physical objects or media. Informal or private sharing of data is allowed in **either** of the following two circumstances:

I. The third-party individuals with whom the media is shared do not intend to use this media for any additional sharing (from the third-party individual to further third-party individuals), and the third-party individuals with whom the media is shared do not intend to use this media for any purpose leading to a release, research paper, educational course, presentation, analysis project, website, blog post, social media post, or other public communication. **OR**

II. If **all** of the following conditions apply:

    i. MorphoSource data was initially downloaded or accessed with the explicit purpose of being used in a collaborative activity where limited sharing is a necessary component of the use case.

    ii. The download/access use case *can not* be achieved through multiple individuals separately downloading copies of the data through MorphoSource.

    iii. The download/access use case does not constitute public or unrestricted third-party re-distribution of the data and/or derivatives

Two examples of allowed informal or private sharing would be the use of 3D models or prints in an educational classroom setting for instructive purposes, or the collaborative creation of derivative 3D media from data by a team of subject experts. Other allowed usages do exist and are not limited to these examples.

Generally speaking, if you know of someone who wants to use or could benefit from the use of data you downloaded or accessed from MorphoSource, but you did not download the data for the purpose of a collaborative effort with that individual, then you should not share that data with that individual. Instead, that individual should download the data directly from MorphoSource.

In cases where limited informal sharing is permitted, you and each individual with whom you share data are bound to this agreement. This agreement does not supersede any copyright license, usage agreement, or other restrictions placed on the data by the data owner or collection organization. A copyright license or usage agreement may impose more or fewer restrictions on sharing compared to this provision, and users are responsible for following both these requirements and other licenses or agreements.

d. **Limits on redistribution through other repositories**. You will not share, re-archive, or deposit any original data downloaded or accessed from MorphoSource in a different publicly accessible digital repository without prior written permission from the data owner or collection organization, as applicable. In such cases, all requirements related to (a) **Acknowledgment of Source** and (b) **Citation** still apply.

e. **Commercial Usage of Data**. Data and derivatives created from data may not be resold or used as part of any activity or purpose that is primarily intended for commercial purposes or monetary gain without explicit approval from the relevant authorities of the collection organization, which might be subject to legal review by lawyers and/or licensing as per the collection organization's policy. Upon written request, MorphoSource staff can help with getting in touch with relevant curators and/or collections staff as necessary.

f. **Derivatives**. If a digital derivative is created from original data downloaded or accessed from MorphoSource **and** that derivative is used in a public-facing or otherwise-distributed product (including but not limited to a release, research paper, educational course, presentation, analysis project, website, blog post, social media post,, etc.), that digital derivative must be archived (but not necessarily openly shared) on MorphoSource within one year of the release of the product in which the derivative is used. Any usage of derivative data must adhere to the standards for (a) **Acknowledgment of Source** and (b) **Citation**. Other restrictions, such as forbidding archiving of derivatives in other publicly accessible digital repositories, may apply through other copyright licenses or usage agreements on original data, and users are responsible for following both these requirements and other licenses or agreements.

g. **3D Printing**. 3D printing of these data downloaded or accessed from MorphoSource is prohibited without prior written permission from the data owner or collection organization, as applicable.

I acknowledge that failure to adhere to these provisions could result in revocation of downloading privileges from MorphoSource, and could result in legal ramifications for me or my institution/company. I understand that I may still have to sign additional agreements deployed directly from the owner or collection organization before I am granted access to data. I also reaffirm my knowledge of and agreement to the MorphoSource terms and conditions generally (https://www.morphosource.org/About/termsAndConditions).

APPENDIX 5E
MorphoSource Restrictive Usage Agreement

MorphoSource Standard Usage Agreement

I agree to comply with each of the following basic requirements (1, 2, and 3) below whenever I download or access from MorphoSource any file(s), image(s), or model(s) (each, individually and all collectively, as applicable, referred to as "data") for use in any release, research paper, educational course, presentation, analysis project, website, blog post, social media post, or other public communication in the form of measurements, analyses, discussion, or 2D and 3D figures. I also agree that if I create digital or physical derivatives of the data (as distinguished from "data" or "original data"), I will comply with standards 1 and 2 immediately below, as well as the additional citation requirements and restrictions listed further below specifically referencing derivatives of data.

Basic Requirements

1. Any instructions for data citation and funding acknowledgment found on the relevant MorphoSource page for the files.

2. Any restrictions and requirements of any copyright licenses or usage agreements that may be attached to these files

3. Additional citation requirements and restrictions of use that are listed below.

Additional Citation Requirements and Restrictions of Use:

a. **Acknowledgement of Source**. Any release, research paper, educational course, presentation, analysis project, website, blog post, social media post, or other public communication referencing the data and/or derivatives of the data must clearly and explicitly identify MorphoSource, the collection organization (i.e., the specific collection, department, or laboratory), and the collection institution (i.e., the museum or university) as the original source of the data or derivative. This same acknowledgement must be made for data and/or derivatives of data originally sourced for MorphoSource that have been re-archived in third-party repositories (see below for any further restrictions on re-archival of this data or derivatives, if any exist).

b. **Citation**. Any release, research paper, educational course, presentation, analysis project, website, blog post, social media post, or other public communication referencing the data and/or digital or physical derivatives of the data will label, cite, or otherwise be captioned both with the institutional catalog number of the physical object or specimen that the data or derivative represents (preferably using the "Darwin Core triplet" of [Institution Code]:[Collection Code]:[Catalog Number]; for example, MCZ:Herp:A-12345), and the MorphoSource media identifier(s) of the downloaded file(s). Specifically, the MorphoSource *Digital Object Identifier* (DOI) should be included when available. If a DOI is not available for the file object in question, then a MorphoSource *Archival Resource Key* (ARK) identifier must be included. See our guide to best practices for citation. These requirements for citation also apply to data and/or derivatives of data originally sourced on MorphoSource that have been re-archived in third-party.

c. **Limits on informal or private sharing**. Informal or private sharing includes both any direct person-to-person transfer of files and the act of making files available to a select limited group of individuals whether using a digital system or by sharing physical objects or media. Informal or private sharing of data is allowed in **either** of the following two circumstances:

I. The third-party individuals with whom the media is shared do not intend to use this media for any additional sharing (from the third-party individual to further third-party individuals), and the third-party individuals with whom the media is shared do not intend to use this media for any purpose leading to a release, research paper, educational course, presentation, analysis project, website, blog post, social media post, or other public communication. **OR**

II. If **all** of the following conditions apply:

   i. MorphoSource data was initially downloaded or accessed with the explicit purpose of being used in a collaborative activity where limited sharing is a necessary component of the use case.

   ii. The download/access use case *can not* be achieved through multiple individuals separately downloading copies of the data through MorphoSource.

   iii. The download/access use case does not constitute public or unrestricted third-party re-distribution of the data and/or derivatives

Two examples of allowed informal or private sharing would be the use of 3D models or prints in an educational classroom setting for instructive purposes, or the collaborative creation of derivative 3D media from data by a team of subject experts. Other allowed usages do exist and are not limited to these examples.

Generally speaking, if you know of someone who wants to use or could benefit from the use of data you downloaded or accessed from MorphoSource, but you did not download the data for the purpose of a collaborative effort with that individual, then you should not share that data with that individual. Instead, that individual should download the data directly from MorphoSource.

In cases where limited informal sharing is permitted, you and each individual with whom you share data are bound to this agreement. This agreement does not supersede any copyright license, usage agreement, or other restrictions placed on the data by the data owner or collection organization. A copyright license or usage agreement may impose more or fewer restrictions on sharing compared to this provision, and users are responsible for following both these requirements and other licenses or agreements.

d. **Limits on redistribution through other repositories**. You will not share, re-archive, or deposit any original data downloaded or accessed from MorphoSource in a different publicly accessible digital repository without prior written permission from the data owner or collection organization, as applicable. In such cases, all requirements related to (a) **Acknowledgment of Source** and (b) **Citation** still apply.

e. **Commercial Usage of Data**. Data and derivatives created from data may not be resold or used as part of any activity or purpose that is primarily intended for commercial purposes or monetary gain without explicit approval from the relevant authorities of the collection organization, which might be subject to legal review by lawyers and/or licensing as per the collection organization's policy. Upon written request, MorphoSource staff can help with getting in touch with relevant curators and/or collections staff as necessary.

f. **Derivatives**. If a digital derivative is created from original data downloaded or accessed from MorphoSource **and** that derivative is used in a public-facing or otherwise-distributed product (including but not limited to a release, research paper, educational course, presentation, analysis project, website, blog post, social media post,, etc.), that digital derivative must be archived (but not necessarily openly shared) on a publicly accessible repository within one year of the release of the product in which the derivative is used. We encourage you to archive this derivative on MorphoSource, but this is not specifically required. Any usage of derivative data must adhere to the standards for (a) **Acknowledgment of Source** and (b) **Citation**. Other restrictions, such as forbidding archiving of derivatives in other publicly accessible digital repositories, may apply through other copyright licenses or usage agreements on original data, and users are responsible for following both these requirements and other licenses or agreements.

g. **3D Printing**. 3D prints of these data downloaded or accessed from MorphoSource are treated as derivatives of data, with all applicable non-digital provisions thereof, but other unique restrictions apply as well. You may create 3D prints of these data downloaded or accessed from MorphoSource only if the data was initially downloaded or accessed for a purpose that explicitly includes 3D printing. If you have downloaded media for another purpose, please re-download that media providing a 3D printing use case or communicate directly with the data owner or collection organization to obtain permission to 3D print the data. You may create a small number of 3D prints necessary for personal use or an allowed limited informal or private sharing use as described in (c) **Limits on informal or private sharing**. You may not redistribute 3D prints of data, unless this usage has been explicitly cleared with the data owner(s) or collection organization. Digital derivatives created during the process of preparing data for 3D scanning – e.g., the creation of a hollowed-out STL with 3D printing supports added – are exempted from the requirement to re-archive digital derivatives described in (f) **Derivatives**. We encourage you to share that derivative on MorphoSource so that others could benefit from this work, but you **are not** required to re-archive that derivative.

I acknowledge that failure to adhere to these provisions could result in revocation of downloading privileges from MorphoSource, and could result in legal ramifications for me or my institution/company. I understand that I may still have to sign additional agreements deployed directly from the owner or collection organization before I am granted access to data. I also reaffirm my knowledge of and agreement to the MorphoSource terms and conditions generally (https://www.morphosource.org/About/termsAndConditions).

APPENDIX 5F

University of Michigan Online Repository of Fossils Terms of Use

The Terms of Use for the University of Michigan Museum of Paleontology Online Repository of Fossils. (University of Michigan, Museum of Paleontology, Online Repository of Fossils, "Terms of Use," https://umorf.ummp.lsa.umich.edu/wp/terms/.)

About | 3D Invertebrates | 3D Vertebrates | Education | Login/Logout | Search

Terms of Use

By downloading, receiving, or otherwise accessing or using UMORF Data provided by the UMMP, you agree to be bound by these Terms of Use.

1. *Usage and Registration*

1.1 The University of Michigan Museum of Paleontology ("UMMP") has made certain digital fossil specimen data ("UMORF Data") available to you in the University of Michigan Online Repository of Fossils ("UMORF"), which can be found at umorf.ummp.lsa.umich.edu (the "Site"). The Regents of the University of Michigan ("Michigan") grants you a non-exclusive, revocable, **non-sublicensable, and non-transferable** license to use the "UMORF Data made available to you by the University of Michigan Museum of Paleontology ("UMMP") in the UMORF for **education and research purposes** only. For the avoidance of doubt, any actual or intended commercial use requires separate and prior written permission from the UMMP. Permission requests can be made by emailing Paleo-Museum@umich.edu or by calling (734)764-0489.

1.2. When you visit the Site, you have the option of creating an account. In order to create an account, you must provide your name, email address, and affiliation. We will handle your personal information in accordance with our Privacy Policy. If you create an account, you are responsible for the confidentiality of your log-in credentials, for keeping your information accurate, and for any activity resulting from the use of your log-in credentials on the Site. If you create an account, you may download UMORF Data and use it in accordance with these Terms. If you do not create an account, you may view but not download the UMORF Data in accordance with these Terms.

1.3 **Redistribution or sublicensing of the UMORF Data is prohibited**. Deposition of UMORF Data in repositories requires separate and prior written permission from the UMMP. Additionally, production of 3D prints or other physical replicas derived from the UMORF Data also requires separate and prior written permission from the UMMP except where permission to create 3D prints is explicitly granted to registered users in the metadata or README.txt files of a specific fossil specimen.

1 of 4 9/30/2020, 7:56 PM

1.4 You agree to provide the UMMP with copies of digital derivatives you produce from the UMORF Data if you create an account and are permitted to download the UMORF Data, including but not limited to measurements, renderings, segmented data, 3D meshes, rigged meshes, animations, and annotations, if requested. Michigan may use these digital derivatives for research and educational purposes.

1.5 You must comply with any additional restrictions (e.g., copyright or privacy rights) specified in metadata and/or README.txt files associated with the UMORF Data. You may not remove any proprietary notices included with or in the UMORF Data. You agree that you will not bypass, circumvent, or attempt to bypass or circumvent, any feature of the Site or any measures that we may use to prevent or restrict access to UMORF Data or the Site, including but not limited to downloading UMORF Data without registering for an account.

2. *Publication*

2.1 You shall notify the UMMP of all authored/co-authored publications in which UMORF Data are used upon acceptance to a publication and provide copies if requested.

3. *Attribution*

3.1 **The author and source of the UMORF Data must be cited in publications and other works** as follows: "UMORF data is provided courtesy of the University of Michigan Museum of Paleontology". Additionally, the catalog number and URL of each specimen represented in the dataset must also be included in the attribution (e.g., "specimen UMMP 117677, URL https://umorf.ummp.lsa.umich.edu/wp/specimen-data/?Model_ID=1306").

4. *Disclaimer of Warranty; Limitation of Liability*

4.1 UMORF Data and associated information are provided on an "as is" basis, and Michigan specifically disclaims any and all warranties or conditions of any kind, whether express or implied, including, without limitation, any warranties or conditions of title, non-infringement, merchantability, or fitness for a particular purpose. Additionally, Michigan has no obligation to provide you with any maintenance, support, or updated UMORF Data at any time.

4.2 IN NO EVENT WILL MICHIGAN, INCLUDING ITS REGENTS, FELLOWS, OFFICERS, EMPLOYEES AND AGENTS (the "Michigan Parties"), BE LIABLE UNDER OR IN CONNECTION WITH THIS AGREEMENT UNDER ANY LEGAL OR EQUITABLE THEORY, INCLUDING BREACH OF CONTRACT, TORT (INCLUDING NEGLIGENCE), STRICT LIABILITY, AND OTHERWISE, FOR ANY (a) DIRECT; (b) CONSEQUENTIAL, INCIDENTAL, INDIRECT, EXEMPLARY, SPECIAL, ENHANCED, OR PUNITIVE DAMAGES; (c) INCREASED COSTS, DIMINUTION IN VALUE, OR LOST BUSINESS, PRODUCTION, REVENUES, OR PROFITS; (d) LOSS OF GOODWILL OR REPUTATION; (e) USE, INABILITY TO USE, LOSS, INTERRUPTION, DELAY, OR RECOVERY OF ANY DATA OR BREACH OF DATA OR SYSTEM SECURITY; OR (f) COST OF REPLACEMENT GOODS OR SERVICES, IN EACH CASE REGARDLESS OF

WHETHER A MICHIGAN PARTY MAY HAVE BEEN ADVISED OF THE POSSIBILITY OF SUCH LOSSES OR DAMAGES OR SUCH LOSSES OR DAMAGES WERE OTHERWISE FORESEEABLE.

5. *Reservation of Rights*

5.1 Michigan reserves all rights not expressly granted to you in these Terms of Use.

5.2 These Terms of Use may change without notice. Changes shall be effective upon posting, and you are encouraged to check regularly for updates.

6. *Applicable Law*

6.1 These Terms of Use will be governed by and construed under the laws of the State of Michigan. Any actions in regard to these Terms of Use will be heard, decided, and resolved in the Michigan Court of Claims.

7. *No Assignment*

7.1 You may not assign or transfer any of your rights or delegate any of your obligations hereunder to any third party.

8. *Compliance with Law*

8.1 You may not use the UMORF Data in any manner or for any purpose that infringes, misappropriates, or otherwise violates any intellectual property right or other right of any person, or that violates any applicable law.

8.2 The rights granted to you in these Terms of Use are subject to any restrictions or any rights required to be granted under prior research or sponsorship agreements, or retained by the U.S. government, for example in accordance with Chapter 18 or Title 35 of U.S.C. 200-212 and the regulations thereunder (37 CFR Part 401), when applicable. You agree to comply in all respects, and shall provide us with all reasonably requested information and cooperation for us to comply with applicable provisions of the same and any requirements of any contracts between us and any agency of the U.S. government that provided funding for the subject matter hereof.

9. *Termination*

9.1 Michigan may terminate your right to use the UMORF Data at any time or. Upon termination, you agree to cease using the UMORF Data, delete or destroy all copies of the UMORF Data, and certify in writing to UMMP that the UMORF Data has been deleted or destroyed.

See Terms of Use and Privacy Policy

Our mission is to support, enhance, and collaborate in the research and instructional activities of the faculty, students, and staff, and contribute to the common good by collecting, organizing, preserving, and sharing the Paleontological record. This mission is accomplished through the establishment of the various collections and programs available through the Museum of Paleontology.

APPENDIX 5G
Resources

For further exploration, consider this very selective list of resources regarding items specifically discussed in the chapter.

Cases Discussed

Burrow-Giles Lithographic Co. v. Sarony, 111 U.S. 53 (1884) considers the camera as new technology and found that photographs are eligible for copyright protection.

Meshwerks, Inc. v. Toyota Motor Sales U.S.A, 528 F.3d 1258 (10th Cir. 2008)

Databases in the EU

European Commission. "Priority, Digital Single Market: Bringing Down Barriers to Unlock Online Opportunities." Accessed February 27, 2019. https://ec.europa.eu/commission/priorities/digital-single-market_en (page discontinued).

———. "Evaluation of Directive 96/9/EC on the Legal Protection of Databases." Commission staff working document. April 25, 2018. https://ec.europa.eu/digital-single-market/en/protection-databases.

Preserving the Public Domain

International Rights Statements Working Group. *White Paper: Recommendations for Standardized International Rights Statements*. RightsStatements.org, October 2015, last modified May 2018. https://rightsstatements.org/files/180531recommendations_for_standardized_international_rights_statements_v1.2.2.pdf. Discusses how photos of public domain works are typically not eligible for copyright and—even if eligible in some cases—should be treated by cultural institutions as in the public domain for reasons of public policy.

On the Freedom of Panorama

de Rosnay, Mélanie Dulong, and Pierre-Carl Langlais. "Public Artworks and the Freedom of Panorama Controversy: A Case of Wikimedia Influence." *Internet Policy Review* 6, no. 1 (February 16, 2017). https://doi.org/10.14763/2017.1.447.

Notes

1. Copyright Act of 1976, 17 U.S.C. § 102 (1990).
2. Feist Publications v. Rural Telephone Service Co., 499 U.S. 340 (1991).
3. *Feist*, 499 U.S. 340.
4. Rockford Map Publishers, Inc. v. Directory Serv. Co. of Colorado, 768 F.2d 145, 148 (7th Cir. 1985).
5. *Feist*, 499 U.S. 340.
6. *Feist*, 499 U.S. 363.
7. *Feist*, 499 U.S. 359.
8. *Feist*, 499 U.S. 364.
9. Bridgeman Art Library v. Corel, 36 F. Supp. 2d 191 (S.D.N.Y. 1999).
10. *Bridgeman*, 191.
11. US Copyright Office, *Compendium of U.S. Copyright Office Practices*, 3rd ed. (Washington, DC: US Copyright Office, 2021), § 313.2.
12. Copyright Act of 1976, 17 U.S.C. § 101 (1990).
13. For more information, see the Local Contexts website, https://localcontexts.org/.
14. Intergovernmental Committee on Intellectual Property and Genetic Resources, Traditional Knowledge and Folklore, "Statement by the Tualip Tribes of Washington on Folklore, Indigenous Knowledge, and the Public Domain," July 9, 2003, https://www.wipo.int/export/sites/www/tk/en/igc/ngo/tulaliptribes.pdf.
15. Copyright Act of 1976, 17 U.S.C. § 101 (1990).
16. US Copyright Office, *Compendium*, § 313.2.
17. Copyright Act of 1976, 17 U.S.C. § 107 (1990).
18. *Campbell v. Acuff-Rose Music, Inc.* 510 U.S. 569, 579 (1994).
19. The complete reconstruction with metadata is available at CISA—Interdepartmental Center for Archeology, "3D Model of Augusteum at Herculaneum with Quadrifrontal Arch," Europeana, 2013–2014, https://www.europeana.eu/portal/it/record/2048703/object_HA_1786.html?q=herculaneum#dcId=1554361549007&p=1.
   The quadrifons is available at CISA—Interdepartmental Center for Archeology, "3D Model of Quadrifrontal Arch of Augusteum at Herculaneum," Europeana, 2013–2014, https://www.europeana.eu/portal/it/record/2048703/object_HA_847.html?q=herculaneum#dcId=1554361549007&p=1.
   The CAD model is available at, CISA—Interdeparmental Center for Archaeology, "3D Model of Augusteum at Herculaneum," Europeana, 2013–2014, https://www.europeana.eu/portal/it/record/2048703/object_HA_1773.html?q=herculaneum#dcId=1554361549007&p=1.
   The four statues are available into Europeana under the 3D collection of the Museo Archeologico Nazionale di Napoli. The photos of the frescoes have kindly been provided by the Museo Archeologico Nazionale di Napoli. The high-resolution model can be downloaded on demand.
20. This work was published at A. D'Andrea, A. Bosco, and M. Barbarino, "A 3D Environment to Rebuild Virtually the So-Called Augusteum in Herculaneum," *Archeologia e Calcolatori* 28, no. 2 (2017): 437–46, https://doi.org/10.19282/AC.28.2.2017.35.

Bibliography

Andrea, A. D', A. Bosco, and M. Barbarino. "A 3D Environment to Rebuild Virtually the So-Called Augusteum in Herculaneum." *Archeologia e Calcolatori* 28, no. 2 (2017): 437–46. https://doi.org/10.19282/AC.28.2.2017.35.

Bridgeman Art Library v. Corel, 36 F. Supp. 2d 191 (S.D.N.Y. 1999).

CISA—Interdepartmental Center for Archaeology. "3D Model of Augusteum at Herculaneum." Europeana, 2013–2014. https://www.europeana.eu/portal/it/record/2048703/object_HA_1773.html?q=herculaneum#dcId=1554361549007&p=1.

CISA—Interdepartmental Center for Archeology. "3D Model of Augusteum at Herculaneum with Quadrifrontal Arch." Europeana, 2013–2014. https://www.europeana.eu/portal/it/record/2048703/object_HA_1786.html?q=herculaneum#dcId=1554361549007&p=1.

CISA—Interdepartmental Center for Archeology. "3D Model of Quadrifrontal Arch of Augusteum at Herculaneum." Europeana, 2013–2014. https://www.europeana.eu/portal/it/record/2048703/object_HA_847.html?q=herculaneum#dcId=1554361549007&p=1.

Copyright Act of 1976, 17 U.S.C § 102 (1990).

Digital Curation Centre home page. https://www.dcc.ac.uk/.

Feist Publications v. Rural Telephone Service Co., 499 U.S. 340 (1991).

Field Museum of Natural History. "Final Media Creation and Assignment Agreement." August 26, 2019. https://mm.fieldmuseum.org/b3198478-7efc-4896-b0ae-ca1a9431263f.

Intergovernmental Committee on Intellectual Property and Genetic Resources, Traditional Knowledge and Folklore. "Statement by the Tualip Tribes of Washington on Folklore, Indigenous Knowledge, and the Public Domain." July 9, 2003. https://www.wipo.int/export/sites/www/tk/en/igc/ngo/tulaliptribes.pdf.

Local Context website. https://localcontexts.org/.

MorphoSource 2.0 User Guide page. March 2, 2022. https://www.morphosource.org/docs/guide.

Research Data Alliance home page. https://www.rd-alliance.org/.

RightsStatements.org home page. https://rightsstatements.org.

Rockford Map Publishers, Inc. v. Directory Serv. Co. of Colorado, 768 F.2d 145 (7th Cir. 1985).

University of Michigan, Museum of Paleontology, Online Repository of Fossils. "Terms of Use." Accessed September 30, 2020. https://umorf.ummp.lsa.umich.edu/wp/terms/.

US Copyright Office. *Compendium of U.S. Copyright Office Practices*, 3rd ed. Washington, DC: US Copyright Office, 2021.

Chapter 6

Accessing 3D Data

Francesca Albrezzi, John Bonnett, Tassie Gniady, Heather Richards-Rissetto, and Lisa M. Snyder

ABSTRACT

The issue of access and discoverability is not simply a matter of permissions and availability. To identify, locate, retrieve, and reuse 3D materials requires consideration of a multiplicity of content types, as well as community and financial investment to resolve challenges related to ***usability****, interoperability, sustainability, and equity. This chapter will cover* ***modes****, audiences, assets and decision points, technology requirements, and limitations impacting access, as well as providing recommendations for next steps.*

Introduction

3D digital data **preservation** and access are complex and multilayered, involving many variables, including standards, best practices, open-source versus proprietary software, migration, and versioning. While 3D models hosted on platforms such as Sketchfab can sometimes (if downloadable with a compatible format) be reused for visualization, they are typically decimated models that are not acceptable for analysis. Other high-**resolution** models that can be downloaded are a step in the right direction; however, they typically require requested access (e.g., CyArk) or are difficult to download on lower bandwidths. As for scholarly reuse and peer review, most academics must not only be able to access 3D data, including raw, derived, and scene data, but they also need included critical **metadata** and **paradata**.[1]

When considering access to 3D materials, it helps to define what is meant by access. We are talking about the means to discover, examine, retrieve, or reuse 3D materials—because the issues of access are not simply a matter of permissions and availability. For example, to reference an audience use case that is discussed in depth later in the chapter, a researcher is attempting to study coral reefs, but some 3D models are **point clouds**, some are **meshes**, and some seem not to be scaled. For trusted reuse of data, the **provenance**, capture data, and apparatus surrounding the final product are essential in building a case using 3D materials from multiple sources. To do so requires consideration of (1) different modes of content type, (2) the needs of different audiences, (3) discoverability, (4) an understanding of technological requirements and limitations, (5) accessibility and inclusivity concerns, (6) the need for community and financial investment, and (7) citability guidelines.

In terms of 3D materials, following standards and best practices that promote interoperability is a viable strategy for ensuring long-term preservation and access of 3D content because they enable reuse of this material across any number of open-source or commercial software applications. Audience scenarios are used to explore the motivation for the use of 3D content by different audience categories. These scenarios emphasize the ways users discover 3D data and how each audience is using those data. In addition, the scenarios are used to evaluate technology needs and constraints as well as considerations surrounding accessibility and inclusivity for the widest set of audience categories. This chapter concludes with an examination of challenges to 3D data access, from discoverability, to use and reuse, to the creation of international standards, and to use of this work in promotion and tenure. Suggested next steps include possible repository solutions, methods to insure interoperability, uses of metadata in access, furthering accessibility and inclusivity, and evolving annotations, standards for peer review, and formats for citing 3D work.

Modes of 3D Data

While the bulk of this volume deals with scanned or photogrammetric reality-based models, 3D work can be made available in many forms, from simple CAD models to fully fledged virtual environments. Nine modes have been identified and examined in the appendix, and this chapter draws upon the diversity of modes when considering the matrix of issues surrounding access. Whether proprietary software impacts the interusability of data or virtual machines are necessary to fully resurrect a virtual environment, each presents challenges from creation to access to preservation. The appendix attempts to deal with the issues surrounding each phase so that creators, users, and archivists (among others) can make informed decisions.*

* Tables in the appendix describe each of these modes, detailing source material, methods associated with capture, hardware and software needed for capture and creation, output format, derivatives, methods of interaction, minimum files needed for access, and maximum files needed for preservation.

Audiences for 3D Data

Good discoverability and access practices begin with understanding the audiences that need to be served. When creating 3D material, institutions and organizations are likely to have established expectations regarding the end use of their content. While the end use of content may be known and is often a driving factor for the generation of 3D material, how the content might be reused in the future is often unknown or an afterthought. A natural history museum, for example, may digitize collections for its own preservation and management purposes, but it is also responsive to the communities it serves; thus, the needs of secondary users critically shape consideration of access and **discovery**. A project such as UMORF (University of Michigan Online Repository of Fossils) provides students, faculty, researchers, and a general audience with a collection of online 3D and 2D fossils from the University of Michigan Museum of Paleontology that can be examined within an online viewer. Designed with these audiences in mind, UMORF contains functionalities that allow users to spin the specimen, zoom in closely to see details and textures, toggle measurements on and off, and even see the object in anaglyph or interlaced 3D. Additionally, the hosting platform is rich with contextual metadata that supplement the visuals. These functionalities enable a wide range of information that can be useful to various audience types.

The following discussion explores how 3D material appeals to a multitude of audiences and defines the six main audience categories that are likely to need access to 3D data and related materials.[2] These categories may be porous but should help to identify what is at stake, of value, and important to users looking to employ 3D content. To further emphasize these perspectives, fictional scenarios are presented as exemplative use cases to demonstrate the particular wants and requirements for the specific audiences described.[3] These examples and scenarios presume noncommercial uses for the 3D content and a share-and-share-alike stance toward distribution. For-profit commercial and professional uses that might require licensing or use fees and rights and reproduction agreements are beyond the scale and scope of this chapter.

Audience Categories

Scholars and Researchers

The category of scholars and researchers refers to academics investigating 3D work or utilizing 3D methods for an evidence-based understanding of design and the development of new knowledge and learning opportunities for the public in formal and informal environments. They should have a knowledge of 3D methods that reinforces their use of 3D data for research.

Scenario: An art historian wants to interact with another scholar's sources-based reconstruction of an ancient site in order to test her own theory regarding the quality of light on a wall painting.

Scenario: A research team has scanned the underwater topography of a reef in Mexico and wants to combine it with similar datasets. They are equally happy working with point clouds or 3D meshes.

Educators

Educators in this context are defined as all instructors working with learners across all age and ability levels. These educators will likely have specific needs based on their students and use the 3D material to advance their own knowledge or incorporate into their pedagogy. Access will depend on whether the available 3D content meets their specific learning objectives or matches with their lesson plans and standards (state and federal). Additionally, access for educators is usually contingent on device availability.

Scenario: A high school teacher wants to teach a class on research methods by having her students reconstruct buildings from Victorian London using a free computer modeling program and so wants to locate existing academically generated models for a classroom discussion at the start of the semester.

Students

Depending on their age and abilities, students will have different sets of expectations and goals for seeking out and interacting with 3D materials. They may be interested in learning more about the content the model represents or the technical processes that are involved in producing 3D data and objects and what they can communicate regarding the physical objects or terrains themselves.

Scenario: Undergraduates in an American studies class are searching for 3D models of Native American pottery. Their assignment is to identify recurring decorative patterns and analyze them across cultures. If there is not a shared repository for such materials or connections among **archives**, this would require them to access multiple archives. The instructor may or may not provide them with links to known websites with Indigenous materials.

Museums, Public Outreach, and Nongovernmental Organizations (NGOs)

Institutions with missions to offer learning opportunities for the public via alternative environments may seek out 3D materials to supplement or support their existing programs and resources. Their motivation is to provide multiple pathways for broadening access to and engagement in learning experiences. For museums specifically, 3D materials enable display of resources that are warehoused due to lack of space and minimize handling of irreplaceable specimens. Virtual models also offer a way to present material that could not be displayed in a museum space, such as a reconstruction of a city's built landscape. Additionally, museums often turn to 3D models to facilitate user interaction and engagement with objects printed from 3D files. This is particularly important for museums that wish to serve those who are visually impaired. 3D printed objects can increase access and allow for haptic learning.

Scenario: A museum wants to 3D print bones from the skeleton of an endangered species as part of a hands-on installation for kids.

Professionals

Covering a wide range of expertise, the professionals category includes artists, architects, medical practitioners, engineers, game designers, animators, and more. These users avail themselves of 3D tools and content regularly within the scope of their work. The needs of the users encompassed within this category can vary greatly, which makes it a difficult category to address. Because this group potentially has commercial interests in the 3D content, they will be interested in intellectual property rights, licensing, technical specifications, and issues surrounding monetization (see chapter 5, "Copyright and Legal Issues Surrounding 3D Data").

Scenario: A mixed-media artist wants to build virtual experiences that incorporate scanned statues from museums across the United States to explore questions of scale and gender identity.

General User/Personal Interest

A general user would be described as anyone interested in material that is presented in a 3D format. The person could be any age or background with undetermined preknowledge. Their needs and expectations could vary widely, but they will likely be looking for a ready-to-use 3D experience that aligns with their personal interests and available technology. Intuitive features are very important to general users.

Scenario: A history enthusiast has just finished reading a book about the Gilded Age in the United States and wants to explore academically generated 3D environments that can immerse him in the time period. While he has some basic knowledge of the era and its architecture, it is critical that these 3D reconstructed environments be fully annotated in order to provide a general user an edifying experience.

Discovering 3D Assets and Decision-Making Issues

Discovery methods for 3D materials and related resources are crucial and also fragmented. At present there is no one unified way to find 3D assets, although some disciplinary silos have begun to occur and may provide a way forward for discovery depending on a user's needs. Currently, finding all 3D cultural heritage materials, even for a given location, can be difficult as different digitizers may have mounted their materials on different platforms. Similarly, libraries don't have a standard way of referencing these data, and it is often difficult to determine what a given institution's 3D holdings are.

Because needs can vary among the six audience categories based on their search experiences, parameters, and goals, we have articulated in table 6.1 the most common ways of finding 3D content and to illustrate the complexity of discovery.

TABLE 6.1

The most common ways of finding 3D content

Discovery Method	Explanation
Web search engines	A web search for "3D models" usually directs users to proprietary online repositories that are designed around consumer-based models; examples include sites like Sketchfab, TurboSquid, and CGtrader.
Online repositories	While proprietary online repositories can be found through web search engines, there are many libraries, archives, and museums that are working to create access for 3D materials online.
Searchable metadata	If 3D objects are shared with searchable metadata, audiences may be directed to them when a user searches for a specific type of object, location, title, creator, etc.
Word-of-mouth	While 3D is a growing community, many still hear about new content through personal channels.
Classroom exposure	Students often learn how to use 3D content and where to find relevant 3D materials during particular classes in their respective disciplines.
Professional training	Vocations like architecture or animation often require specialized training in particular 3D modeling techniques and software, and in the course of that training, students and professionals alike are directed to known caches of relevant 3D material.
Entertainment	The general public has exposure to 3D material through popular entertainment like 3D films, virtual reality, and video game play. This exposure may spark a search for 3D content that employs one or more of the above methods.

Once material has been discovered, certain conditions play a critical role in the use and reuse of 3D materials, and many of them inform the development of infrastructure and metadata schemata. Four out of the six audience categories (scholars and researchers; educators; students; museums, public outreach, and NGOs) will likely share common concerns regarding academic rigor of the project. However, professionals and general audiences may not find all of the concerns in table 6.2 to be of interest.

TABLE 6.2
Common concerns about 3D data.

Need	Explanation
Digital literacy	Provide 3D content in a way that is accessible to a given audience allowing users to successfully engage with and evaluate 3D content.
Ease/availability	The 3D material is discoverable and accessible, and the audience is able to reuse the content in a way that suits their goals for engaging with the material. Also, 3D content is provided in a way (such a web-based viewer and a smaller sized dataset) that requires the least specialized hardware and software.
Trust	It is readily apparent that the models are accurate and truthful, and there is readily available information about the construction or generation of the models.
Ethics	It is apparent that the 3D materials were generated or created with appropriate permissions and acknowledgment of intellectual property, considerations regarding the use of the material (e.g., immersion in educational settings), and providing visible credit and citation for work produced.
Consistency	The models include metadata fields that are generated in accordance with accepted community standards.
Utility	Use of the models is justified by a basic return-on-investment calculation (i.e., the personal time required to locate a model and learn the necessary technologies for use can be justified by the benefit of the engagement).
Interoperability	The models can move easily across platforms as desired, and critical metadata can be transferred from the 3D models.
Accessibility	The models include accommodations for differently abled users. At the moment there is very little offered in terms of virtual 3D materials that make for suitable accommodations for visually impaired users.[a] This is not just an academic concern. It should be a concern for any user.

a. For a more detailed discussion, see the document "Policies and Standards." https://www.hhs.gov/web/policies-and-standards/index.html

Additionally, two of the six main audience categories have specific concerns for decision-making issues regarding delivery systems for 3D materials. Educators will value classroom time, —available time in the classroom or within the lesson plan to integrate 3D materials—as well as the pedagogical return on investment for teaching and learning (i.e., Does the 3D material significantly outpace other forms of instructional technologies?). Professionals, on the other hand, have specific concerns that will vary across professions but could influence use or rejection of 3D materials. For instance, architects searching for 3D models to provide context for their own designs will have very specific requirements concerning rigor, dimensionality, and visual style.

The areas of focus listed in Table 6.2 should inform decisions made in terms of discovery and access. Most critical is that metadata developed for the 3D materials expose information to the users so that they can make informed decisions about the available content. At minimum, these metadata must include detailed information about the technical and academic pedigree of the material—information about the creation of the data and their reuse, the level of rigor and veracity used during their construction, and statements from the content creator about the project's objectives.

Technology Requirements and Limitations Impacting Access

Moving beyond **source material** and capture, it is important for this chapter to consider *how* the data will be accessed. This chapter privileges the creator's intended use so as to limit scope. However, the technology required for interaction needs to be examined as different audiences have access to different kinds of technology. When considering all six audience categories, modes of access can vary greatly based on hardware and internet access.

For example, while in the United States, about 75% of American households have broadband internet service,[4] in Mexico in 2018, only 44.9 percent of households have a personal computer, and only 13.26 percent have fixed broadband subscriptions.[5] Even smartphones increase market penetration only to 56%.[6] Public libraries and internet cafés provide ways to get online, but many will not allow specialized software or large file sizes to be downloaded. Thus viewers that allow access over the internet provide a distinct advantage when considering access for the broadest category, the general public. A virtual world or environment or a model that needs to be accessed in high resolution to evaluate its integrity can be a permanent barrier to entry (see table 6.3).

TABLE 6.3

Good/Better/Best recommendations for online 3D data file types

Tier	Recommendation	Examples
Good	Agnostic file type that is uploadable to a web viewer and loads relatively quickly or can be viewed with free or open-access desktop software	Final decimated model
Better	High-resolution files and access copy	High-resolution model available for download (when copyright allows) and final decimated model for web display
Best	Raw data, output files, high-resolution files, access copy	If the model was captured photogrammetrically, link to raw photographs or model as output from modeling software before cleaning, high-resolution model, final decimated model

If access is the predominant concern across audiences, however, then web-based viewers such as Sketchfab will aim to accept the differing formats of these models and allow additional annotation. While at this writing Sketchfab is the most popular commercial software available, it is important to note that models hosted on Sketchfab are subject to size limits,* and while users retain ownership over their content, Sketchfab is a hosting solution, not a repository. Others are working on viewers that would be self-hosted, such as the Smithsonian's Voyager or 3DHOP from the Visual Computing Laboratory at the Istituto di Scienza e Tecnologie dell'Informazione.[7] A self-hosted viewer removes constraints on the size of the model (although many models need a decimated or optimized version for distribution so as to make loading times reasonable or to meet hardware constraints) and allows more control over generated data. That said, desktop applications such as MeshLab or CloudCompare for scanned and photography-based models are necessary if high-resolution versions of a model are available for detailed inspection, measurement, or comparison.[8] This is also where repositories such as MorphoSource come in, as they often provide high-quality models for download and inspection.[9]

Use Case

Researchers at Indiana University investigated how differences in capture were reflected in resulting photogrammetric models.[10] They used several models from the Stanford 3D Repository that were scanned at high resolution and synthetically photographed the models in Blender before processing the results in PhotoScan to recreate the models.[11] These results were loaded into CloudCompare to determine best practices for capture and investigate tolerances, as seen in figure 6.1. This methodology could also be used to compare captures by different entities where scientific tolerances are important or artistic integrity of the object is paramount.†

---

* Sketchfab is currently prototyping streaming for "massive" models, but the feature has not been rolled out at the time of this writing. (Bart Veldhuizen, "Stream Massive 3D Models, Now with Texture Support," *Sketchfab Blog*, July 31, 2019, https://sketchfab.com/blogs/community/stream-massive-models-now-with-texture-support.)

† For example, there are two reconstructed versions of the Palmyran Arch of Triumph blown up by ISIS in 2015, one 3D printed by the Institute for Digital Archaeology (IDA) and an online model by The Arc/k Project (Arc/K). However, the online version is not downloadable, and the printed arch is patented with limited accessibility, both by the people whose cultural losses are meant to be represented and by those who would be educated in that loss. (Roshni Khunti, "The Problem with Printing Palmyra: Exploring the Ethics of Using 3D Printing Technology to Reconstruct Heritage," *Studies in Digital Heritage* 2, no. 1 [2018]: 1–12, https://doi.org/10.14434/sdh.v2i1.24590.)

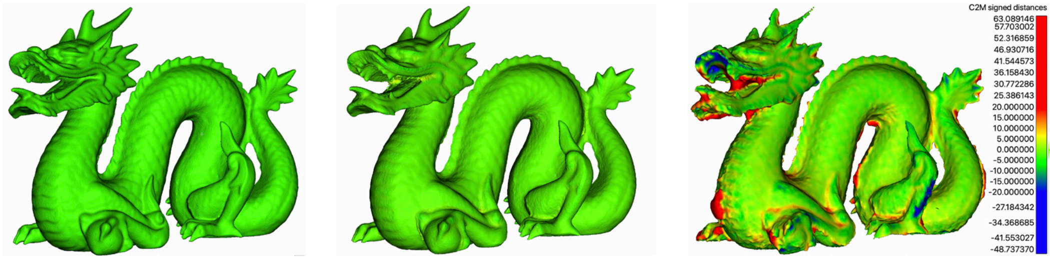


Figure 6.1
Using desktop applications to interrogate 3D models

For immersive worlds and virtual environments, there is no standardized access format at the time of this writing. The **emulation** strategies employed by the Internet Archive and championed by organizations such as the Video Game History Foundation are preserving the user experience of many classic computer games, but one-off academic projects and virtual models and environments are in danger of disappearing from the scholarly record.

Challenges and Outstanding Questions

To facilitate and foster 3D data reuse, we must take into account the considerations mentioned in the appendix; however, four areas rise to the surface as the most critical: discoverability, interoperability, citability, and peer review. The following sections demonstrate how the ability to locate, use, evaluate, and reference 3D materials affects the audiences and technology requirements listed earlier in this chapter. Adding to the complexity of the topic of access, these four factors are interdependent. As a result, consideration of any one factor requires consideration of the others. The essential work being done within these areas will establish consensus regarding practitioner and archivist **workflows** and infrastructures for preservation and exchange.

Discoverability

At the moment, there is no central repository for all 3D scholarship, even for the modes that can be gathered together. For example, DigiMorph.org (University of Texas) went live in 2002 to serve visualizations derived from high-resolution **X-ray computed tomography** (HRXCT); it was not, however, designed to serve the HRXCT data themselves. Since then, Duke University's MorphoSource has made inroads, as it is designed for **volumetric data** of biological and paleontological specimens with downloadable files ranging in format from the raw data (e.g., TIFF HRXCT slice stacks) to derivatives (e.g., .stl, .obj, .ply). This approach speaks to the possibility of separating 3D work by discipline rather than method of production. Cultural heritage work, for example, could be deposited into central repositories that would ideally allow options for both viewing and downloads.[12] 3D ICONS—a Europeana project focused on cultural heritage—does include appropriate metadata, but models are often not downloadable or cannot be viewed in an interactive 3D web display.[13] The integration of a viewer would be more important here, as rights restrictions come into play often with cultural objects, meaning that the raw dataset would never be available for download if intellectual property rights are not released. That said, further challenges occur when one moves beyond models to virtual worlds, environments, and games, each of which may necessitate specialty software for playback. These modes and their need for software (some of which might be proprietary) mean that not all end users may be able to access every piece of 3D work in a repository (if on a public library computer where software cannot be downloaded, for example).[14] However, if there were central repositories for 3D data, at least similarly cataloged work could be found—pointing to the need for robust and standardized metadata. Such repositories need not actually host all the work if rights management or scope of storage and management becomes an issue. Rather they could be **aggregators** with **persistent identifiers** linking out to work hosted elsewhere.

Some university libraries have begun hosting 3D content, but their metadata and cataloging strategies are not consistent. Some libraries refer to the method of digitization, and some simply call their material by the type of holding (e.g., OBJ, PLY). As a result, the major metadata aggregators such as WorldCat would need to incorporate additional parameters to successfully return comprehensive results. The first step toward finding available 3D content lies in standardized metadata surrounding each mode of 3D content.

In chapter 4, "Metadata Requirements for 3D Data," attention is given to how standardization of language plays into discoverability by non-3D practitioners, such as librarians who may be assisting patrons. One must also consider the role of verifiable provenance and tools for recreation (when rights permit) as well as evolving citation standards. In addition, each dataset requires a globally unique identifier (GUID), digital object identifier (DOI), or Archival Resource Key (ARK) if it is going to be findable by a catalog like WorldCat. Preferably, the identifier should be a globally unique persistent and resolvable identifier (GUPRI). Chapter 4 also rightly points out that a physical specimen may need multiple identifiers if different derivative or digitally constructed versions are available.

Interoperability

Interoperability for 3D data remains a major challenge. Numerous file formats exist for proprietary software that often are not interchangeable—it is difficult to achieve interoperability without a legal framework surrounding both licensing and open-access data. Data types such as point clouds and meshes are based on ASCII, binary, or both. While ASCII is recommended for long-term archiving and is essentially interoperable, it does not necessarily facilitate access and reuse because it stores minimal metadata and lacks paradata. While no standards exist, common file formats for 3D models include OBJ, PLY, DAE, and STL, which can be used in many software packages, and this is where API converters could help bridge the gap between different file types. However, each of these formats has pros and cons, and file conversion for interoperability can change the initial raw data.

Virtual environments tend to be more complex because they are typically proprietary and often originate from multiple datasets, and the viewers required to interact with them include additional elements such as lighting, sound, and collision detection. Additionally, game engines such as Unity3D and Unreal Engine have numerous versions that are not backward compatible; that is, files created with newer versions cannot be opened in earlier versions. While many (but not all) older projects can be opened in newer versions, incompatibility between versions still exists requiring editing code to ensure original visualization and functionality. Based on open web standards, WebGL, along with 3D libraries and APIs such as three.js, provides an alternative for 3D visualization; few current software packages, however, are based on these open standards because they require intense coding by experts as well as consistent updates. Thus, while 3D visualization options exist (both proprietary and open-source) that are in theory interoperable, for example, they often do not have cross-compatible file formats, and the difficulties associated with migration and versioning are also often a roadblock. 3D analysis is a greater challenge because to carry out scholarly research requires having access to metadata and paradata. Additionally, it is critical that researchers have access to original 3D data (not simply derived models) to facilitate interoperability with other software as a single 3D visualization software is often insufficient for analytical purposes. As for CAD data, they are particularly challenging because not only are there numerous file types (extensions), but there are also many CAD-software-using native formats that are not interoperable.

3D data lack official standards, and this lack lessens their interoperability. While best practices and standards are slowly are emerging,[15] because of obsolete and diverse formats, versioning, specialized technologies, and rapid development of new software, there is no consensus on standards,[16] and it can be difficult to adhere even to recommendations.[17] For example, while OGC (Open Geospatial Consortium) highlights standards for some 3D formats, such as LAS, CityGML, and I3S, most of the commonly used formats, such as PLY, OBJ, and DAE, are not included. A key challenge for geospatial 3D data is that many 3D file formats cannot store or work with real-world coordinate information; thus, data integration is difficult. For example, DEMs and shapefiles lose their real-world reference in 3D gaming engines, which makes it difficult to easily ingest other georeferenced 3D models to create virtual worlds. To move forward with standards and best practices requires that communal work take place across disciplines and organizations to develop a set of 3D data standards that promote data exchange and interoperability for now and into the future.

Another major challenge for interoperability are the differences in 3D point clouds (acquired from **laser scanning**) versus 3D meshes comprised of faces (generated from 3D point clouds). While conversion from point clouds to meshes is commonplace, it is essential to realize the potential data loss and **transformation** from such conversion on the raw data. In other words, decisions are already being made that alter the data based on use purposes—public dissemination and research have different requirements.

Even if 3D data are interoperable and reusable for visualization, they are rarely reusable for scholarly purposes. While reality-based 3D models derived from **photogrammetry** and laser scanning can theoretically be reused, available models are typically decimated for web-based visualization. The decimation process sacrifices elements of the original data (e.g., geometry or **texture maps**) in exchange for viewing efficiency, resulting in data loss and potentially limiting the models' usefulness for secondary analysis because the optimized models no longer contain the original data that made them desirable to researchers in the first place. Models that are not optimized, however, are often too heavy for many computers to visualize or run computational analyses requiring both large amounts of RAM (memory) and processing power as well as expensive, powerful video cards (depending on 3D model mode). Reuse of 3D models and virtual environments created using 3D Studio Max, Maya, or AutoCAD, for example, is more complex, not simply because of proprietary formats, but also because of associated metadata and documentation. 3D reconstructions typically comprise multiple data sources such as GIS data, architectural drawings, photographs, field notes, and so on, and it is essential to know the data sources and modeling decisions (parameters) made in the reconstruction for scholarly reuse and peer review.[18] Similarly, 3D models and reconstructions that are repurposed mostly are not cited despite the scholarly work that goes into creating them.

Accessibility and Inclusivity

Ideally, content creators would consider different audiences and delivery platforms as they develop their 3D work and tools so that the materials serve the widest possible array of audiences. For example, decimated versions of models can be made available in web delivery players with links to higher resolution models linked in the metadata. Providing two versions of the material enables both web interaction and more detailed and stable offline exploration. It is also worth considering inclusivity when discussing access to 3D data. At the time of this writing, while the World Wide Web Consortium (W3C) promotes its Web Content Accessibility Guidelines (WCAG), recommendations for making 3D material inclusive are just beginning to be discussed.[19] The issues of inclusion can be wide-ranging, assessing both access to and the quality of software, hardware, and internet connectivity. Inclusive practices also address concerns regarding differences in digital literacy and skill sets, economic situation, education, geographic location, language, age, and disability. For online content and digital tools, accessibility-compliant materials would allow users with disabilities to "perceive, understand, navigate, and interact" with websites and tools so that they "can contribute equally without barriers."[20]

A project that could inform this inclusivity discussion is the Project Gap Analysis Rubric developed by Jasmine Clark to help practitioners assess the extent to which a digital project is accessible, usable, and inclusive. Through seven layers of criteria, practitioners rate a total of twenty-one elements as Weak (1), Average (2), or Strong (3) and tally their results. Including specific and detailed project information will yield a gap analysis that will be both concrete and actionable. The rubric elements combine well with something like Francesca Albrezzi's XR Implementation Checklist as a way to think about accessibility, usability, and inclusion within the early stages of a project. Clark stresses that even if practitioners do not have the time or resources to accomplish everything within the rubric, considering such matters is a substantial step.[21]

In terms of digital publication platforms, the University of Michigan Press/Michigan Publishing's e-book platform Fulcrum acknowledges that accessibility is a core value of its user experience design, adhering to the latest WCAG Level 2 AA Standards and providing users information about known web accessibility issues.[22] Fulcrum was used to publish 3D content with its 2016 release of *A Mid-Republican House from Gabii*.[23] Again, Jasmine Clark has helped pave new ground for digital publication in terms of accessibility and inclusion by creating a VR Accessibility Resource Sheet and a Web Accessibility Primer.[24] These resources serve to better educate and assist web designers, students, librarians, and scholars on how to make their immersive technology endeavors meet current standards and to help differentiate between web accessibility, usability, and inclusivity.

Annotation of 3D Research

Increasingly, academics assert that 3D models—particularly reconstruction models—can and should be seen as rigorous scholarly arguments in and of themselves.[25] Enabling that transformation from dataset to scholarly object requires the ability to associate the 3D models with supplemental information beyond the basic metadata required for discoverability. This supplemental information could be textual (e.g., spatially aware "footnotes" in 3D space that provide context, references to source material, paratext, explanations about interpretive decisions), expressed by the model itself (e.g., visual elements that signal areas of uncertainty, strategies for representing gaps in the available evidence, multiple reconstruction alternatives), or overlaid on the 3D model as a linear argument akin to the Tour feature built into the Sketchfab and Smithsonian 3D viewers or the Narrative feature in VSim.

When considering the infrastructure necessary to access and use 3D data, it is vital to consider the characteristics of the data. Depending on the domain of practice, the 3D data could represent a simple object, a large-scale virtual environment, or a complex spatiotemporal object that changes its morphology and its surface appearance in response to user interaction or changes in virtual world time. For example, several projects have proposed using photogrammetric models to monitor coral growth and die-off around the world, and the ability to compare models of the coral reefs generated over time will be critical to these efforts.[26] Similarly, Bernard Frischer uses computer-simulated shadows over time in his article on the Montecitorio Obelisk and the Ara Pacis to reveal "over 230 hitherto unrecognized solar and shadow alignments" to "create a recurrent sun and shadow spectacle that would have impressed the ancient viewer with [Augustus]'s learning, power, and religious commitment."[27]

Further confounding the requirements for reuse and preservation, 3D models themselves can also be considered as objects of study. Janet Delve describes them as complex and multimodal objects that are internally differentiated, hierarchical, and heterogeneous.[*] In this reframing, the model, as a scholarly object, is not merely the finished product as defined by the content creator, but an array of model iterations that illustrate its development over time. Thus, tools developed to support reuse and preservation must supports three functions:

- the display of the model's changing morphology over time,
- the display of the final and preceding versions of the model, and
- the display of the surface appearance of the model so that it symbolically represents one or more ontologies of data.

With respect to the last function, the 3D model plays a role akin to the 2D polygons used in geographic information systems (GIS). It can change its surface appearance from photorealistic to a symbolic color to show things such as the ethnicity of a given building's inhabitants, the provenance of a given tool's manufacture, or the reliability of a given building component's reconstruction.[28] Given this radically different conceptualization of 3D models as knowledge constructs, it will be crucial to devise new expressive, attestive, and workflow conventions that support the critical apparatus

* Janet Delve explains, "An essential first step when considering the nature of complex digital objects is to recognize that there are multiple layers of difficulty encountered when attempting to analyse them. These layers could be superficially likened to Georg Cantor's 'levels of infinity' in terms of mapping out the size of the problem space to be analysed. The first 'level of infinity' is that of detail: the problem of drilling down through many layers of technical elements, showing levels of interconnectedness both within digital objects themselves, and also with their technical environments." (Janet Delve, "Introduction to POCOS E-book 1: Preserving Visualisations and Simulations," in The Preservation of Complex Objects, ed. Janet Delve, David Anderson, Milena Dobreva, Drew Baker, Clive Billenness, and Leo Konstantelos [Portsmouth, UK: University of Portsmouth, 2012]: 10–11.)

surrounding digital scholarship and provide for citability of models and environments in different states.*

Citability and Peer Review

As mentioned above, citability and peer review are critical to encouraging scholars to make 3D data accessible and to reuse 3D data for academic scholarship. To enable the citation of 3D data by secondary scholars requires developing standards and best practices for referencing 3D scholarship. Because 3D data encompass geometry, metadata (publication and bibliographic), and paradata, citation is not straightforward. The use of open standards for file types and best practices for exporting and importing 3D data from multiple platforms can facilitate and foster broader publication and thus expand opportunities for discovering, using, and citing 3D material. Exporting 3D models and scenes using (still emerging) standards would allow them to more easily be reused as originally intended (without modification) for visualization and explanatory purposes in scholarly arguments because users could employ open-source, rather than only proprietary software to interact with and peer-review both models and arguments made with the models.

One way to approach the challenge of data structure is to develop and publish workflows or develop tools, such as the Digital Lab Notebook by Cultural Heritage Imaging (CHI), that provide guidance and easy-to-implement tools for documenting models using ISO-standard-compliant metadata to standardize and package geometry and metadata.[29] Providing workflows that offer step-by-step guidelines of best practices is a critical step toward creating citable 3D models that can be peer-reviewed and reused for new scholarly research. However, because of a lack of standards and infrastructure, it is still a challenge to carry original model attribution and metadata across many generations of derived models, thus inhibiting citation even when original models are properly cited.[30]

Moreover, as discussed above, web-accessible citation formats—using machine-readable formats that are "fixed to a specific file or bundle of files over the lifetime of those objects, even if their location on the internet changes"[31]—need to be employed for 3D models to allow them to be discovered and cited. For example, the Virtual Hampson Museum, with a specific focus on reusing and repurposing 3D models, hosts 3D objects. Originally each model had a URI and was downloadable.[32] However, because of

* Building information modeling (BIM) allows professionals to annotate and track data within virtual structures during their development and throughout their life cycle. By adding new metadata fields, scholars can adapt the schema for research documentation purposes. However, some, like Susan Schreibman and Costas Papadopoulos, are considering the efforts that are needed to produce a digital scholarly edition with 3D content. (Costas Papadopoulos and Susan Schreibman, "Towards 3D Scholarly Editions: The Battle of Mount Street Bridge," *Digital Humanities Quarterly* 13, no.1 [2019], http://www.digitalhumanities.org/dhq/vol/13/1/000415/000415.html.)

misuse of Creative Commons licensing, the models are now available only via 3DHOP and no longer downloadable, thus introducing another roadblock to reuse.

As for infrastructure, a few options exist for 3D publication. Journals such as *Digital Applications in Archaeology and Cultural Heritage* allow simple, low-resolution (typically decimated) models, or *Studies in Digital Heritage*, interactive 3D scenes (using Unity 3D–based platforms) to be included with traditional text. Recently, publishers have been experimenting with digital books with interactive 3D models linked to a database to allow data queries of model attributes (descriptive data) that form part of the scholarly argument. VSim, while developed for pedagogical purposes, also offers a way to reuse existing models to construct scholarly arguments that could potentially be used for peer review; however, as a desktop application, it must be downloaded for use.

Recommendations for Next Steps

Given these critical challenges, the following recommendations provide actionable interventions for the 3D community. In response to the previous challenges, these areas have been identified as opportunities for strategic development to improve the quality of access for 3D material.

Develop Repository Solutions for 3D Materials

While there are notable commercial repositories, the 3D community is primed for an aggregated repository or portal that would allow many 3D archives to be searched at once. An undertaking such as this would likely require the founding of a consortium, which would act in accordance and collaboration with others like IIIF, W3C, International Internet Preservation Consortium, and the Software Preservation Group. More than one such repository may need to be established in order to address particularities regarding content type and discipline-specific needs. To support these efforts, strong standards would need to be formed around linked open data (LOD) to take full advantage of 3D initiatives. Additionally, this work should establish a mechanism to include 3D materials in WorldCat and similar systems so added 3D objects have an appropriate level of inclusion in search aggregators but do not overwhelm the user or the platform.

The standard 3D metadata schema should also be expanded to include optional fields that enable Geoweb (i.e., finding all assets for a region simply by drawing a box and having 3D assets come up).[33] At a minimum, this would require latitude, longitude, and altitude for 3D assets. For even greater searching, the element of time could be added with date ranges. While Geoweb 3D work of this nature raises issues that would need to be addressed with sensitive site locations, the workflow could allow 3D data to cross disciplines and solve many problems of discovery that currently exist. In this case, the metadata would use geographic location as a standard feature.

Enhance Interoperability

While the hardware and software involved in producing a 3D model may be specific, interoperability standards could allow different 3D data types to be shared across various viewers. The IIIF 3D community group is currently assessing features of available viewers to identify common requirements and map the landscape of available options for cultural heritage content. Goals are

- To explore possibilities for viewing, search, discovery, and annotating 3D data.
- To collect and document use cases from existing and new IIIF community members that suggest the need for interoperability of 3D data.
- To collect, discuss, and evaluate the state of the art with respect to 3D requirements for use by the cultural heritage community on the web.
- To coordinate and connect through outreach to internal and external partners, technical experts, and related initiatives.
- To explore best practices for interoperability and possibilities with existing IIIF specifications and open APIs through articulating use cases and experimentation.[34]

Employ Standard Metadata and Cataloging Schemata

Agreed-upon metadata standards should be employed wherever possible, and the table of best practices in chapter 4 of this book, "Metadata Requirements for 3D Data," should be employed, as well as looking toward the use of RDF and OWL as described below in Table 6.4.

Design for Accessibility and Inclusivity

For a 3D access environment or platform to be inclusive, access needs to be equitable, addressing a person's requirements until their experience aligns with the standard that is set for all. Adopting universal design principles and building multimodal systems can help increase a project's usability and inclusiveness. Designing to reach the greatest audience possible avoids the need for adaptations, which can cause users to feel excluded or singled out.[35]

Employ Robust Annotations

Annotations, whether existing in the viewing environment or as supplementary material, need to be robust enough to meet the needs of the highest level user envisioned. The elements below may be considered as a starting point for this discussion:

> An annotation system designed to address the needs of academics working with 3D content must address five layers of information relative to the modeled environment: the **source material** used by the content creator to inform the reconstruction, **introductory information** to explicate the environment for users, **paradata** documenting the processes used during its creation [and interpretive decisions], **academic argumentation**, and **paratextual** information created by peer reviewers, editors, or secondary users.[36]

When dealing exclusively with in-environment annotations, Papadopoulos and Schreibman write that in-environment annotations are meant to explain and contextualize and offer scholarly scaffolding:

> For example, the 3D (re)constructions may offer one version of a building; however, evidence that supports alternative versions of certain architectural features may be represented by other models accessible in-world through a pop-up box or by replacing the current version of a feature with other possible versions; areas of uncertainty may be rendered in different colours and shading to indicate hypotheses, sources, and surviving evidence; or, ambiguous features may be toggled on and off or replaced by alternative versions, also indicating how other elements will be affected by these changes (e.g. a larger door opening may indicate a lighter roof structure).[37]

Moving forward, the 3D community will need to consider ways that scholarly annotation can be standardized to increase interoperability across platforms and allow for greater publication opportunities. Concerns regarding issues of version control and editing pipelines for annotated 3D materials need to be addressed. If agreement about managing annotation workflows can be met, publishers will be furthered empowered to take on 3D projects.

Set Standards for Peer Review

Peer review for 3D scholarship could be modeled on past projects that were created explicitly to review digital scholarship such as NINES (Networked Infrastructure for Nineteenth-Century Electronic Scholarship), and its sister sites 18thConnect, MESA, ReKN, and ModNets.[38] However, these projects, which were once robustly active, have lapsed into silence—from either lack of funding, lack of staffing, or both. They still serve as a snapshot of best practices for a window in time in the realm of electronic scholarship, but without new accessions, any cutting-edge work being done is not represented.

The danger of orphaned, well-intentioned, and even successful projects looms large in the digital realm, and, for the sake of promotion and tenure, a new construct for peer review in the 3D realm should look seriously at sustainability. In lieu of a formal evaluating body, the 3D community could follow Geoffrey Rockwell's "Short Guide to Evaluation of Digital Work"; the more recently penned *Guidelines for the Professional Evaluation of Digital Scholarship by Historians* put out by the American Historical Association; the "Guidelines for Evaluating Digital Scholarship" in the Society for American Archaeology's *Report of the SAA Task Force on Guidelines for Promotion and Tenure for Archaeologists in Diverse Academic Roles*, which includes recommendations for evaluating 3D modeling and VR scholarship; and the College Art Association and the Society of Architectural Historians' *Guidelines for the Evaluation of Digital Scholarship in Art and Architectural History* .[39]

Agree Upon Format for Citing 3D Data

It is crucial that scholars, information specialists, commercial and construction practitioners, and other users of 3D (or 4D, if one includes a time element) content begin to formalize and establish their respective documentation practices. While scholars have been painfully aware of the need for proper attestation practices since the 1990s, and that realization has given rise to initiatives such as the Cultural Virtual Reality Organisation (CVRO) and the London Charter, no initiative has led to the articulation of a concrete set of practices that we might find affiliated with the International Organization for Standardization (ISO) or articulated as a 3D equivalent to the *Chicago Manual of Style*.[40] A good deal of discussion has emerged from literatures ranging from virtual heritage to digital construction, historical GIS, the digital humanities, and other fields articulating documentation requirements, and based on those writings we recommend the development of 3D citation practices that meet the requirements of 3D scholarship (see table 6.4).

TABLE 6.4

Good/Better/Best recommendations for 3D data citation practices

Tier	Recommendation	Examples
Good	Citations contain three components: publication data, bibliographic data, and paradata.	Publication data should communicate the name of the dataset, the identities of its creators, the name of its publishers or host institutions, the object's metadata (i.e., keywords), **its location and its publication and copyright status (e.g., proprietary versus open source)**. Bibliographic data should identify the name and provenance of all data sources that gave rise to the model. Paradata is a concept that has generated interest in multiple domains and with it multiple definitions, but for our purposes its definition can be reduced to the following: paradata communicate the decision-making and methodologies that gave rise to the model. See table 4.1 in chapter 4, "Metadata Requirements for 3D Data."

Tier	Recommendation	Examples
Better	Multimodal citations	Discussion and examples of 3D model documentation have typically centered on metadata and the use of software with text-entry fields to attach inscribed metadata to the given model. However, as mentioned earlier in the chapter, CHI has a Digital Lab Notebook that can provide a pathway from creation to publication. Scholars seeking to describe the workflow underlying a given model, for example, might also take a page from the construction industry and use schematic diagrams to describe the workflow and indicate important decision points associated with the data's construction (see figure 6.2). Other scholars, wishing to follow the interpretive reasoning behind a model, might wish to see prior versions of that model expressed in 3D.
Best	3D models that are extensible and semantic	Further structuring data with Semantic Web technologies such as RDF (Resource Description Framework) and OWL (Web Ontology Language) support the rapid aggregation from multiple sources of data relating to a given domain or topic.

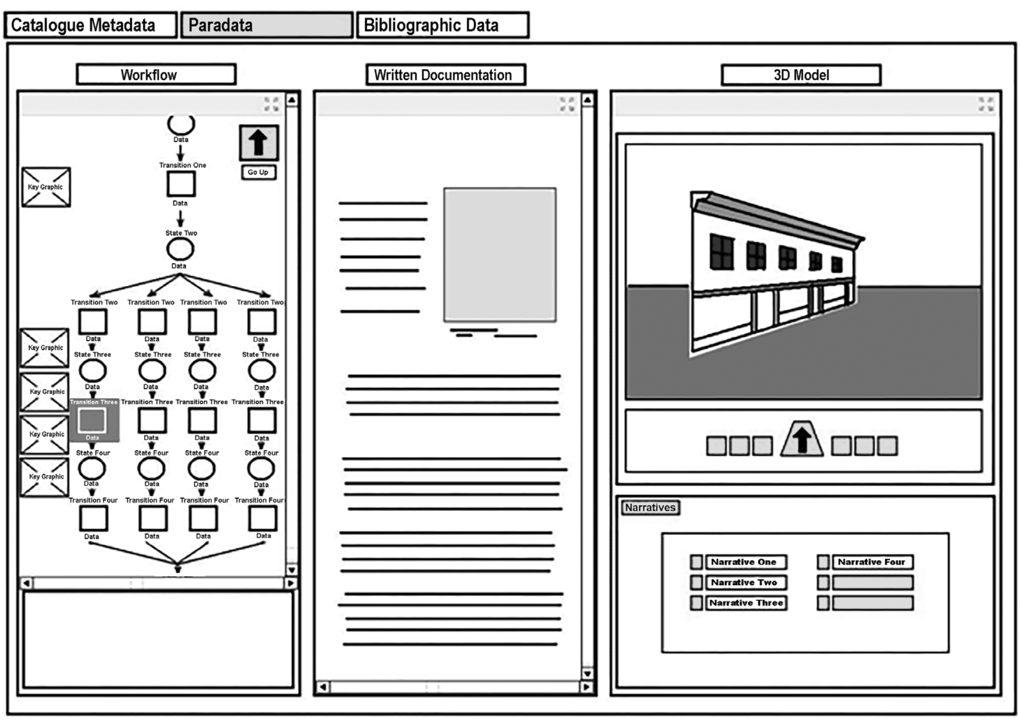


Figure 6.2
Mockup of multimodal expression of paradata

Conclusion

Discovery and access are essential for the dissemination and use of 3D material. By defining the various modes of 3D data, this chapter has touched on a myriad of production methods and file outputs that frame conversations about how people store and share 3D materials. This discussion is critical for the construction of useful discovery tools and interoperability standards as the field seeks long-term sustainable workflows. Standards will need to not only address crosswalking metadata schemata generated in connection to methods of production, but also find a solution for viewing 3D material across platforms. The issues around viewing platforms also highlight the need to formalize how supplemental materials in the forms of annotation, embedded resources, tour features, and the like are integrated with the computer model and translated to other platforms as content is moved and preserved.

At the heart of this discussion are the audiences producing, using, and reusing 3D data. Reflecting on user needs can assist the 3D community in developing technical requirements and identifying implementation or use limitations. Future work for 3D practitioners needs to address gaps around accessibility, whether by those without access to high-speed internet connections or by the disabled. Additionally, in terms of scholarship and publication, if a community goal is for the model to become the site of academic argument, issues regarding citation and peer review will need further attention. Citation for 3D material must include far more than 3D coordinates (*x*, *y*, *z*, *h*, *p*, *r*), but address geometry, metadata (publication and bibliographic), and paradata. It is recommended that peer review of digital work have a solid framework for evaluation that is posited on access and annotation of the life cycle of 3D materials. Discovery and access become possible with the careful integration of the lessons and best practices communicated in the other chapters of this book and with a focus and dedication to the audiences that make up the 3D ecosystem.

APPENDIX 6A
Information needed to support long-term access to different modes 3D Data

TABLE 6.A.1
Modality: Manual

Information	Examples
Source Material	Creative expression, documents, or photographs
Method of capture	Geometry constructed with modeling software, which could be a polygonal surface modeler, solids modeler, in-world modeler, or similar. Software ranges in complexity from SketchUp to 3ds with data structures dependent on output (e.g., real-time vs. high-resolution animation). The models are possibly augmented with texture, materials, and/or lighting.
Hardware/software needed	There are over 100 modeling software packages that run on a variety of hardware platforms. The challenge is the amount of data in proprietary software that are dependent on specific versions of specific software applications. See Wikipedia, s.v. "List of 3D modeling software," last modified September 25, 2021, 1:10, https://en.wikipedia.org/wiki/List_of_3D_modeling_software.
Output files	Raw model file(s) whether connected or individually; digital research files related to the modeled environments; related textures maps, materials files, palettes, and shadow maps; physical archives related to the project (e.g., notes, physical photos, collected reference material); metadata; paradata; and text publications related to projects (see also maximum files for preservation below).
Derivatives	Could include interactive environment, animation, static images, models transferred to other file formats for 3D printing, secondary models created by others to explore alternative reconstructions, teaching and learning objects; documentaries or film productions that include content from the model; or VSim files formats (.vsim, .nar., .ere), in instances of reuse by original creator or secondary scholars, depending on the research objectives.
Methods of interaction	Depending on the use case, could be with a real-time viewer, uploaded into other 3D content types (e.g., virtual world, virtual immersive environment), or mixed with other 3D content types for use in other viewers.

Information	Examples
Minimum files for access	Presuming intent is reuse as 3D geometry and interaction as content creator intended: final version of the 3D model(s) in preservation file format and native file format; final textures maps, materials files, palettes, and shadow maps (file structure may be critical); stable version of the software required to view and interact with said files (e.g., Creator, 3ds, Maya, SketchUp/Google Earth); metadata and paradata that describe the projects and decisions made during the creation of the model. If a real-time environment, the final aggregated binary files and the software necessary to "fly" the model. If a Google Earth model, any GIS-related files and spreadsheets.
Maximum files for preservation	3D computer model(s) in different formats (e.g., .obj, .dae, native file formats); versioned copies of the 3D model files; textures maps, materials files, palettes, and shadow maps that go with the different model versions (file structure may be critical); stable version of the software required to view and interact with said files (e.g., Creator, 3ds, Maya, SketchUp/Google Earth); videos generated from the computer model; static images generated from the computer model (screenshots and **renderings**); metadata and paradata that describe the projects and decisions made during the creation of the model; research files and documents in various formats (scans, PDFs, bibliographic information, etc.). If a real-time environment, various iterations of the aggregated binary files and the software necessary to "fly" the model. If a Google Earth model, GIS-related files and spreadsheets if time periods are included. If used for creation of other materials (e.g., teaching resources or a film production), work files related to the final output and copies of that final output; analog documents and artifacts relating to the 3D model; correspondence related to the project; and publicity/marketing related to the project.

TABLE 6.A.2

Modality: Scanned volumetric

Information	Examples
Source material	Real-world object, time series volumes (fMRI).
Method of capture	Computed tomography (CT), magnetic resonance imaging (MRI), functional MRI, positron emission tomography (PET).
Hardware/software needed	Scanner (make, model, setting).
Output files	Package of images in sequence.

Information	Examples
Derivatives	JPG stacks, rendered/interpolated surfaces or volumes that could include information about the characteristics of the object (isosurfaces), colormaps, or a color lookup table (which can also have opacity from alpha channel).
Method of interaction	Desktop or web applications (ImageJ, Box DICOM, or a plethora of other (especially medical) applications.
Minimum files for access	Ordered stack of images (generally TIFF or JPG) with resolution, spacing of slices, number of slices.
Maximum files for preservation	Original DICOM files, TIFF stacks (or JPG derivatives), interpolated volumes or surfaces, documentation of capture and workflow.

TABLE 6.A.3

Modality: Scanned surface

Information	Examples
Source material	Real-world object.
Method of capture	Contact, active, conoscopic, structured light, modulated light, laser, microscribe.
Hardware/software needed	Scanner, software, lighting rig, enclosure, turntable.
Output files	OBJ, PLY, STL, X3D.
Derivatives	Lower poly count models for better web accessibility, portion of model for preservation of detail via web deliverable.
Method of interaction	Desktop application such as MeshLab, online viewer such as Sketchfab.
Minimum files for access	Constituent files of 3D model and metadata.
Maximum files for preservation	Original scan files, cleaned scan files, decimated files, documentation of capture and workflow.

TABLE 6.A.4

Modality: Photography-based

Information	Examples
Source material	Real-world objects captured through a variety of image formats from historical photos to terrain photos to photos of objects.
Method of capture	Flyover (landscape), light tent with turntable, circling object.
Hardware/software needed	GoPro, DLSR camera. See Wikipedia, s.v. "Comparison of photogrammetry software," last modified October 24, 2021, 20:55, https://en.wikipedia.org/wiki/Comparison_of_photogrammetry_software.

Information	Examples
Output files	OBJ, PLY, STL, X3D (raw and cleaned-up model; metadata, paradata on production).
Derivatives	Lower poly count for web display, watertight for printing.
Method of interaction	Sketchfab, VR environments, stand-alone players like 3DHOP.
Minimum files for access	Constituent files of 3D model, metadata.
Maximum files for preservation	RAW, unaltered photos; derived JPGs for stitching; unaltered, stitched 3D model; cleaned model; metadata of workflow.

TABLE 6.A.5

Modality: Procedural/algorithmic

Information	Examples
Source material	"Direct" import: GIS, laser scans (3D point clouds imported as polygons), photogrammetric data (imported as polygons/mesh such as .obj, .dae), photos (as textures). "Indirect" import: used in process to create data imports and GIS attributes—architectural plans, excavation maps, architectural drawings, photos, ethnographic/ethnohistoric descriptions.
Method of capture	Geometry generated from GIS data and rule-based script (computer graphics architecture—CGA) as well as expertise/interaction with data in software; qualitative and quantitative.
Hardware/software needed	Esri CityEngine (proprietary) and Terragen (work with GIS data); Acropora, Bryce, Modo, Cinema 4D, Esri CityEngine, Grome, Houdini, HyperFun, OpenSCAD, Softimage, VUE, PlantFactory, Xfrog, SpeedTree, Grasshopper 3D
Output files	Esri CityEngine (proprietary) and Terragen (work with GIS data); Acropora, Bryce, Modo, Cinema 4D, Esri CityEngine, Grome, Houdini, HyperFun, OpenSCAD, Softimage, VUE, PlantFactory, Xfrog, SpeedTree, Grasshopper 3D3D terrain models, and subsets of 3D models with terrain; .dae, .dxf, .fbx, .gdb, .kml, .kmz, .obj, .osm (import only), .vob (export only), .abc (export only), .rib (export only); Unreal Engine (export only), .3ws (CityEngine webscene), .3VR (standard VR format—export only); unlike exporting these data from a GIS, you can specify whether you want to retain materials and textures, whether you need to write the normals or even triangulate the meshes; .cga file (text file with script/code) serves as paradata and possibly as metadata depending on comments.
Derivatives	3D single object models (as polygons), 3D terrain models, 3D terrain models with 3D architectural models/textures to be used in various software.
Method of interaction	Within CityEngine or CityGML. Export to ArcGIS Pro, WebGL environments, game engines (e.g., Unity, Unreal Engine), 3D object viewers (e.g., 3DHOP, Sketchfab).

Information	Examples
Minimum files for access	CGA file (text), GIS data (specifically speaking of CityEngine).
Maximum files for preservation	Presuming intent is reuse as 3D geometry and interaction as content creator intended: final version of the 3D model(s) in preservation file format and native file format; final textures maps, materials files (file structure may be critical); stable version of the software required to view and interact with said files (e.g., CityEngine, but other 3D software programs for exports); metadata and paradata that describe the projects and decisions made during the creation of the model. If a real-time environment, the final aggregated binary files and the software necessary to "fly" the model. If a Google Earth model, any GIS-related files and spreadsheets.

TABLE 6.A.6
Modality: Digital terrain

Information	Examples
Source material	Satellite imagery, GIS data (vector [contours] or raster [DEM]), aerial imagery (photogrammetry/stereo pairs), airborne lidar, height maps.
Method of capture	Vector to raster conversion (topo to raster)—interpolation; 3D points to raster (e.g., total station, lidar)—interpolation; stereo pairs (aerial imagery); photogrammetric methods (satellite/aerial imagery); TINs.
Hardware/software needed	GIS software, e.g., GIS, ArcGIS; photogrammetric (SFM) software, e.g., Agisoft PhotoScan (now Metashape); lidar processing software, e.g., Global Mapper, CloudCompare). Hardware: high processing power and RAM.
Output files	Raster files such as Esri GRID, GeoTIFF, DEM, ASCII, STRM1, STRM3, STRM30, ASTER, GTOPO30.
Derivatives	Lower resolution raster files; heightmaps; urban DEM; digital surface model (DSM); digital terrain model (DTM); digital elevation model (DEM).
Modes of interaction	Desktop, mobile, web applications.
Minimum files for access	Raster file (e.g., Esri GRID, GeoTIFF)
Maximum files for preservation	Original data sources (e.g., total station points, 3D points, contour lines), interpolated raster surface (e.g., Esri GRID, GeoTIFF); surface derivatives (e.g., heightmaps, DSM, etc.).

TABLE 6.A.7

Modality: Virtual world

Information	Examples
Source material	The most useful definition of a virtual world contends that it is a computer-simulated representation of a world with specific spatial and physical characteristics, and users of virtual worlds interact with each other via representations of themselves called "**avatars**." Virtual worlds are three-dimensional environments in which you can interact with others and create objects as part of that interaction. Confusion over the term has resulted in a fragmented understanding in the existing literature of what a virtual world is and is not. There are a range of virtual worlds to choose from, which include fantasy, sport, historical, and science fiction. Some are loosely based upon the real world, but others, such as fantasy worlds, are completely disconnected from the real world, which is also part of their attraction. To further complicate this problem, a variety of terms are used in the literature to label the technology: virtual world (VW); virtual environment (VE); multiuser virtual environment (MUVE); massively multiplayer online game (MMOG); immersive virtual world (IVW); serious virtual world; social virtual world; and synthetic virtual world. Most recently, a virtual world has been defined as "Shared, simulated spaces which are inhabited and shaped by their inhabitants who are represented as avatars. These avatars mediate our experience of this space as we move, interact with objects and interact with others, with whom we construct a shared understanding of the world at that time" (Carina Girvan, "What Is a Virtual World? Definition and Classification," Educational Technology Research and Development 66, no. 5 [2018]: 1099).
Method of capture	Geometry constructed with modeling software, which could be a polygonal surface modeler, solids modeler, in-world modeler, or similar. Software ranges in complexity from SketchUp to 3ds with data structures dependent on output (e.g., real-time vs. high-resolution animation). The models are possibly augmented with texture, materials, and/or lighting.
Hardware/software needed	There are over 100 modeling software packages that run on a variety of hardware platforms. The challenge is the amount of data in proprietary software that are dependent on specific versions of specific software. See Wikipedia, s.v. "List of 3D modeling software," last modified September 25, 2021, 1:10, https://en.wikipedia.org/wiki/List_of_3D_modeling_software.
Output files	Raw model file(s) whether connected or individually; digital research files related to the modeled environments; related textures maps, materials files, palettes, and shadow maps; physical archives related to the project (e.g., notes, physical photos, collected reference material); metadata; paradata; and text publications related to projects. See also maximum files for preservation below.

Information	Examples
Derivatives	Could include interactive environment, animation, static images, models transferred to other file formats for 3D printing, secondary models created by others to explore alternative reconstructions, teaching and learning objects; documentaries or film productions that include content from the model; or VSim files formats (.vsim, .nar., .ere), in instances of reuse by original creator or secondary scholars, depending on the research objectives.
Method of interaction	Depending on the use case, could be with a real-time viewer, uploaded into other 3D content types (e.g., virtual world, virtual immersive environment), or mixed with other 3D content types for use in other viewers.
Minimum files for access	Presuming intent is reuse as 3D geometry and interaction as content creator intended: final version of the 3D model(s) in preservation file format and native file format; final textures maps, materials files, palettes, and shadow maps (file structure may be critical); stable version of the software required to view and interact with said files (e.g., Creator, 3ds, Maya, SketchUp/Google Earth); metadata and paradata that describe the projects and decisions made during the creation of the model. If a real-time environment, the final aggregated binary files and the software necessary to "fly" the model. If a Google Earth model, any GIS-related files and spreadsheets.
Maximum files for preservation	3D computer model(s) in different formats (e.g., .obj, .dae, native file formats); versioned copies of the 3D model files; textures maps, materials files, palettes, and shadow maps that go with the different model versions (file structure may be critical); stable version of the software required to view and interact with said files (e.g., Creator, 3ds, Maya, SketchUp, Google Earth); videos generated from the computer model; static images generated from the computer model (screenshots and renderings); metadata and paradata that describe the projects and decisions made during the creation of the model; research files and documents in various formats (scans, PDFs, bibliographic information, etc.). If a real-time environment, various iterations of the aggregated binary files and the software necessary to "fly" the model. If a Google Earth model, GIS-related files and spreadsheets if time periods are included. If used for creation of other materials (e.g., teaching resources or a film production), work files related to the final output and copies of that final output; analog documents and artifacts relating to the 3D model; correspondence related to the project; and publicity/marketing related to the project.

TABLE 6.A.8
Modality: Immersive virtual environment

Information	Examples
Source material	Virtual "environments" are distinguished from virtual "worlds" to emphasize the use of a headset for the experience and that they do not necessarily require a sense of physical place. A virtual environment is more focused on creating an immersive experience, as dictated by the creator. In addition, to distinguish it from Mark Bell's definition of "virtual worlds," immersive virtual environments do not usually have (as yet) a time element in the way that things like Second Life do, where the world continues on without the user's engagement.
Method of capture	Construction of immersive virtual environments can happen two ways: outside the environment or inside the environment. Production outside the environment can range from geometry constructed with modeling software to experimental. In-environment creation includes tools such as Tilt Brush or Sketchbox, and will likely grow as immersive technology develops.
Hardware/software needed	Typically a headset display will interface with a software application through the device or through an application on a mobile phone to allow for experiences to run. Sometimes hand controller, sensor stands, and/or headphones may be necessary to navigate the software within the headset. While the combination of technologies can vary, there are tailored specifications for display refresh rates, graphics cards, screen resolution that are needed for certain digital immersive software to run. Often released by hardware producers, software development kits (SDKs) are used in the development of many extended reality products. Some hardware producers offer emulators for software developers to use if they do not have headsets to test with. In order to be disseminated more widely, these digitally immersive extended reality products usually need to be packaged for one of the app stores where they can be released and played.
Output files	Application/package release to an app store such as the Apple App Store, Google Play Store, YouTube 360, Facebook 360, Little Star, Jaunt, 360 RIZE/360Heros, Discovery VR, WAVRP, 360s.tv (adult content), Oculus App Store, Steam VR.
Derivatives	Reuse by original creator or secondary scholars: Could include interactive environment, animation, static images, models transferred to other file formats for 3D printing, secondary models created by others to explore alternative reconstructions, teaching and learning objects; transfer to virtual worlds
Method of interaction	Headset, controllers.
Minimum files for access	Adherence to specific app store or platform distribution requirements.

Information	Examples
Maximum files for preservation	3D computer model(s) in different formats (e.g., .obj, .dae, native file formats); versioned copies of the 3D model files; textures maps, materials files, palettes, and shadow maps that go with the different model versions (file structure may be critical); stable version of the software required to view and interact with said files (e.g., Creator, 3ds, Maya, SketchUp, Google Earth); videos generated from the computer model; static images generated from the computer model (screenshots and renderings); metadata and paradata that describe the projects and decisions made during the creation of the model; research files and documents in various formats (scans, PDFs, bibliographic information, etc.). If a real-time environment, various iterations of the aggregated binary files and the software necessary to "fly" the model. If a Google Earth model, GIS-related files and spreadsheets if time periods are included. If used for creation of other materials (e.g., teaching resources or a film production), work files related to the final output and copies of that final output; analog documents and artifacts relating to the 3D model; correspondence related to the project; and publicity/marketing related to the project; emulator program if hardware no longer exists.

TABLE 6.A.9

Modality: Games

Information	Examples
Source material	Games will have unique preservation challenges, based on what kind access the preserved material should provide, e.g., a walk-through that reduces the game to a singular experience or a virtual machine that allows interactive play. Games can have components from the real or imagined world, and components can also be algorithmically generated.
Method of capture	Constructed base structures, manipulatives, scenes that define "relationships between objects, including location and size" (Wikipedia, s.v. "3D computer graphics, last modified November 21, 2021, 4:47, https://en.wikipedia.org/wiki/3D_computer_graphics).
Hardware/software needed	See the following lists on Wikipedia: • "List of 3D computer graphics software," https://en.wikipedia.org/wiki/List_of_3D_computer_graphics_software • "List of 3D modeling software," https://en.wikipedia.org/wiki/List_of_3D_modeling_software • "List of 3D rendering software," https://en.wikipedia.org/wiki/List_of_3D_rendering_software.
Output files	Package released to platform of choice.

Information	Examples
Derivatives	Audiences watching games, video walk-throughs, virtual machine preservation.
Method of interaction	First- or third-person POV (third-person perspective is often used as a camera position in the game) with additional actions dependent on the controller and platform in use.
Minimum files for access	Minimum package required for publication to platform of choice.
Maximum files for preservation	Versions of the game in development, walk-through, metadata of construction and intended use, virtual machine of game.

Notes

1. GO FAIR, "FAIR Principles," accessed May 4, 2020, https://www.go-fair.org/fair-principles/.
2. For literature regarding the use of personas for research, see Lene Nielsen and Kira Storgaard Hansen, "Personas Is Applicable: A Study on the Use of Personas in Denmark," in *CHI '14: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York: Association for Computing Machinery, 2014), 1665–74, https://doi.org/10.1145/2556288.2557080; John Pruitt and Jonathan Grudin, "Personas: Practice and Theory," in *DUX '03: Proceedings of the 2003 Conference on Designing for User Experiences* (New York: Association for Computing Machinery, 2003), 1–15, https://doi.org/10.1145/997078.997089; Cynthia Putnam, Beth Kolko, and Siri Wood, "Communicating about Users in ICTD: Leveraging HCI Personas," in *ICTD '12: Proceedings of the Fifth International Conference on Information and Communication Technologies and Development* (New York: Association for Computing Machinery, 2012), 338–49, https://doi.org/10.1145/2160673.2160714; Jack M. Maness, Tomasz Miaskiewicz, and Tamara Sumner, "Using Personas to Understand the Needs and Goals of Institutional Repositories," *D-Lib Magazine* 14, no. 9/10 (September/October 2008), http://www.dlib.org/dlib/september08/maness/09maness.html.
3. Richard M. Young and Phil Barnard, "The Use of Scenarios in Human-Computer Interaction Research: Turbocharging the Tortoise of Cumulative Science," *ACM SIGCHI Bulletin* 18, no. 4 (April 1987): 291, https://doi.org/10.1145/1165387.275645.
4. Pew Research Center. "Internet/Broadband Fact Sheet." April 7, 2021. https://www.pewresearch.org/internet/fact-sheet/internet-broadband/.
5. Statista Research Department, "Share of Households Equipped with a Computer in Mexico from 2001 to 2018," April 18, 2019, https://www.statista.com/statistics/712803/household-ownership-computer-mexico/ (page content changed).
6. Statista Research Department, "Smartphone Users as Share of the Population in Mexico from 2017 to 2023," January 17, 2019, https://www.statista.com/statistics/625424/smartphone-user-penetration-in-mexico/ (page content changed).
7. "Smithsonian Digitization, "Smithsonian Voyager: Open Source 3D Explorer and Authoring Tool Suite," accessed April 15, 2020, https://smithsonian.github.io/dpo-voyager/; 3DHOP home page, accessed April 15, 2020, http://vcg.isti.cnr.it/3dhop/.
8. MeshLab home page, accessed April 15, 2020, https://www.meshlab.net/; CloudCompare home page, accessed April 15, 2020, https://www.danielgm.net/cc/.
9. MorphoSource home page, accessed April 15, 2020, https://www.morphosource.org/.
10. Jenny Zhao et al., "Validating Photogrammetric Processes Using Jetstream and Synthetic Images," presentation, IUScholarWorks, Indiana University Bloomington, July 28, 2019, https://doi.org/10.5967/6VC0-1M43.

11. Stanford Computer Graphics Laboratory, "The Stanford 3D Scanning Repository," accessed April 15, 2020, http://graphics.stanford.edu/data/3Dscanrep/; Blender home page, accessed April 15, 2020, https://www.blender.org/; Agisoft home page, accessed April 15, 2020, https://www.agisoft.com/.
12. David Koller, Bernard Frischer, and Greg Humphreys, "Research Challenges for Digital Archives of 3D Cultural Heritage Models," *Journal on Computing and Cultural Heritage* 2, no. 3 (December 2009): 7, https://doi.org/10.1145/1658346.1658347.
13. 3D ICONS portal, last modified June 7, 2015, http://3dicons.ceti.gr/index.php.
14. Heinz Pampel et al., "Making Research Data Repositories Visible: The re3data.org Registry," *PLOS ONE* 8, no. 11 (2013): e78080, https://doi.org/10.1371/journal.pone.0078080.
15. Web3D Consortium, "Recommended Standards," accessed October 1, 2019, https://www.web3d.org/standards.
16. Library of Congress, "Born to Be 3D: Digital Stewardship of Intrinsic 3D Data" (workshop, Washington, DC, November 2, 2019), https://www.loc.gov/preservation/digital/meetings/b2b3d/b2b3d2018.html.
17. Martina Trognitz, Kieron Niven, and Valentijn Gilissen, *3D Models in Archaeology: A Guide to Good Practice*, Guides to Good Practice (Archaeology Data Service and Digital Antiquity, 2016), https://guides.archaeologydataservice.ac.uk/g2gp/3d_Toc; Franco Niccolucci, "Setting Standards for 3D Visualization of Cultural Heritage in Europe and Beyond," in *Paradata and Transparency in Virtual Heritage*, ed. Anna Bentkowska-Kafel, Hugh Denard, and Drew Baker (Farnham, UK: Ashgate, 2016), 23–36.
18. Heather Richards-Rissetto and Jennifer von Schwerin, "A Catch 22 of 3D Data Sustainability: Lessons in 3D Archaeological Data Management and Accessibility," *Digital Applications in Archaeology and Cultural Heritage* 6 (2017): 38–48, https://doi.org/10.1016/j.daach.2017.04.005.
19. W3C Web Accessibility Initiative, "Web Content Accessibility Guidelines (WCAG) Overview," 2005, last modified April 29, 2021, https://www.w3.org/WAI/standards-guidelines/wcag/.
20. Shawn Lawton Henry, Shadi Abou-Zahra, and Kevin White, eds. "Accessibility, Usability, and Inclusion." W3C Web Accessibility Initiative, March 2010, last modified May 6, 2016. https://www.w3.org/WAI/fundamentals/accessibility-usability-inclusion/.
21. Jasmine Clark, "Project Gap Analysis Rubric" included as Appendix A in Francesca Albrezzi, "Virtual Actualities: Technology, Museums, and Immersion." (PhD diss, University of California, Los Angeles, 2019), 234, 250–53, 257–60, https://escholarship.org/uc/item/5tc2q2dt.
22. Fulcrum. 2018. "Accessibility." Last updated: October 15, 2018, https://www.fulcrum.org/accessibility.
23. Rachel Opitz, Marcello Mogetta, and Nicola Terrenato, eds., *A Mid-Republican House from Gabii* (Ann Arbor: University of Michigan Press, 2016); Rachel Opitz, "Publishing Archaeological Excavations at the Digital Turn," *Journal of Field Archaeology* 43, no. sup1 (2018): S68–S82.
24. Jasmine Clark, "Designing Accessible, Usable, and Inclusive Digital Projects" (webinar. January 2019). https://charlesstudy.temple.edu/event/4920107, cited in Albrezzi, "Virtual Actualities," appendix 2.
25. See Anthony Caldwell and John A. Lynch, "Approaches to Sharing Modeling Data" (presentation, CHNT 20: Conference on Cultural Heritage and New Technologies, Vienna, Austria, November 2–4, 2015), https://www.chnt.at/approaches-to-sharing-modeling-data/; Elaine A. Sullivan and Lisa M. Snyder, "Digital Karnak: An Experiment in Publication and Peer Review of Interactive, Three-Dimensional Content," *Journal of the Society of Architectural Historians* 76, no. 4 (2017): 464–82; Albrezzi, "Virtual Actualities"; Susan Schreibman and Costas Papadopoulos, "Textuality in 3D: Three-Dimensional (Re)constructions as Digital Scholarly Editions," *International Journal of Digital Humanities* 1, no. 2 (2019): 221–33.
26. Luis Gutierrez-Heredia et al., "End to End Digitisation and Analysis of Three-Dimensional Coral Models, from Communities to Corallites," *PLOS ONE* 11, no. 2 (2016): e0149641, https://doi.org/10.1371/journal.pone.0149641; G. C. Young et al., "Cost and Time-Effective Method for Multi-scale Measures of Rugosity, Fractal Dimension, and Vector Dispersion from Coral Reef 3D Models," *PLOS ONE* 12, no. 4 (2017): e0175341, https://doi.org/10.1371/journal.pone.0175341; Renata Ferrari et al., "3D Photogrammetry Quantifies Growth and External Erosion of Individual Coral Colonies and Skeletons," *Scientific Reports* 7, no. 1 (2017): article 16737, https://doi.org/10.1038/s41598-017-16408-z.
27. Bernard Frischer, "New Light on the Relationship between the Montecitorio Obelisk and Ara Pacis of Augustus," (html version) *Studies in Digital Heritage* 1, no. 1 (2017): 18, 20, https://doi.org/10.14434/sdh.v1i1.23331.

28. John Bonnett, "A Plea for Design: Historians, Digital Platforms, and the Mindful Dissemination of Content and Concepts," *Journal of the Canadian Historical Association* 25, no. 2 (2014): 189–231, https://doi.org/10.7202/1032846ar.
29. Cultural Heritage Imaging Digital Lab Notebook page, accessed February 24, 2022, http://culturalheritageimaging.org/Technologies/Digital_Lab_Notebook/.
30. Heather Richards-Rissetto et al., Keeping Data Alive: Supporting Reuse and Repurposing of 3D Data in the Humanities website, accessed October 1, 2019, https://cdrhsites.unl.edu/keeping-data-alive/workflows.html.
31. Ben Marwick and Suzanne Birch, "A Standard for the Scholarly Citation of Archaeological Data as an Incentive to Data Sharing," *Advances in Archaeological Practice* 6, no. 2 (May 2018): 125-143, https://doi.org/10.1017/aap.2018.3.
32. Virtual Hampson Museum home page, accessed October 2, 2019, https://hampson.cast.uark.edu.
33. See Arno Scharl and Klaus Tochtermann, eds., *The Geospatial Web* (London: Springer-Verlag, 2007).
34. IIIF 3D Community Group website, accessed April 15, 2020, https://iiif.io/community/groups/3d/.
35. Centre for Excellence in Universal Design, "What Is Universal Design?" National Disability Authority, accessed April 15, 2020, https://universaldesign.ie/What-is-Universal-Design/.
36. Snyder, Lisa M. "VSim: Scholarly Annotations in Real-Time 3D Environments,"in *DH-CASE '14: DH-CASE II: Collaborative Annotations on Shared Environments: Metadata, Tools and Techniques in the Digital Humanities*, ed. Patrick Schmitz and Laurie Pearce (New York: Association for Computing Machinery, 2014), 1–8, https://doi.org/10.1145/2657480.2657483.
37. Costas Papadopoulos and Susan Schreibman, "Towards 3D Scholarly Editions: The Battle of Mount Street Bridge," *Digital Humanities Quarterly* 13, no. 1 (2019), http://www.digitalhumanities.org/dhq/vol/13/1/000415/000415.html.
38. Laura Mandell, "Promotion and Tenure for Digital Scholarship," *Journal of Digital Humanities* 1, no. 4 (2012), http://journalofdigitalhumanities.org/1-4/promotion-and-tenure-for-digital-scholarship-by-laura-mandell/.
39. Geoffrey Rockwell, "Short Guide to Evaluation of Digital Work," *Journal of Digital Humanities* 1, no. 4 (2012), http://journalofdigitalhumanities.org/1-4/short-guide-to-evaluation-of-digital-work-by-geoffrey-rockwell/; Ad Hoc Committee on the Evaluation of Digital Scholarship by Historians, *Guidelines for the Professional Evaluation of Digital Scholarship by Historians* (Washington, DC: American Historical Association, June 15, 2015), https://www.historians.org/teaching-and-learning/digital-history-resources/evaluation-of-digital-scholarship-in-history/guidelines-for-the-professional-evaluation-of-digital-scholarship-by-historians; Jon Driver et al., *Report of the SAA Task Force on Guidelines for Promotion and Tenure for Archaeologists in Diverse Academic Roles*, white paper (Washington, DC: Society for American Archaeology, January 31, 2018), 11–15, https://documents.saa.org/container/docs/default-source/doc-careerpractice/promotion-and-tenure-white-paper-02-12-19-web.pdf?sfvrsn=8a5815cd_4; Task Force to Develop Guidelines for Evaluating Digital Art and Architectural History for Promotion and Tenure, *Guidelines for the Evaluation of Digital Scholarship in Art and Architectural History* (College Art Association and the Society of Architectural Historians, January 2016), https://www.collegeart.org/pdf/evaluating-digital-scholarship-in-art-and-architectural-history.pdf.
40. Bernard S. Frischer et al., "From CVR to CVRO: The Past, Present, and Future of Cultural Virtual Reality," in *Territorial Marketing: Proceedings of the VAST EuroConference, Arezzo, Italy, 24–25 November 2000*, ed. Franco S. Niccolucci, Oxford, UK. Archeopress.: 2002), 7–18; London Charter, home page, accessed May 31, 2019, http://www.londoncharter.org/.

Bibliography

3DHOP home page. Accessed April 15, 2020. http://vcg.isti.cnr.it/3dhop/.

3D ICONS portal. Last modified June 7, 2015. http://3dicons.ceti.gr/index.php.

Ad Hoc Committee on the Evaluation of Digital Scholarship by Historians. *Guidelines for the Professional Evaluation of Digital Scholarship by Historians*.

Washington, DC: American Historical Association, June 15, 2015. https://www.historians.org/teaching-and-learning/digital-history-resources/evaluation-of-digital-scholarship-in-history/guidelines-for-the-professional-evaluation-of-digital-scholarship-by-historians.

Agisoft home page. Accessed April 15, 2020. https://www.agisoft.com/.

Albrezzi, Francesca. "Virtual Actualities: Technology, Museums, and Immersion." PhD diss., University of California, Los Angeles, 2019. https://escholarship.org/uc/item/5tc2q2dt.

Blender home page. Accessed April 15, 2020. https://www.blender.org/.

Bonnett, John. "A Plea for Design: Historians, Digital Platforms, and the Mindful Dissemination of Content and Concepts." *Journal of the Canadian Historical Association* 25, no. 2 (2014): 189–231.

Caldwell, Anthony, and John A. Lynch. "Approaches to Sharing Modeling Data." Presentation, CHNT 20: Conference on Cultural Heritage and New Technologies, Vienna, Austria, November 2–4, 2015. https://www.chnt.at/approaches-to-sharing-modeling-data/.

Centre for Excellence in Universal Design. "What Is Universal Design?" National Disability Authority. Accessed April 15, 2020. https://universaldesign.ie/What-is-Universal-Design/.

Clark, Jasmine. "Designing Accessible, Usable, and Inclusive Digital Projects." Webinar. January 2019. https://charlesstudy.temple.edu/event/4920107.

CloudCompare home page. Accessed April 15, 2020. https://www.danielgm.net/cc/.

Delve, Janet. "Introduction to POCOS E-book 1: Preserving Visualisations and Simulations." In *The Preservation of Complex Objects*. Edited by Janet Delve, David Anderson, Milena Dobreva, Drew Baker, Clive Billenness, and Leo Konstantelos, 9–14. Portsmouth, UK: University of Portsmouth, 2012.

Driver, Jon, Ted Goebel, Lynne Goldstein, P. Nick Kardulias, Fred Limp, Heather Richards-Rissetto, and LuAnn Wandsnider. *Report of the SAA Task Force on Guidelines for Promotion and Tenure for Archaeologists in Diverse Academic Roles*. White paper. Washington, DC: Society for American Archaeology, January 31, 2018. https://documents.saa.org/container/docs/default-source/doc-careerpractice/promotion-and-tenure-white-paper-02-12-19-web.pdf?sfvrsn=8a5815cd_4.

Ferrari, Renata, Will F. Figueira, Morgan S. Pratchett, Tatiana Boube, Arne Adam, Tania Kobelkowsky-Vidrio, Steve S. Doo, et al. "3D Photogrammetry Quantifies Growth and External Erosion of Individual Coral Colonies and Skeletons." *Scientific Reports* 7, no. 1 (2017): article 16737. https://doi.org/10.1038/s41598-017-16408-z.

Frischer, Bernard. "New Light on the Relationship between the Montecitorio Obelisk and Ara Pacis of Augustus." *Studies in Digital Heritage* 1, no. 1 (2017): 18–119. https://doi.org/10.14434/sdh.v1i1.23331.

Frischer, Bernard, Franco S. Niccolucci, Nick S. Ryan, and Juan S. Barceló. "From CVR to CVRO: The Past, Present, and Future of Cultural Virtual Reality." In *Territorial Marketing: Proceedings of the VAST EuroConference, Arezzo, Italy, 24-25 November 2000*. Edited by Franco S. Niccolucci, Oxford, UK. Archeopress.: 7-18.2002.

Fulcrum. 2018. "Accessibility." Accessed January 21, 2022. www.fulcrum.org/accessibility.

Girvan, Carina. "What Is a Virtual World? Definition and Classification." *Educational Technology Research and Development* 66, no. 5 (2018): 1087–1100.

GO FAIR. "FAIR Principles." Accessed May 4, 2020. https://www.go-fair.org/fair-principles/.

Gutierrez-Heredia, Luis, Francesca Benzoni, Emma Murphy, and Emmanuel Reynaud. "End to End Digitisation and Analysis of Three-Dimensional Coral Models, from Communities to Corallites." *PLOS ONE* 11, no. 2 (2016): e0149641. https://doi.org/10.1371/journal.pone.0149641.

Henry, Shawn Lawton, Shadi Abou-Zahra, and Kevin White, eds. "Accessibility, Usability, and Inclusion." W3C Web Accessibility Initiative. March 2010, last modified May 6, 2016. https://www.w3.org/WAI/fundamentals/accessibility-usability-inclusion/.

IIIF 3D Community Group website. Accessed April 15, 2020. https://iiif.io/community/groups/3d/.

Khunti, Roshni. "The Problem with Printing Palmyra: Exploring the Ethics of Using 3D Printing Technology to Reconstruct Heritage." *Studies in Digital Heritage* 2, no. 1 (2018): 1–12. https://doi.org/10.14434/sdh.v2i1.24590.

Koller, David, Bernard Frischer, and Greg Humphreys. "Research Challenges for Digital Archives of 3D Cultural Heritage Models." *Journal on Computing and Cultural Heritage* 2, no. 3 (December 2009): 1–17, https://doi.org/10.1145/1658346.1658347.

Library of Congress. "Born to Be 3D: Digital Stewardship of Intrinsic 3D Data." Workshop, Washington, DC, November 2, 2019. https://www.loc.gov/preservation/digital/meetings/b2b3d/b2b3d2018.html.

London Charter home page. Accessed May 31, 2019. http://www.londoncharter.org/.

Mandell, Laura. "Promotion and Tenure for Digital Scholarship." *Journal of Digital Humanities* 1, no. 4 (2012). http://journalofdigitalhumanities.org/1-4/promotion-and-tenure-for-digital-scholarship-by-laura-mandell/.

Maness, Jack M., Tomasz Miaskiewicz, and Tamara Sumner. "Using Personas to Understand the Needs and Goals of Institutional Repositories." *D-Lib Magazine* 14, no. 9/10 (September/October 2008). http://www.dlib.org/dlib/september08/maness/09maness.html.

Marwick, Ben, and Suzanne Birch. "A Standard for the Scholarly Citation of Archaeological Data as an Incentive to Data Sharing." *Advances in Archaeological Practice* 6, no. 2 (May 2018). https://doi.org/10.1017/aap.2018.3.

MeshLab home page. Accessed April 15, 2020. https://www.meshlab.net/.

MorphoSource. Home page. Accessed April 15, 2020. https://www.morphosource.org/.

Niccolucci, Franco. "Setting Standards for 3D Visualization of Cultural Heritage in Europe and Beyond." In *Paradata and Transparency in Virtual Heritage*. Edited by Anna Bentkowska-Kafel, Hugh Denard, and Drew Baker, 23–36. Farnham, UK: Ashgate, 2016.

Nielsen, Lene, and Kira Storgaard Hansen. "Personas Is Applicable: A Study on the Use of Personas in Denmark." In *CHI '14: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1665–74. New York: Association of Computing Machinery, 2014. https://doi.org/10.1145/2556288.2557080.

Opitz, Rachel. "Publishing Archaeological Excavations at the Digital Turn." *Journal of Field Archaeology* 43, no. sup1 (2018): S68–S82.

Opitz, Rachel, Marcello Mogetta, and Nicola Terrenato, eds. *A Mid-Republican House from Gabii*. Ann Arbor: University of Michigan Press, 2016.

Pampel, Heinz, Paul Vierkant, Frank Scholze, Roland Bertelmann, Maxi Kindling, Jens Klump, Hans-Jürgen Goebelbecker, et al. "Making Research Data Repositories Visible: The re3data.org Registry." *PLOS ONE* 8, no. 11 (2013): e78080. https://doi.org/10.1371/journal.pone.0078080.

Papadopoulos, Costas, and Susan Schreibman. "Towards 3D Scholarly Editions: The Battle of Mount Street Bridge." *Digital Humanities Quarterly* 13, no.1 (2019). http://www.digitalhumanities.org/dhq/vol/13/1/000415/000415.html.

Pew Research Center. "Internet/Broadband Fact Sheet." April 7, 2021. https://www.pewresearch.org/internet/fact-sheet/internet-broadband/

Pruitt, John, and Jonathan Grudin. "Personas: Practice and Theory." In *DUX '03: Proceedings of the 2003 Conference on Designing for User Experiences*, 1–15. New York: Association for Computing Machinery, 2003. https://doi.org/10.1145/997078.997089.

Putnam, Cynthia, Beth Kolko, and Siri Wood. "Communicating about Users in ICTD: Leveraging HCI Personas." In *ICTD '12: Proceedings of the Fifth International Conference on Information and Communication Technologies and Development*, 338–49. New York: Association for Computing Machinery, 2012. https://doi.org/10.1145/2160673.2160714.

Richards-Rissetto, Heather, Rachel Optiz, Karin Dalziel, Jessica Dussault, and Greg Tunink. Keeping Data Alive: Supporting Reuse and Repurposing of 3D Data in the Humanities website. Accessed October 1, 2019. https://cdrhsites.unl.edu/keeping-data-alive/.

Richards-Rissetto, Heather and Jennifer von Schwerin, "A Catch 22 of 3D Data Sustainability: Lessons in 3D Archaeological Data Management & Accessibility." *Digital Applications in Archaeology and Cultural Heritage* 6 (2017): 38-48. https://doi.org/10.1016/j.daach.2017.04.005.

Rockwell, Geoffrey. "Short Guide to Evaluation of Digital Work." *Journal of Digital Humanities* 1, no. 4 (2012). http://journalofdigitalhumanities.org/1-4/short-guide-to-evaluation-of-digital-work-by-geoffrey-rockwell/.

Scharl, Arno, and Klaus Tochtermann, eds. *The Geospatial Web: How Geobrowsers, Social Software and the Web 2.0 Are Shaping the Network Society*. London: Springer-Verlag, 2007.

Schreibman, Susan, and Costas Papadopoulos. "Textuality in 3D: Three-Dimensional (Re)constructions as Digital Scholarly Editions." *International Journal of Digital Humanities* 1, no. 2 (2019): 221–33.

Smithsonian Digitization. "Smithsonian Voyager: Open Source 3D Explorer and Authoring Tool Suite." Accessed April 15, 2020. https://smithsonian.github.io/dpo-voyager/.

Snyder, Lisa M. "VSim: Scholarly Annotations in Real-Time 3D Environments." In *DH-CASE '14: DH-CASE II: Collaborative Annotations on Shared Environments: Metadata, Tools and Techniques in the Digital Humanities*. Edited by Patrick Schmitz and Laurie Pearce,1–8. New York: Association for Computing Machinery, 2014. https://doi.org/10.1145/2657480.2657483.

Stanford Computer Graphics Laboratory. "The Stanford 3D Scanning Repository." Accessed April 15, 2020. http://graphics.stanford.edu/data/3Dscanrep/.

Statista Research Department. "Share of Households Equipped with a Computer in Mexico from 2001 to 2018." April 18, 2019. https://www.statista.com/statistics/712803/household-ownership-computer-mexico/ (page content changed).

Statista Research Department. "Smartphone Users as Share of the Population in Mexico from 2017 to 2023." January 17, 2019. https://www.statista.com/statistics/625424/smartphone-user-penetration-in-mexico/ (page content changed).

Sullivan, Elaine A., and Lisa M. Snyder. "Digital Karnak: An Experiment in Publication and Peer Review of Interactive, Three-Dimensional Content." *Journal of the Society of Architectural Historians* 76, no. 4 (2017): 464–82.

Task Force to Develop Guidelines for Evaluating Digital Art and Architectural History for Promotion and Tenure. *Guidelines for the Evaluation of Digital Scholarship in Art and Architectural History*. College Art Association and the Society of Architectural Historians, January 2016. https://www.collegeart.org/pdf/evaluating-digital-scholarship-in-art-and-architectural-history.pdf.

Trognitz, Martina, Kieron Niven, and Valentijn Gilissen. *3D Models in Archaeology: A Guide to Good Practice*. Guides to Good Practice. Archaeology Data Service and Digital Antiquity, 2016. https://guides.archaeologydataservice.ac.uk/g2gp/3d_Toc.

US Department of Health and Human Services. "Policies and Standards." Accessed January 21, 2022. https://www.hhs.gov/web/policies-and-standards/index.html.

Veldhuizen, Bart. "Stream Massive 3D Models, Now with Texture Support." *Sketchfab Blog*, July 31, 2019. https://sketchfab.com/blogs/community/stream-massive-models-now-with-texture-support.

Virtual Hampson Museum home page. Accessed October 2, 2019. https://hampson.cast.uark.edu.

W3C Web Accessibility Initiative. "Web Content Accessibility Guidelines (WCAG) Overview." 2005, last modified April 29, 2021. https://www.w3.org/WAI/standards-guidelines/wcag/.

Web3D Consortium. "Recommended Standards." Accessed April 15, 2020. https://www.web3d.org/standards.

Young, G. C., S. Dey, A. D. Rogers, and D. Exton. "Cost and Time-Effective Method for Multi-scale Measures of Rugosity, Fractal Dimension, and Vector Dispersion from Coral Reef 3D Models." *PLOS ONE* 12, no. 4 (2017): e0175341. https://doi.org/10.1371/journal.pone.0175341.

Young, Richard M., and Phil Barnard. "The Use of Scenarios in Human-Computer Interaction Research: Turbocharging the Tortoise of Cumulative Science." *ACM SIGCHI Bulletin* 18, no. 4 (April 1987): 291–96. https://doi.org/10.1145/1165387.275645.

Zhao, Jenny, Matthew Mercer, Tassie Gniady, and Eric Wernert. "Validating Photogrammetric Processes Using Jetstream and Synthetic Images." Presentation. IUScholarWorks, Indiana University Bloomington, July 28, 2019. https://doi.org/10.5967/6VC0-1M43.

Chapter 7

Conclusion

Jennifer Moore, Adam Rountrey, and Hannah Scates Kettler

Over the course of producing this volume, needs and goals have changed as the technical landscape of 3D and applications of 3D data have continued to expand and evolve. Yet there is a lurking danger of 3D data being lost or becoming obsolete. Therefore, it is critical that we take the steps of documenting the state of the field and making recommendations based on current needs to provide a starting point for what will need to persist as an adaptable, extensible system of standards. Identifying some **preservation** commonalities across disciplines and use cases helps avoid further development of the already initiated "siloed standards," and it allows creation of generalized repositories that at least meet some of all users' needs. Thus, while we see this volume as an important milestone in 3D data preservation, the assumption that adaptation and extension of recommendations will occur as the field changes is also part of the CS3DP system.

To conclude the volume, we will address a few core ideas that emerged through the process, critically evaluate the organization and processes of the community, and share some thoughts on future work. We recognize that many will not read the volume as a whole but will look to it for recommendations on specific topics covered in the chapters. With data preservation, the devil is certainly in the details, and this conclusion cannot be a useful summary of the work presented in the chapters; it is a review of select ideas, processes, and expectations.

How Are 3D Data Different?

The question of how 3D data differ from other types of data we already preserve in standardized ways came up at several points in CS3DP discussions, and it is also a common entry question from those already familiar with data preservation best practices. True, 3D data are *just* data, and many of the same preservation difficulties that occur with digital images, video, or code are shared. For example, we encounter large file sizes, many proprietary formats, interdependent file structures, post-processing decisions affecting data integrity, and the lack of an archivable physical analogue that permits all or most end uses. It seems the features that necessitate treating preservation of 3D data differently emerge from the co-occurrence of multiple data types (e.g., grid-based volumes, point-based clouds and **meshes**, and curve-based shapes) and multiple intents for end use (e.g., use to create fixed 2D visualizations, use to produce physical objects, use to allow interactive inspection, or use to produce a virtual experience). While one could argue that **workflows** could be separated such that there would be, for example, a preservation standard for the laser-scan-to-3D-print workflow, the same scan data might commonly be used in an interactive visualization. Indeed, any of the example data types might be used for any of the end uses, and standards for preservation should take this matrix of type and use into account. The preservation of 3D data may actually be most similar in scope to the preservation of physical objects (with an array of types and end uses) but without the physical forms. For physical objects, we have conservators, and for 3D data, we may also need highly specialized experts, trained in the various ways 3D data are created and consumed, to support preservation of these data.

Ideas from the Community

The concept of the 3D community has been historically too narrow and, while this book is an attempt at diversifying voices and perspectives in the development, we acknowledge the continued growth that needs to occur as the community realizes itself. Relationship building is an organic process that has played out over the number of years since CS3DP began (and indeed before that!). The work of CS3DP is moving 3D data preservation forward in the United States, and it has also made an impact internationally. The community has broken ground, planted seeds, connected people, harvested ideas, and spread them widely. We have seen evidence of this in papers, in presentations, and in social media. Most importantly, an invested community of creators, curators, and specialists has been built and is growing.

Given that CS3DP recommendations have emerged from this diverse community, we would like to highlight a few ideas that we think owe their inclusion here to the community dynamic. One of the most pervasive of these is the understanding that not all institutions or individuals can meet the requirements of "best practices" and that outlining only what is required at the highest level of preservation can be discouraging and overwhelming for those who would otherwise be interested in trying to meet preservation standards. Additionally, preservation and **access** are interdependent, and preservationists must take into account the access constraints imposed by disparity in users' computational and network resources. Therefore, the **Good/Better/Best** approach to preservation was adopted. This tiered model for preservation practices allows for flexible recommendations that are applicable at various levels of resource and time availability, aiding institutions in making practical, actionable 3D data preservation decisions. The Good/Better/Best model resonated throughout the community regardless of disciplinary focus, and it is clear to us that it applies broadly to the issues we are addressing. This resource-sensitive framework is used explicitly in chapter 3, "Management and Storage of 3D Data," and chapter 4, "Metadata Requirements for 3D Data," and it was discussed and implemented conceptually by all working groups while developing their contributions to this book.

Another idea that emerged as a result of the diversity of participant backgrounds was emphasis on the critical importance of data creators working with curators to ensure preservation. The term "**preservation intervention point**" (PIP) was used to recognize the stages in the process of 3D data capture or creation at which the creator should pause and consider documenting aspects of the process or make decisions about expectations for long-term preservation. For example, a technician using **photogrammetry** to produce a 3D model of a museum object might, after completing necessary photography, record information about the photography equipment used, the photographer's name and role (are the photos a work for hire?), and the workflow used to create the photos. Such information, which may be important for understanding how the data can later be used, are easily forgotten or overlooked if one waits to document the process at completion of the project. What level of documentation is possible is dependent on resource availability, and a Good/Better/Best approach is also needed here. Part of the preservationist's role in this PIP is making recommendations about how the information should be organized. Perhaps the most critical PIP is the planning stage. Consulting with a **data curator** when planning a project is helpful for building a data management and preservation plan that identifies specific PIPs for seamless and efficient recording of relevant information as part of the creator's workflow.

Data curators and creators each have important roles to play in ensuring preservation of 3D data. However, the complexity of data types, audiences, and uses means that finding a curator with the expertise required to properly advise on all aspects of 3D data preservation is likely impossible. Of course, this need for guidance in preserving 3D data is the very reason the CS3DP community came together. There is a crucial requirement for agreed-upon, interdisciplinary preservation practices for various 3D data types that can serve as references for curators in the process of preservation. Although this volume may not be a 3D data curator's guidebook, it provides a foundation to inform a generalized approach to 3D data preservation that is fundamental. From here, standards and practices can be extended to meet the specific needs of particular communities, but it will be important to maintain communication among specialists in particular fields to maintain core generality and prevent drift.

Assessing Our Approach (CoP)

When we surveyed the community in 2017,[1] three things were clear from the results:

1. Most institutions were not using standards for documenting and preserving 3D data.
2. Many institutions were creating their own bespoke solutions or best practices to fill the gaps.
3. Most individuals were interested in collaboratively establishing standards/best practices that would be useful broadly.

From the very beginning of formulation of CS3DP, we identified some ideals that would shape how we approached the issue of 3D data preservation. We believe that effective standards must be broadly adopted in order to serve their purpose and that community investment in the process of creating the standards increases both the likelihood of adoption and the suitability of the resulting system for use in a variety of contexts. Of course, working with a diverse group of stakeholders presents difficulties due to varying technical backgrounds, functional roles, vocabularies, and end goals, but exposing community members to such variance is essential in producing a broadly applicable, adoptable system.

The first CS3DP discussions were admittedly difficult, with lack of shared vocabulary (and an overabundance of acronyms and discipline-specific terminology) standing out as a barrier to communication on day one of our first forum event. However, the response to the difficulty was a desire for a unified glossary and an acronym list that could help individuals of a given background better understand the ideas coming from individuals of other backgrounds. The end result of these conversations is the included glossary, which leverages the expertise and perspectives of many and distills our various disciplinary and methodological approaches into a list of concepts that can be used across these intellectual boundaries. The community now has a familiar and consistent grounding from which to continue further discussion and refinement of the best practices and standards for 3D data.

While it was clear that we needed input from the full group on all of the issues related to 3D data preservation, we decided early at the first forum event to establish smaller work groups to focus on particular topics, as writing coherent text in large groups can be problematic. To ensure feedback across the work groups, each presented its work for feedback at multiple points in the process, and all groups had access to the draft documents being produced by other groups. Chapter contributions were also formally reviewed (anonymously) by members of other work groups prior to an external peer review. This system seemed to work well, and it was notable that a degree of overlap between the scope of chapters required that members of work groups keep track of what developments were emerging in other groups. Overall, the system of multiple work groups with slightly overlapping scopes producing content reviewed informally through presentations and formally through anonymous internal review was a suitable compromise between the entire group trying to produce documents together and a set of tightly scoped work groups addressing particular areas in isolation. A drawback of this system of organization was the excessive time required to integrate feedback from so many rounds of review.

Considering aspects of the work that need further development, we identify 3D Data Ethics and more international perspective. The five work groups that were initially created align with the five core chapters in this volume—best practices, management, rights, access, and metadata—and these areas were considered during recruitment of experts for the first forum event. An area that was not considered separately at that point was ethics. While ethical issues related to colonialism, social justice, accessibility, and Indigenous rights were raised in discussions, a separate work group was not created to delve into the complex area of 3D Data Ethics, largely because few experts in this area were part of the CS3DP community. In hindsight, we feel it would have been appropriate to have sought the necessary expertise and created an additional work group between the two main forum events. As standards for 3D data preservation are refined, we suggest that a deep analysis of the ethical implications of 3D data creation and preservation is required. It will also be necessary to consider preservation and access needs and concerns outside of the United States. The CS3DP community has only a small number of participants from outside the United States and Europe, which has limited our ability to recognize and address preservation issues present in other regions. This limitation was a result of granting and logistical constraints as well as the need to draw from existing networks to identify interested parties for participation. Differences in legal frameworks and access infrastructure are among areas needing study.

Communication among CS3DP participants was largely maintained following the two forum events because of member investment in producing and reviewing chapters of this volume. With this aspect of the work now complete, we find ourselves looking to next steps and continued solidarity. There have been calls for a CS3DP Phase 2, which would involve implementation of the recommendations and evaluation of that process. It is not clear how the broad community could take part in such efforts, which are likely to play out at an institutional level, but involving participants in discussions or consultations as standards are implemented seems critical to continue the sense of community ownership of the standards. The editors welcome institutions implementing recommendations found here to reach out with feedback and comments on the process.

Going Forward

We have demonstrated through this work that it is possible for a diverse, committed group of stakeholders to come together and create grassroots standards. Continuing to approach standards as a collective need, rather than on an ad hoc basis by discipline or by modality, allows the creation of protocols that can be applied broadly, and it facilitates flexibility and generality in its design. However, the chapters in this volume have established the basis for a standard for 3D data preservation that requires further development and refinement. We view the next steps in the development process as testing and refinement of the system in practice. While many of us discussed previous experiences and hypothetical scenarios when considering proposed recommendations, the practices here have not been tested together. Just as we typically see software moving through testing and refining phases (e.g., betas), the CS3DP recommendations will need to be tested and developed in an iterative fashion to ensure that they meet the specified goals and can be implemented in efficient ways.

Maintaining community involvement through this testing phase may be challenging because the implementations must happen at the level of institutions and individuals. For the flexibility of recommendations to be properly assessed, they will need to be implemented in a variety of contexts, ranging from large museums and libraries with existing repository infrastructure to individual creators and curators with minimal preservation resources. Similarly, the implementation testing should cover multiple disciplinary foci to assess generality and be scaled to a manageable number of pilots and experiments (e.g., five to ten versus 100 simultaneous tests). Some test implementers will likely be previous CS3DP contributors, but in order to truly assess applicability, implementers should be a mix of folks who have and have not previously engaged with CS3DP. Members of the CS3DP community will connect with these testers by using a guiding set of research questions to critically evaluate the process:

- How well do the Good/Better/Best recommendations support data preservation and access in diverse communities and contexts?
- What is the impact of 3D data preservation activities applied throughout the creator-to-repository pipeline?
    - How feasible is it for creators to document at proposed PIP stages?
    - What aspects of documentation or preservation might be automated to reduce workload and increase adoption?
- How does a **repository manager** integrate 3D-specific metadata elements into existing repository systems?
- How effective is a limited set of shared, interdisciplinary, 3D-specific metadata?
    - Does it actually support an inclusive ecosystem?
    - Do users value the cross-disciplinary accessibility provided by this?
- What topics related to 3D preservation need to be targeted for increased education of and outreach to creators, users, and curators to enhance adoption and compliance?

After implementation testing and analysis comes a refinement phase. Ultimately, we envision the result to be a sustainable, yet evolving 3D preservation ecosystem rooted in community design and collaboration.

Notes

1. Community Standards for 3D Data Preservation (CS3DP), "CS3DP 2017 Community Survey," accessed January 10, 2020, https://osf.io/tcn6h/.

Bibliography

Community Standards for 3D Data Preservation (CS3DP). "CS3DP 2017 Community Survey." Accessed January 10, 2020. https://osf.io/tcn6h/.

Glossary

Term	Also Known As	Definition
access		The ability and/or permission to discover, examine, retrieve, and/or reuse information (e.g., 3D data) through the use of search engines, catalogs, indexes, finding aids, documentation, or other tools.
accuracy		The proximity of a measured value to a standard or known value. High accuracy, similar to high precision, implies that the difference between the values is small. This term is commonly misused.
administrative metadata		Administrative metadata are used in managing and administering collections and information resources and can comprise both technical and preservation metadata. They are generally used for internal management of digital resources. They can include acquisition information, rights and reproduction tracking, documentation of legal access requirements, location information, selection criteria for digitization, and use analytics. Sources: NDSA, "Glossary," 2013, https://ndsa.org/glossary/; Anne J. Gilliland, "Setting the Stage," in *Introduction to Metadata*, 3rd ed., ed. Murtha Baca (Los Angeles: Getty Publications, 2016), http://www.getty.edu/publications/intrometadata/setting-the-stage/.
aggregator		An entity that is gathering assets from other institutions and making them available (i.e., sharing/disseminating assets).

Term	Also Known As	Definition
Archival Information Package (AIP)		As defined by the OAIS, an AIP is the set of content and metadata managed by a preservation repository and organized in a way that allows the repository to perform preservation services. Source: Society of American Archivists, "Open Archival Information System (OAIS)," https://www2.archivists.org/groups/standards-committee/open-archival-information-system-oais.
archive		An entity that looks to preserve assets for future use (noun); to file or collect with the intent to provide long-term access (verb).
avatar		An icon or figure representing an actor in a digital setting.
bitstream preservation	preservation	Technical measures such as checksums and redundancy that aim to ensure the bits comprising a file remain unchanged over time and after changes in technology. It is a basic prerequisite for digital long-term preservation.
born digital		We do not recommend using this term in a context as diverse as CS3DP, as its application with respect to 3D data is complicated and may be misleading.
checksum		A checksum on a file is a "digital fingerprint" whereby even the smallest change to the file will cause the checksum to change completely. Checksums are typically created using cryptographic techniques and can be generated using a range of readily available and open-source tools. It is important to note that while checksums can be used to detect if the contents of a file have changed, they do not tell you where in the file that the change has occurred. Source: Digital Preservation Coalition, "Fixity and Checksums," ***Digital Preservation Handbook***, 2nd ed., (Glasgow, UK: Digital Preservation Coalition, 2015), https://www.dpconline.org/handbook/technical-solutions-and-tools/fixity-and-checksums.
CIDOC CRM (Conceptual Reference Model)		An ISO standard semantic ontology created by the International Council of Museums (ICOM) and recently adopted by the International Federation of Libraries. The CIDOC Conceptual Reference Model (CRM) provides definitions and a formal structure for describing the implicit and explicit concepts and relationships used in cultural heritage documentation. See CIDOC CRM home page, https://cidoc-crm.org.

Term	Also Known As	Definition
computer-aided design (CAD)		Computer system to develop or optimize a design. It is often used in fields such as architecture to create 3D representations. Common output is a NURBS model.
content standards		Guidelines or rules for how elements are selected, formatted, and recorded, e.g., ***Anglo-American Cataloging Rules***, 2nd edition (AACR2), Resource Description and Access (RDA), ***Cataloging Cultural Objects*** (CCO), Describing Archives: A Content Standard (DACS).
controlled vocabulary		Domain-specific lists of allowable values for certain metadata elements. Classification schemes are often connected to a chosen vocabulary, e.g., Library of Congress Subject Headings (LCSH), Library of Congress Name Authority File (LCNAF), Virtual International Authority File (VIAF), Medical Subject Headings (MeSH), ***Art and Architecture Thesaurus*** (AAT), Getty Thesaurus of Geographic Names (TGN), Getty Union List of Artist Names (ULAN).
data curator		A person who reviews and modifies (or suggests modification for) datasets and their documentation to ensure they are preserved for long-term access and reusability.
data model		"An abstract model that organizes elements of data and standardizes how they relate to one another and to the properties of real-world entities." Source: Wikipedia, s.v. "Data model," last modified October 9, 2021, 04:38, https://en.wikipedia.org/wiki/Data_model.
descriptive metadata		Metadata that describe a digital entity for purposes such as discovery and identification (e.g., title, creator, summary, tags, etc.). Source: NDSA, "Glossary," 2013, https://ndsa.org/glossary/.
digital asset life cycle	digital curation life cycle	A conceptual model used to identify the stages required for successful curation and preservation of data from initial conceptualization through the iterative curation cycle. This model can be used to plan activities within a specific research project, organization, or consortium to ensure all necessary stages are undertaken.

Term	Also Known As	Definition
digital curation		The "active and on-going management of data through [their] lifecycle of interest and usefulness to scholarship, science, and education." Data curation enables data discovery and retrieval, maintains data quality, adds value, and provides for reuse over time through activities including authentication, archiving, management, preservation, and representation. Source: University of Illinois Urbana-Champaign School of Information Sciences, "Data Curation," https://ischool.illinois.edu/research/areas/data-curation.
digital preservation		"Series of managed activities necessary to ensure continued access to digital materials for as long as necessary." Digital preservation seeks to prevent loss of digital information from medium failures and hardware/software obsolescence. Source: Digital Preservation Coalition, "Glossary," *Digital Preservation Handbook*, 2nd ed. (Glasgow, UK: Digital Preservation Coalition, 2015), https://dpconline.org/handbook/glossary.
digital repository		System in which digital assets are deposited, stored, managed, and sometimes archived.
discovery		Identification of potentially relevant materials through the process of searching. The application of metadata and persistent identifiers is a method by which data can become discoverable.
Dissemination Information Package (DIP)		As defined by the OAIS, a DIP is distributed to a consumer by the repository in response to a request and may contain content spanning multiple AIPs. Source: Society of American Archivists, "Open Archival Information System (OAIS), https://www2.archivists.org/groups/standards-committee/open-archival-information-system-oais.
emulation	preservation	A means of overcoming technical obsolescence of hardware and software using tools and techniques for imitating obsolete systems.
fixity		The state of being unchanged, usually used in reference to a digital file or object. Fixity is usually checked by calculating, storing, and comparing a checksum value for the file or object.

Term	Also Known As	Definition
Good/Better/ Best		A continuum of practices ranging from the minimum, which includes required practices that can be accomplished with limited resources, to the maximum, which assumes greater resource availability and can ensure robust support of a digital asset through its life cycle.
image resolution		Number of pixels per fixed space in raster data.
laser scanning		The process of recording precise three-dimensional information of a real-world object or environment. This can include handheld laser line probes, ground-based terrestrial lidar, and aerial-based lidar that can be used to scan at a range of scales from very small objects to very large monuments or entire sites. Laser scanners rapidly sample or scan an object's surface, recording shape and often visual properties (intensity and/or RGB information). The information is often returned to the unit as a dense collection of precisely measured ***x***, ***y***, ***z*** points referred to as a point cloud. Source: Angie Payne, "Section 1. Introduction to the Laser Scanning Guide," in ***Laser Scanning for Archaeology: A Guide to Good Practice***, Guides to Good Practice, (Archaeology Data Service and Digital Antiquity, 2009), http://guides.archaeologydataservice.ac.uk/g2gp/LaserScan_1-2.
manifold mesh	watertight model	A hole-free mesh with no open edges. All edges have exactly two incident triangles. All vertices have exactly one cycle of incident triangles. All triangles in the mesh are consistently oriented. Source: Blender Development Forums, "manifold mesh" https://www.blender.org/forum/viewtopic.php?t=1012 (page discontinued).
mesh	3D model, polygonal model	A collection of vertices (points), edges (straight line segments connecting two vertices), and faces that describe the shape of a 3D object.
metadata		Information about data. See ***technical metadata***, ***administrative metadata***, ***descriptive metadata***, or ***preservation metadata***.
mode	3D mode	In this text, ***mode*** refers to methods of 3D generation (e.g., photogrammetry) or output. It may also refer to methods of interaction (e.g., virtual world).

Term	Also Known As	Definition
multimodal model	hybrid model, 3D model	A 3D model resulting from a combination of methods (modes), including reality-based capture (multi-sensor) and/or any other creation method.
NURBS model	3D model	Nonuniform rational basis spline—a mathematical model commonly used in computer graphics for generating and representing curves and surfaces. Source: Wikipedia, s.v. "Non-uniform rational B-spline," last modified November 9, 2021, 18:01, https://en.wikipedia.org/wiki/Non-uniform_rational_B-spline.
paradata	documentation; codebook; process data; methods documentation	Data regarding hypothesis, concept, or intent behind a model; the human intent; process behind creating the model. They may include methods used or variables. A recipe to recreate a 3D model.
persistent identifier		"A long-lasting reference to a digital resource" such as a digital object identifier (DOI) or ARK (Archival Resource Key). Source: ORCID Support, "What Are Persistent Identifiers (PIDs)?" https://support.orcid.org/hc/en-us/articles/360006971013-.
photogrammetry		"Digital photogrammetry falls under the broader category of Geomatics, and, according to the American Society for Photogrammetry and Remote Sensing, is defined as, 'the art, science and technology of obtaining reliable information about physical objects and the environment through the process of recording, measuring and interpreting photographic images' …(Salma, 1980). A simplified definition could be the extraction of three-dimensional measurements from two-dimensional data (i.e. images). Close-range (terrestrial) photogrammetry refers to the acquisition of photographs from a lesser distance than traditional aerial (or orbital) photogrammetry." Source: Adam Barnes, "Section 1. Introduction," in ***Close-Range Photogrammetry: A Guide to Good Practice***, Guides to Good Practice (Archaeology Data Service and Digital Antiquity, 2009), https://guides.archaeologydataservice.ac.uk/g2gp/Photogram_1-1.
platform migration	preservation	A means of overcoming technical obsolescence by transferring digital resources from one hardware/software generation to the next.

Term	Also Known As	Definition
point cloud		A set of 3D data points usually resulting from surface capture methods. Point clouds differ from a mesh in that points are not connected by edges or faces. May include additional information such as point color.
precision		The proximity of two or more repeated measurements to each other. Precision is achievable without accuracy. High precision, similar to high accuracy, implies that the difference between the values is small. This term is commonly misused.
preservation		See ***digital preservation***.
preservation intervention point	PIP	Point in the progression of the digital asset life cycle at which it might be advisable to stop and document, save versions of files for preservation, or take other preservation actions.
preservation metadata		The information necessary support the long-term retention and accessibility of digital content. Preservation metadata also establish the authenticity of digital content and record the chain of custody and provenance for a digital object. Source: NDSA, “Glossary,” 2013, https://ndsa.org/glossary/.
provenance		Information about entities, activities, and people involved in producing a piece of data or thing, which can be used to form assessments about its quality, reliability, or trustworthiness. It establishes the chain-of-custody information needed for users to make trust decisions about digital data.
README file		A file containing “information about other files in a directory or archive of computer software. A form of documentation, it is usually a simple plain text file” with the text README in the file name. Source: Wikipedia, s.v. “README,” last modified October 9, 2021, 13:08, https://en.wikipedia.org/wiki/README.)
realia		A library science term for physical three-dimensional objects that do not fit easily in the categories for printed material.
registration	processing step	Transforming multiple sets of 3D data to one coordinate system so they are in meaningful positions relative to each other. A manual or automated process that aligns two or more datasets (e.g., surfaces or volumes) to each other.

Term	Also Known As	Definition
render	processing step/output	Rendering is the process of generating a 2D image from a 3D model/scene by means of computer programs. Also, the image which results from rendering can be called a render. Source: Wikipedia, s.v. "Render," last modified April 1, 2021, 20:22, https://en.wikipedia.org/wiki/Render.
repository manager		A person who oversees infrastructure and handles ingestion, management, and discoverability of data in a data repository.
reprojection error		A geometric error corresponding to the image distance between a projected point and a measured point. It is used to quantify how closely an estimate of a 3D point recreates the point's true projection.
resolution		Usually refers to the level of detail in a 2D raster image or 3D voxel-based model. For a 3D mesh, ***resolution*** is not an appropriate term; instead, consider using ***point count*** or ***face count***. See also ***image resolution***.
RTI (Reflectance Transformation Imaging)		A technique for creating detailed surface information from a single camera point of view. The resulting data can be interactively relit and mathematically enhanced. Used in art conservation, rock art, the deciphering of inscriptions, and other uses where very fine surface details are needed. The format is pixel-based (i.e., 2D) but includes 3D information along with color and reflectance data—this is sometimes referred to as 2½D. Also referred to as PTM (polynomial texture mapping) .
sidecar		"Sidecar files, also known as buddy files or connected files, are computer files that store data (often metadata) which is not supported by the format of a source file." Source: Wikipedia, s.v. "sidecar file," last modified September 16, 2021, 1:39, https://en.wikipedia.org/wiki/Sidecar_file.
software emulation		The process of recreating, on current hardware, the technical environment required to view and use digital objects from earlier times. Source: David Holdsworth and Paul Wheatley, "Emulation, Preservation, and Abstraction," *RLG DigiNews* 5, no. 4 (August 15, 2001), http://worldcat.org/arcviewer/1/OCC/2007/08/08/0000070519/viewer/file990.html.

Term	Also Known As	Definition
source material		The data and reference material used to create a manually built 3D model (e.g., drawings, notes, measurements).
structured light		A method of scanning real objects or environments that relies on the distortion of projected light to calculate surface form. A known pattern (often a grid or horizontal lines) of light projected onto a surface appears distorted from perspectives other than that of the projector, and this can be used for geometric reconstruction of the surface shape.
Submission Information Package (SIP)		As defined by the OAIS, an SIP is the content and metadata received from an information producer by a preservation repository. Source: Society of American Archivists, "Open Archival Information System [OAIS]," https://www2.archivists.org/groups/standards-committee/open-archival-information-system-oais.
technical metadata		Metadata that describe the technical state of and process used to create a file. Often closely related to either its file format or the original software used to create the file, e.g., scanning equipment and settings used to create or modify a digital object. Source: NDSA, "Glossary," 2013, https://ndsa.org/glossary/.
texture map	3D model, UV map	An image file that is applied to the surface of a 3D model (often in UV space) to provide color, specularity, or other surface properties. Examples of named texture maps would include, but are not limited to, normal maps, diffuse maps, ambient occlusion maps, displacement maps, and bump maps.
3D creator		Someone who is creating 3D models of real-world or imaginary objects, visualizations, environments, or theoretical arguments.
transformation	format migration	Transformation of a file from an obsolete, unsupported, or closed format into a newer, open, or supported format.
usability		The extent to which a product can be used to achieve articulated goals with effectiveness, efficiency and satisfaction for a specific context of use. Modified from the official ISO 9241-11 definition of ***usability***. Interaction Design Foundation, "Usability," https://www.interaction-design.org/literature/topics/usability.

Term	Also Known As	Definition
UV space	3D model	The coordinates that map one or more 2D texture maps to the 3D geometry.
volumetric (capture) data		3D raster data, usually represented by a DICOM or image stack, or binary file that represents data on a regularly spaced (usually) three-dimensional grid. Each unit, voxel, can contain extra information such as opacity, density, and color. Volumetric data can be captured by various technologies such as magnetic resonance imaging (MRI) and X-ray computed tomography (CT).
voxel		The basic volume element (often cubic) in a 3-dimensional array, as a pixel is the basic area element (often square) in a 2-dimensional array.
watertight mesh	3D model, manifold	A mesh that has no holes and is completely enclosed. If you filled it with water, it would not leak. A mesh must be watertight for it to be 3D printed.
workflow		The sequence of events and processes that comprise the life cycle of a 3D work from initiation to curation.
X-ray computed tomography (CT)	CAT scan	"A CT scan [also known as] …computed tomography scan, [and] formerly known as a computerized axial tomography scan or CAT scan," makes use of computer-processed combinations of many "X-ray measurements taken from different angles …to produce tomographic (cross-sectional) images (virtual 'slices')" of specific areas of a scanned object, allowing the user to see inside the object without cutting. Source: Wikipedia, s.v. "CT scan," last modified November 28, 2021, 7:02, , https://en.wikipedia.org/wiki/CT_scan.

Bibliography

Barnes, Adam. "Section 1. Introduction." *Close-Range Photogrammetry: A Guide to Good Practice*. Guides to Good Practice. Archaeology Data Service and Digital Antiquity, 2009. https://guides.archaeologydataservice.ac.uk/g2gp/Photogram_1-1.

Blender Development Forums, "manifold mesh." https://www.blender.org/forum/viewtopic.php?t=1012 (page discontinued).

CIDOC CRM. Home page. https://cidoc-crm.org.

Digital Preservation Coalition. "Fixity and Checksums." *Digital Preservation Handbook*, 2nd ed. Glasgow, UK: Digital Preservation Coalition, 2015. https://www.dpconline.org/handbook/technical-solutions-and-tools/fixity-and-checksums.

Digital Preservation Coalition. "Glossary." *Digital Preservation Handbook*, 2nd ed. Glasgow, UK: Digital Preservation Coalition, 2015. https://dpconline.org/handbook/glossary.

Gilliland, Anne J. "Setting the Stage." In *Introduction to Metadata*, 3rd ed., ed. Murtha Baca. Los Angeles: Getty Publications, 2016. http://www.getty.edu/publications/intrometadata/setting-the-stage/.

Holdsworth, David, and Paul Wheatley. "Emulation, Preservation, and Abstraction." *RLG DigiNews* 5, no. 4 (August 15, 2001). http://worldcat.org/arcviewer/1/OCC/2007/08/08/0000070519/viewer/file990.html.

Interaction Design Foundation. "Usability." https://www.interaction-design.org/literature/topics/usability.

NDSA. "Glossary." 2013. https://ndsa.org/glossary/.

ORCID Support. "What Are Persistent Identifiers (PIDs)?" https://support.orcid.org/hc/en-us/articles/360006971013-.

Payne, Angie. "Section 1. Introduction to the Laser Scanning Guide." *Laser Scanning for Archaeology: A Guide to Good Practice*. Guides to Good Practice. Archaeology Data Service and Digital Antiquity, 2009. https://guides.archaeologydataservice.ac.uk/g2gp/LaserScan_1-2.

Society of American Archivists. "Open Archival Information System (OAIS)." https://www2.archivists.org/groups/standards-committee/open-archival-information-system-oais.

University of Illinois Urbana-Champaign School of Information Sciences. "Data Curation." https://ischool.illinois.edu/research/areas/data-curation.

About the Authors

Best Practices for 3D Data Preservation

Kristina Golubiewski-Davis is the director of the Digital Scholarship Department at the University of California, Santa Cruz. She leads the library's outward-facing efforts to support the integration of digital methods into teaching and research. Kristy received her PhD in anthropology from the University of Minnesota, where she incorporated 3D scanning of archaeological objects for research. In her position at UCSC, she leverages her experience with 3D to expand library support of 3D data across their entire life cycle, including scanning, modeling, printing, VR, and preservation.

Jessica Maisano is a research associate and the facility manager of the University of Texas High-Resolution X-ray CT Facility (UTCT). She arrived at UTCT in 2000 as a postdoc to help build DigiMorph.org (the Digital Library of Morphology) and subsequently collaborated on the NSF Assembling the Tree of Life project for lizards ("Deep Scaly"). Currently, Jessie is PI of a PEN grant partnering with OVert (the Open Vertebrate network) to upload approximately twenty years (approximately 9 TB) of UTCT scans of fossil and Recent vertebrates to MorphoSource.org and teach short courses to train the next generation of high-resolution X-ray CT data consumers. She is also the chair of the North American branch of ToScA (Tomography for Scientific Advancement).

Marcia McIntosh received her master's in information studies from the University of Texas at Austin's School of Information. She is the digital production librarian at the University of North Texas, where she assists in the coordinating, management, and hiring necessary to create digital collections.

Jennifer Moore is head of data services at Washington University in St. Louis. She leads a team of experts to address areas of digital data management, curation, analysis, and visualization. She's been working on 3D capture and preservation since 2014. Moore is a co-PI on the IMLS grant Community Standards for 3D Data Preservation (CS3DP), which brings stakeholders together to establish agreement on how to make 3D data long-lasting. She is also a member of the Data Curation Network (DCN) Project Team, which is focused on building a shared data curation model across institutions, and a co-PI on the IMLS Specialized Data Curation Workshop grant, which builds on the DCN's CURATE model to provide training and develop curation resources. https://orcid.org/0000-0001-6628-6820

Kieron Niven is a digital archivist and data standards lead at the Archaeology Data Service, University of York, UK. Since 2003 he has undertaken work on a number of large project archives, including Crossrail, the Historic Landscape Characterisation program, and the Channel Tunnel Rail Link. As data standards lead he has been responsible for the ongoing development of the ADS Guides to Good Practice through various projects, most notably the 2009–2011 Andrew W. Mellon Foundation–funded project in conjunction with Digital Antiquity and tDAR.

Will Rourk is currently the 3D data and content specialist with the Scholars' Lab at the University of Virginia Library. He has over twenty-five years' experience in 3D modeling and ten years' experience with 3D data and GIS. He has been a chief architectural consultant for the Tibetan and Himalayan Library (UVA), with which he has done field research into the preservation of historical buildings of the Tibetan capital city of Lhasa. His current work focuses on teaching and research in the department of architectural history at UVA, defining methods of 3D cultural heritage informatics in the collection, processing, preservation, and distribution of 3D data of historical objects, buildings, and sites. His work employs the use of 3D scanning technologies and photogrammetry techniques for collecting geometric and color data of historical and cultural context at various levels of scale and complexity from artifact to architecture. He specializes in 3D scan-to-print of artwork and artifacts for preservation and public access. He has been a key component in working with the UVA Library toward open access of 3D data by the scholarly community and has published work on 3D cultural heritage informatics. More information about his work can be found at https://pages.shanti.virginia.edu/Cultural_Heritage_Data/ or by following @rezn8r on Twitter.

Rebecca Snyder is the acting informatics branch chief at the National Museum of Natural History, Smithsonian Institution (SI). Rebecca is responsible for the digital stewardship and preservation of collections and research data. Recent projects include the application of persistent identifiers for SI collections data and media; the Smithsonian Open Access initiative, where she was responsible for designing the data flows between all SI collections systems; assisting in the development of a system of record for 3D data; and data quality improvement projects adhering to FAIR data principles. She is also a member of the Audubon Core data standard maintenance group, focusing on creating standards for the sharing of 3D data.

Management and Storage of 3D Data

Doug Boyer is associate professor of evolutionary anthropology and director of MorphoSource at Duke University.

Rachel Fernandez is the digital preservation program manager at the Center for Digital Antiquity, which runs the Digital Archaeological Record (tDAR), an international repository for digital archaeological data. She assists in the development, implementation, and management of digital curation projects hosted within tDAR. She also provides workshops and training in data management to archaeologists and cultural resource managers in the US and abroad. Rachel holds a master's in classical archaeology from the University of Colorado Boulder and has conducted fieldwork on several archaeological excavations and surveys throughout the Mediterranean. https://orcid.org/0000-0002-7697-4149

Monique Lassere is the inaugural digital archivist at Houghton Library, Harvard's rare books and manuscripts, literary and performing arts archives. There she stewards born-digital archival materials in the form of manuscript collections, 3D objects, and web archives. Previously she was the digital preservation librarian at the University of Arizona in Tucson. Monique earned her MLIS in 2017 at the University of Illinois at Urbana-Champaign.

Marcia McIntosh [see Best Practices for 3D Data Preservation]

Jennifer Moore [see Best Practices for 3D Data Preservation]

Francis P. McManamon, PhD, is a faculty member in the MA program in cultural heritage management at Johns Hopkins University (JHU) and adjunct research professor in the School of Human Evolution, Arizona State University (ASU). He is teaching online courses in cultural heritage management at JHU. From 2009 to 2019, he was the executive and founding director of the Center for Digital Antiquity at ASU, an organization devoted to broadening and improving the long-term preservation and ease of access to archaeological information. Before Digital Antiquity, he served as the chief archeologist of the National Park Service (NPS; 1995–2009) and the departmental consulting archaeologist of the Department of the Interior (1991–2009). Earlier in his career Frank served as regional chief of cultural resources and regional archaeologist for the North Atlantic Region of the NPS and staff archaeologist for the Massachusetts Historical Commission. He has worked on a variety of curation issues for physical and digital collections.

Albert Rozo is a digital preservation specialist at Penn State University.

Todd P. Swanson is assistant director, head of Getty Digital Imaging at the J. Paul Getty Trust and formerly digitization lead at the Walt Disney Archives.

Kate Webbink is an information systems specialist in the technology department at the Field Museum of Natural History in Chicago. More info about Kate can be found at https://orcid.org/0000-0002-8347-0942.

Metadata Requirements for 3D Data

Jon Blundell is a 3D program officer at the Smithsonian Institution, where he focuses on the technical challenges of the department, developing workflows and IT infrastructure to support 3D capture, processing, data management, and delivery to the public. When he's not uploading the Smithsonian's collection to the Matrix, he can be found playing pinball and tabletop games.

Jasmine L. Clark is the digital scholarship librarian at Temple University. Her primary areas of research are accessibility and metadata in emerging technology. Currently, she is leading the Virtual Blockson, a project to recreate the Charles L. Blockson Afro-American Collection in virtual reality, while also doing research in 3D metadata and the development of Section 508–compliant guidelines for virtual reality experiences. Jasmine has experience in a variety of functional areas and departments, including metadata, archives, digital scholarship, and communications and development. She is interested in the ways information organizations can integrate accessible, inclusive practices into their services, hiring, and management practices.

Katherine E. [Katie] DeVet, PhD, is the section supervisor for document delivery in the Texas Tech University Libraries. Holding a PhD in fine arts from Texas Tech University, Katie uses her diverse experiences to help colleagues and scholars in all disciplines find the resources they need for their research and use the technology available to assist with interdisciplinary collaboration across the world. Within the department, she works to guide document delivery, SHAPES, and course reserves staff and student assistants to give excellent service to all patrons, both at Texas Tech and abroad. Through the SHAPES program (Sharing and Helping Academics Prepare for Educational Success), Katie and her colleagues are working to bring 3D data and models to the classroom, facilitating conversation and collaboration between 3D practitioners and non-practitioners, with an eye toward eventual interlibrary loan of 3D objects.

Juliet L. Hardesty is the metadata analyst at Indiana University Libraries. She establishes standards and requirements for discoverability, access, and sharing of digital collections held and managed by the IU Libraries and works to ensure that these collections are preserved and will remain usable into the future. Her research interests include bias and inclusivity in metadata, metadata standards for 3D digital objects, and migration of metadata from XML (Extensible Markup Language) into RDF (Resource Description Framework) and linked data.

Copyright and Legal Issues Surrounding 3D Data

Andrea D'Andrea took part in the archaeological explorations of many Italian sites (Herculaneum, Pompeii, Paestum), taking care of the digital surveys and the coordination of the computerization of the archaeological documentation and the implementation of the relative GIS. Since 2006 he has been a member of the research groups of the archaeological missions of the University of Naples L'Orientale in Egypt, focusing in particular on topographical surveys and detailed surveys with laser scanning and photogrammetry technique. He has also participated in the archaeological missions in Cyprus, Saudi Arabia, Ethiopia, Jordan, Turkey, and Yemen. Since 2010 he is director of the archaeological investigations at Sun Temple of Niuserra at Abou Gurab (Egypt) granted by Italian Ministry of Foreign Affairs, and in 2012–2015 he was one of the coordinators of the European Project 3D ICONS (ICT Policy Support Programme, project no. 297194). More information about Andrea can be found at https://orcid.org/0000-0002-5274-0786.

Michael Conyers is the digital archivist for the Arc/k Project, a nonprofit organization dedicated to empowering communities and individuals to digitally preserve their own endangered cultural heritage objects and locations in 3D and VR around the world; most recently in Venezuela and Syria. He holds a BA in history and an MLIS degree from the University of Washington, is an active member of the Society of American Archivists, and currently serves on the steering committee of the Seattle Area Archivists. Michael is a working advocate for progressive ethics that acknowledge and respect the moral rights of Indigenous communities to control and preserve their own digitized cultural heritage.

Kyle K. Courtney is the copyright advisor at Harvard University, working out of Harvard Library's Office for Scholarly Communication. He works closely with Harvard University to establish a culture of shared understanding of copyright issues among Harvard staff, faculty, and students. His Copyright First Responders initiative is in its sixth year and has spread beyond Harvard to reach libraries, archives, museums, and cultural institutions in Alaska, Arizona, California, Colorado, Massachusetts, New Hampshire, Oregon, Rhode Island, and Washington. He presently teaches two sections of Legal Research as part of the first-year legal research and writing program at Harvard Law School. He also has a dual appointment at Northeastern University, teaching Cyberlaw: Privacy Ethics and Digital Rights and the Legal Writing Workshop. Kyle holds a JD with distinction in intellectual property law and an MSLIS. He is a published author and a nationally recognized speaker on the topic of copyright, technology, cultural institutions, and the law.

Emily G. Finch currently serves as Kansas State University's scholarly communication and copyright librarian. She is passionate about history, the humanities, and the law as it pertains to publishing, intellectual property, and cultural heritage. Emily Finch holds a BA from Kalamazoo College in English and history with a minor in political science and concentration in American studies and an MSI from the University of Michigan School of Information, where she specialized in libraries, archives, and digital curation. She is also in the process of completing the requirements for a graduate certificate in museum studies from the University of Michigan Rackham Graduate School.

Melissa Levine is the director of the Copyright Office at the University of Michigan Library, where she provides expertise to support research, scholarship, education, and expression. She holds a BA in art history and history from Emory University and received her JD from the University of Miami School of Law. She is a member of the steering committee for the U-M Museum Studies Certificate program, is a lecturer for the U-M School of Information, and has taught for Johns Hopkins University's master in museum studies online program. More information about Melissa can be found at https://orcid.org/0000-0003-1104-2911.

Nicole Meyer is the corporate archivist for Morphosis Architects, an international architecture and planning firm recognized as an industry leader in BIM (building information modeling) and digital design technologies. She holds a BA in art history from the University of Colorado Boulder and an MLIS from the University of California Los Angeles, where her research focused on archiving and preservation concerns for born-digital design records and contemporary architectural practice.

Adam Rountrey, PhD is a research museum collection manager and 3D specialist at the University of Michigan Museum of Paleontology. He has been involved with acquisition, analysis, visualization, preservation, and dissemination of 3D specimen data at this institution since 2004. During this time, Adam developed the photogrammetry workflows and 3D web viewer for the University of Michigan Online Repository of Fossils, and he currently manages the online repository. He is a co-PI on the IMLS-funded Community Standards for 3D Data Preservation project and is particularly interested in issues related to rights and ownership of 3D data in museum settings. https://orcid.org/0000-0003-0939-9102

Hannah Scates Kettler leads digital scholarship projects from inception to preservation, managing the process of creation as well as providing research and development support as the head of digital scholarship and initiatives at Iowa State University. She is active in concerns regarding 3D creation and preservation, diverse representations in cultural heritage collections, and digital humanities. Scates Kettler holds a BA from the University of Iowa in anthropology with minors in art history and classics. She also holds an MA from King's College London in digital humanities, where she specialized in virtual cultural heritage. She is co-PI on the IMLS-funded Community Standards for 3D Data Preservation project. https://orcid.org/0000-0001-7706-713X

Kate Webbink [see Management and Storage of 3D Data]

Ann Baird Whiteside is the librarian/assistant dean for information services at Harvard University and the PI on Building for Tomorrow.

Accessing 3D Data

Francesca Albrezzi works as a digital research consultant at UCLA's Office of Advanced Research Computing (OARC) and teaches for the digital humanities program and the World Arts and Culture/Dance department.

John Bonnett is an associate professor of history and a former Canada research chair in digital humanities at Brock University.

Tassie Gniady is the team lead for digital exhibits and data management in the IU3D, a part of research technologies at Indiana University. For more, see a list of modalities and projects she has worked on at https://iu3d.sitehost.iu.edu/iu3d/items/show/57.

Heather Richards-Rissetto is an associate professor of anthropology in the School of Global Integrative Studies (SGIS) and faculty fellow at the Center for Digital Research in the Humanities (CDRH) at the University of Nebraska-Lincoln.

Lisa M. Snyder is UCLA's Office of Advanced Research Computing (OARC) director of campus research initiatives and acting director of the Research Technology Group (RTG).