3D PRESERVATION MODELS AND MODALITIES: ADVANCING RESEARCH REPRODUCIBILITY AND CAPACITY AT VIRGINIA TECH

Alex Kinnaman
Virginia Tech
United States
alexk93@vt.edu
0000-0001-8943-8946

Maureen Saverot
Virginia Tech
United States
mesaverot@vt.edu

Yinlin Chen
Virginia Tech
United States
ylchen@vt.edu

Kara Long
Virginia Tech
United States
karal@vt.edu

Nicholas Polys
Virginia Tech
United States
npolys@vt.edu

Sterling Nesbit
Virginia Tech
United States
sjn2104@vt.edu

Michelle Stocker
Virginia Tech
United States
stockerm@vt.edu

Shuhai Xiao
Virginia Tech
United States
xiao@vt.edu

Nathan Hall
Virginia Tech
United States
nfhall@vt.edu

Abstract – This paper presents findings from a multi-method study in archiving, accessing, and preserving 3D data. Specifically, we seek to develop a reusable and robust model for preserving and accessing 3D data. We further seek to identify links between digital triage, preservation actions, and archiving in multiple disciplines in order to make recommendations to embed data stewardship processes in research ecosystems. This work happens through interdisciplinary collaboration to validate curation practices across fields and disciplines.

Numerous disciplines and sectors are producing 3D data for research and instruction, but without the guidance of published standards or practices generalizable across fields. This paper will present our findings from identifying researcher needs, preservation and access requirements, and metadata models for our 3D models of specimens from the Virginia Tech Entomology Department. Our outcomes are an evaluation of the output using quantitative and qualitative methods, a working 3D metadata schema for access and preservation, and an access platform for our 3D models.

Keywords – digital preservation; 3D curation; 3D access; 3D metadata; 3D modeling

Conference Topics – Exploring the New Horizons; Building the Capacity & Capability

I. INTRODUCTION

Numerous disciplines and sectors are producing 3D data for research and instruction, but without the guidance of published standards or practices generalizable across fields. While some methods for curating 3D data have emerged, and while there are data sets available for specific disciplines, this is a dynamic area of study and there are few practices that are widely used across disciplines and prepare 3D content for long-term preservation. There are some national initiatives addressing this issue, including the IMLS supported projects CS3DP,[1]

[1] CS3DP: https://osf.io/ewt2h/

17th International Conference on Digital Preservation
iPRES 2021, Beijing,China.

DOI: 10.1145/nnnnnnn.nnnnnnn

LIB3DVR,[2] and Building for Tomorrow.[3] Guidance and outcomes from these initiatives is based on broad community input across many disciplines, however the guidance has not been validated through test cases. The goal of this paper is to provide real applications of 3D curation using some of the strategies offered in those initiatives, and our own case studies and survey in order to evaluate practices and identify potential areas for improvement.

II. PROBLEM STATEMENT

To contextualize this paper, accepted data curation models and potential metadata standards which may be applied to 3D data and metadata were explored.

A. Unique Challenges of 3D Preservation

3D data sets present several challenges. They are often created using proprietary software with opaque algorithms which inhibit reproducibility in research. Furthermore, some of the most widely adopted file formats used for display (such as STL and OBJ) do not meet some of the criteria for archival standards. For example, the STL is a proprietary format, and while OBJ is an open format, it is not under active development. X3D is an open format with an active development community, and is supported by an ISO standard, however it does not have a large user community.

Virtual environment data is distinct from 3D data in that instead of an object in a vacant space, the entity is a space with objects that can be interactive--doors can open, for example--which is not possible in PLY, STL, or OBJ files. This interactive element in virtual space is handled differently in X3D than in 3D gaming systems, which use declarations.

3D data sets may be produced with numerable and combined methods, using various software and equipment. The creation method determines the digitization workflow, provenance information, and technical metadata. The objects being digitized determine the descriptive metadata and context. These factors affect the ability to have a single standard or best practice for 3D curation.

B. Approaches to Data Management and Digital Curation

There are several approaches to the management and curation of digital content. For example, The FAIR Principles for Data [1] are guidelines for increasing the Findability, Accessibility, Interoperability, and Reusability of data and its metadata, as well as the system's infrastructure. This method applies a license to the data and increases the discoverability of the data, but is not as applicable to data that is not open, such as preliminary datasets or non-anonymized datasets.

Conversely, the Smithsonian Institute's Digital Program Office manages the Smithsonian 3D Metadata Model [2] as a granular tool specifically for 3D preservation and stewardship. The model is still evolving and is designed to include all documentation to describe a 3D capture event. The 3D Metadata Model is more robust and granular than anything else we have found and is more granular than the level we are examining, but we consider it aspirational for our program.

The Open Archival Information System (OAIS) Reference Model [3] addresses six major components of long-term curation and digital preservation processes; ingest, archival storage, data management, preservation planning, access, and administration. These components cover all of the general technical and organizational aspects needed to plan for and implement a robust data management system. The model itself is "necessarily vague" [4] so it can be used at higher levels rather than as a specific standard. OAIS is a popular model in the digital preservation community, but not widely applied by users and creators of 3D data.

Other individual 3D projects were also reviewed, such as Open Data and Digital Morphology [5] in which the authors recommend practices for what to include in 3D and VR data. In addition to technical metadata about the hardware and software used to create the model, the authors recommend which images and processes to include in a data deposit. Another project, MayaArch3D Project [6] in which the authors explore methods for making 3D archaeology more sustainable. This and other relevant work, notably what has been accomplished with the 3D

[2] LIB3DVR: www.lib3dvr.org

[3] Building for Tomorrow: https://projects.iq.harvard.edu/buildingtomorrow

Reconstruction from multi-view images of a Granary Weevil from Australia's Commonwealth Scientific and Industrial Research Organisation (CSIRO),[4] and MorphoSource [7], hosted at Duke University, offer additional context and recommendations to consider.

III. Methods

We used a multi-methods approach, with data from a multi-case study, and a web survey conducted over the past year. The multi-case study used datasets unique to Virginia Tech. The web-based survey investigates what methods the greater 3D community and digital preservation community use in 3D curation.

A. Case Studies

This multi-case study uses data sets created for research and education. The study uses three data samples, each created with different modalities, and these preliminary results [8] were presented at the 2019 International Data Curation Conference[5] in Dublin, Ireland. These participants are listed as co-authors due to this use of their research data and because of this mutual exploration of archival workflows. As an initial study, our approach to the topic was to describe the curation activities we complete on these data sets and then provide reflections to inform the work of the wider research data management community through the lessons that we learn.

The first data set in this study is photogrammetry of entomology specimens from the Virginia Tech Insect Collection. For each model, the 3D artist captures several hundred images with a D-SLR camera, and then imports the images into Agisoft Metashape, which generates point clouds and a heavy mesh model. The artist lightens the model in Autodesk Maya, and restores detail in ZBrush, and re-imports into Metashape to re-apply the color to the modified geometry of the model. There are several lossy derivative steps and files in this process, and many of the steps use algorithms within proprietary software as well as subjective decisions of the 3D artist. The artist preserves each step, however, so that every decision can be reversed. Multiple steps must be preserved so that future researchers can determine the validity of measurements made on the 3D model as an accurate proxy for the physical specimen so that any future research may be done with greater scientific transparency.

The second data set in this study is high-resolution computed tomographic (CT) data of a skull fragment of the species Parringtonia gracilis. The technician scanned the physical specimen in two parts on an Xradia and stitched using the Xradia plugin. The technician then imports the data into Mimics 17.0 and 19.0, and manually segmented parts of the anatomy with the lasso tool for each slice on a Wacom Cintiq 24HD pen display.

The third data set, called 3D Blacksburg, is a virtual environment replicating Blacksburg, Virginia, pieced together from a variety of data sets over time. It is an example of data re-use since different features and aspects of the data were not collected for the purpose of developing this resource but were collected and saved for other purposes. Building heights were derived from LIDAR (Light Detection and Ranging) data. Building footprints were derived from campus and municipal GIS data and satellite imagery.

Each researcher in the multi-case study described their production and storage process, their experiences with formats and softwares, and their common challenges. The researchers also opened an unfamiliar data set and provided feedback on their successes and challenges. The case studies were practitioner-oriented and relied on narrative feedback from the researcher on their process for locating, opening, and manipulating the unfamiliar data sets. This feedback is the basis of our findings and recommendations for improvements and guided the development of questions for the web survey.

B. Web Survey

This web-based survey (Virginia Tech IRB #20-479) consists of thirteen questions designed to explore preservation workflows for 3D content, including creation, description, management, and preservation activities. The survey also provides a guided summarization of curation workflows. We

[4] CSIRO: https://data.csiro.au/dap/landingpage?pid=csiro%3A8309

[5] IDCC 2019 programme: https://www.dcc.ac.uk/events/idcc20/programme

distributed the survey to known listservs in the communities of digital preservation, digitization, 3D modeling, and digital curation, as well as on social media and as part of several presentations in relevant conferences. No questions required an answer to continue or participate, and most questions were open-ended. We designed each question to understand our participants, gain information on workflows and documentation strategies, and define common problems in the community.

The questions in the survey included:

1) What is your field?
2) What file format(s) are you producing/have you produced?
3) What purpose are these file formats for? Check all that apply. - Selected Choice
4) What is the size of your project? Please indicate this by the total data volume, number of 3D objects, how complicated the objects are in terms of files per 3D object.
    a) Please note any dependencies between these objects.
5) What software is used to capture, process, edit, and access your 3D models?
    a) What are the file specifications used for these datasets?
    b) Is this information included in the preservation package?
6) What camera rig and other equipment are used in the capture of 3D objects?
    a) Is this information included in the preservation package?
7) What methods of capture are you using? (e.g. photogrammetry, virtual environment rendering, CT scan, lidar, structured light, laser, etc.)
8) Please describe the essential steps in your digitization process.
9) Are your essential steps exclusive to your field or capture method?
10) Please describe your documentation process during the digitization workflow.
11) What kind of metadata are you collecting?
    a) Is it an existing schema, or a custom schema? Please provide a link if appropriate.
12) What are your common problems or bottlenecks in preserving 3D objects?
13) Please provide any other comments you would like to share.

After the survey closed, survey results were anonymized, and each question's responses were condensed into broader topics for analysis.

C. *Limitations*

The multi-case study data includes samples of convenience. They are not generally representative of all 3D data sets from all disciplines, but we tried to mitigate this by choosing three different types of data which were each captured with different modalities. We chose these disciplines because we know we must serve them, but we ultimately want to serve other disciplines as well. Another limitation is that this study only uses data that is open and non-sensitive. We did not consider data modeling for sensitive data subject to export control, HIPAA, FERPA, NIH Common Rule, or other protocols that restrict access. Finally, this is an initial study and the participants are all Virginia Tech faculty, which limits the validity and generalizability of the results.

The web-based survey only garnered sixteen responses after three months of distribution. We believe this is due to the field of 3D/VR being small and because the general lack of standards in 3D curation could contribute to a lack of guidance in project planning. However, the community responses helped guide our understanding. The survey also included several detailed responses on equipment and digitization workflows, as well as a range of expected common problems and limitations. Broadening the targeted audience and identifying specific individuals may have been more successful in increasing participation.

IV. Results

A. *Case Studies*

The case studies revealed several common threads, particularly the risks and threats to 3D/VR data:

1) Software obsolescence is a primary issue. Most older datasets either need to be redone or are generally unusable in the long-term, with few exceptions.
2) Large files can be difficult to manage and to open, and information can be lost if files are opened in software that is not its creation software.
    a) We experienced this when our 3D artists traded datasets; the X3D viewer we developed didn't automatically apply a specific map to a dataset and when it opened an

entire layer was missing and had to be manually replaced.

3) File format support and updates for OBJ, one of the most popular file formats, ceased approximately 15 years ago when it's creation company Wavefront Technologies was bought out in 1995.[6] This format is popular because there are no repercussions to using it and it has sustained due to this popular use and community support, but it is also a fragile format with built-in metadata that is not human readable. However, it is very difficult to develop a workflow that completely omits OBJ files in the entire process, and many projects rely on OBJ at some point even if it is not the final file format.
4) There is a heavy reliance on proprietary software and hardware, which makes for an attractive 3D dataset, but can be difficult to duplicate workflows.
    a) One example is the inability to extract metadata or track metadata and provenance within much of the software, so metadata is collected manually.
5) There is a lack of standards for creation and preservation that is reasonable for large-scale projects and for interdisciplinary use. Individual workflows largely rely on the method of 3D modeling, the object being modeled, and the purpose of the model, which further complicates the idea of a single best practice.
6) Limited to no documentation on provenance can seriously inhibit reproducibility of a dataset or the ability to open the dataset correctly.

B. *Web Survey*

The survey results were widely varied depending on the question. Below is a brief overview of the response to each question. The full and anonymized dataset is publicly available in the Virginia Tech Data Repository [9].

1) Field of study
   The top three fields are Information Technology (19%), 3D production (19%), and Libraries (19%). Other fields ranged from Palaeontology, Biology, Geology, Geosciences, Entomology, Archaeology, Evolution Biology, and Immersive Experiences. These responses are not surprising given the communities the survey targeted and the nature of the questions.
2) File formats
   Multiple file formats are necessary to create a 3D object. Participants were asked to provide all of the file types they used. Participants responded with one to a dozen formats. There were 55 unique file formats submitted. Of those formats, the top 3 were OBJ (22%), STL (13%), and FBX (7%). Other top formats included JPG, TIF, X3D, DAE, GLTF, PLY, and XYZ. Most of the file formats were expected and commonly used in 3D modeling.
   The use of proprietary versus non-proprietary is mixed. 60% of the file formats are non-proprietary, and 40% are proprietary. However, of the top 3 formats, OBJ and FBX are proprietary, and STL is non-proprietary. It is worth noting that OBJ is only legally proprietary,but has been openly documented and acts as a non-proprietary format.
3) File format use
   Participants were asked if their file formats are for access, preservation, both, or another purpose. 81% responded with "both"; 13% responded with "access"; and 6% responded with "other." This indicates that there is little no difference between access and preservation file formats, reinforcing the lack of standards in 3D/VR creation. The "other" option was included to note any additional file formats were created throughout the workflow, such as combining OBJ and MTL into the final X3D file.
4) Project sizes
   Participants were asked the project size in terms of total data volume, number of 3D objects, and complexity of the objects. This open-ended question garnered a variety of answers ranging from the number of files, size of the files, number of photographs taken, and broader answers such as "small."

[6] Library of Congress. (2020). Wavefront OBJ File Format: https://www.loc.gov/preservation/digital/formats/fdd/fdd000507.shtml

In general project sizes ranged from small to large, including 50GB to 1-2TB with hundreds to thousands of scans/photos total.

5) Creation/editing software
Multiple softwares may be used to create and edit 3D objects. There were 49 unique softwares submitted. Of these, the top software with more than 4 participants were Blender (44%); Agisoft Metashape (31%); and Zbrush, Artec Studio 15, Meshlab, and Reality Capture (25%). All of the software reported by the participants is expected and commonly used in 3D modeling.

6) Camera Rig and Capture Equipment
Capture equipment ranged from smartphones and tablets with related applications, to professional equipment and camera rigs. DSLR and some sort of Artec software (Leo, Spider, Space, Eva, Studio) were the most used (25%), followed by a mix of laser scanning, structured light, smartphones (13%), as well as a variety of scanners, cameras, and applications. Two participants also noted that they do not perform any capture and instead focus on processing.
When asked if this information is included in a preservation package, 25% responded "Yes," 19% responded with "Not Applicable" or wrote in "Not Usually," and the rest did not respond at all. This may indicate that a preservation package is not being created for most content. This does not imply that no preservation action is being taken on the content, but perhaps that preparation specifically for preservation purposes may not be a priority.

7) Capture Method
There were 13 unique responses for capture methods, significantly less than capture software. The majority of respondents use photogrammetry (63%), followed by structured light and lidar (25%), and virtual environment rendering (19%).

8) Digitization process
56% of the respondents included some description of their digitization process, ranging from a few words on capturing and modeling, to several different detailed workflows. 58% of respondents noted processing is a primary step, and 42% noted assessment of the projects as an initial step. A common thread for many workflows is that equipment choices and settings are dependent on the project and may vary based on the content, the method of capture, and purpose of the model. However, most equipment choices stay the same for each modality.

9) Essential steps
Only 6 participants answered this question of the 9 who offered their digitization process. 50% responded "No;" and 17% responded "Yes," "I don't think so," and "want to use open source where possible" each. The phrasing of this question may have been too specific for many participants to answer confidently.

10) Documentation Process
Participants did not have a common method for documenting their digitization process. Many documented some metadata, the scanner settings, and the original photos of the object. Most documentation were personal notes, task tracking, or not applicable. This is not surprising as standards do not exist, and documentation can take many forms.

11) Metadata Collection
The majority of participants are collecting metadata. Descriptive and technical metadata were the most commonly collected, dependent on the field and purpose of the content. When describing the technical metadata, respondents noted aspects like file characteristics, equipment settings, and capture dataset characteristics like file size and number of files. A few respondents also noted process metadata to track tasks and provenance. This was particularly interesting because various aspects are being recorded overall but very few have a comprehensive metadata schema including descriptive, technical, provenance, and preservation in one.

12) Metadata Schema
When asked if the metadata they collected was adherent to an existing or custom schema, 19% have a custom schema, and if they were collecting metadata, used existing schema including PREMIS, CSIRO, MODS, and DublinCore.

13) Common Problems

The top common problems and bottlenecks the participants noted were the long processing time, large files and storage, and quality. Time commitment in general was a common obstacle, from processing time, quality control review time, and organizing large numbers of files.

The results of the survey overall lend to the idea that there is a lack of best practices and standards for creating, manipulating, and preserving 3D content, but also help reinforce the need for identifying a preservation workflow and appropriate metadata schema for 3D content.

V. Interventions/Design Considerations in Practice

Through the case study interviews and test cases and the wider survey, we gained valuable information that has informed our decisions on what metadata and steps we will preserve in this project. Our final file format, metadata schema, strategy for collecting provenance and technical information, and documenting project context have been further defined. The project to which we are applying these design considerations is the photogrammetry of entomology specimens from the Virginia Tech Insect Collection, and will expand to a broader 3D curation program.

A. 3D Digitization Workflow

We are using eXtensible 3D (X3D) as the preservation file format. X3D is ISO Standard 19775-1: 2013 [10]. It is non-proprietary and has a dedicated user community supporting it. Human-readable metadata can be stored in the file itself. It is also possible to generate a smaller file size by consolidating less stable file formats like MTL and OBJ into an X3D. This decision alone helps to mitigate some of the risks discussed, such as avoiding as much proprietary reliance as we can and minimizing file size, and also providing an international standard we can rely on long-term. We have noted the issue that X3D is very difficult to open in multiple softwares. The Virginia Tech Digital Library Platform (VTDLP) repository will have an X3D viewer embedded into the repository. We also provide examples of software able to open X3D in the collection's Permissions statement.

Our team selected photogrammetry as the most accessible 3D scanning approach for these insects. DSLR cameras are widely available compared to scanning devices and photogrammetry processing softwares are improving quickly over time. If the images are captured and preserved properly, the photo sets can continue to be processed to reproduce 3D models as technologies evolve. To capture an object for photogrammetry, hundreds of photographs are taken of the subject from every possible angle, usually with the aid of a turntable. For each of these insects, approximately 365 images are taken in 5 rotations of 73 images. After capture the backgrounds are masked away and they are processed through photogrammetry software, Agisoft Metashape in our case, to create a mesh as an OBJ file. The mesh data at this stage is rough with extraneous or missing geometry and messy topology. To ameliorate this, we retopologize each mesh in Autodesk Maya to create a cleaner, lighter model that is more accessible and functional for use on the web and in 3D environments. Detail from the original model is then projected on the retopologized model to create a high-resolution mesh that has clean geometry and all the intricacies of the insect. We export a variety of models for user access: the original model created in Metashape, then high resolution (millions of polygons), medium resolution (hundreds of thousands of polygons), and low resolution (under 100,000 polygons) models. Once the models and derivatives are complete, textures and X3D models are created for each mesh in Metashape.

Moving files back and forth between different softwares, while not the most fluid workflow, ensures that we are extracting and preserving as much data as possible from the initial photographs. Each step is key to maintaining the tiny, authentic details of each insect. However, there are instances where data is lost during the photogrammetry process, for example: the very fine antennae or limbs of an insect might not hold enough data through images to appear in the mesh. In these situations, the geometry is restored, somewhat subjectively, by the artist creating the retopologized model. This discrepancy might be avoided with improvement in future technologies, such as scAnt [11], that are optimized for capturing these fine details.

Our meshes are stored and manipulated as OBJ files at each stage until the process is complete and they can be exported as X3D. Unfortunately, many standard 3D softwares do not yet support X3D but we hope to convert our workflow to entirely open-source software, like Meshroom and Blender, that will import X3D files in the future. X3D has an active and supportive development base and is easily readable for humans, making it an excellent file format for our preservation purposes. However, OBJ as a filetype is so pervasive and widely used, despite its proprietary nature and lack of active development, that some of the most popular softwares do not offer X3D import support. By making our standards for 3D preservation accessible and thorough, we hope to encourage X3D support across more 3D softwares.

B. *Documentation and Metadata*

Task tracking in a shared spreadsheet occurs from specimen intake from the Entomology Department and continues through capture and modeling. Initial provenance metadata is documented in this spreadsheet, as well as object-level technical metadata.

The preservation metadata schema is a combination of select elements from the Smithsonian 3D Metadata Model, DarwinCore, DublinCore, and PREMIS to encompass both what we believe is important and what we have the capacity to collect on a wide scale. The Department of Entomology faculty member collaborator already collected descriptive metadata aligning with DarwinCore. Our access and preservation system was built with DublinCore terms, so we collaborated with our Digital Libraries team and the Metadata Coordinator to incorporate DarwinCore into our Digital Libraries Platform for increased searchability and discoverability of the specimens.

Technical metadata is useful for reproducing the exact model. The 3D Insect metadata is organized by select elements from the Smithsonian 3D Metadata Model. This model contains 130 elements, many of which are still undergoing development. The schema we developed consists of specific elements selected from this model that relate directly to project descriptions, photogrammetry, and provenance. This particular model is tailored to photogrammetry, but adaptable for future projects that may have different capture methods or modalities. The technical metadata is highly detailed and may not be useful to the average user, but we want our data and metadata to be as open as possible.

We opted to create a JSON file with a README file to allow the technical metadata to be available for each object, but without the added complication of incorporating such a schema into our system. A preliminary JSON file containing the combined DublinCore, DarwinCore, and Smithsonian 3D Metadata Model can be viewed online.[7]

Provenance is vital to understanding how a dataset or model was created, so we are recording provenance information as granularly as we can, which appears in both the DarwinCore and the Smithsonian elements where appropriate. This outlines the entire lifecycle of the physical specimens transforming to 3D models and identifies the date and person associated with a particular stage of development.

Finally, it's important to understand that, like other digital content, the purpose of the object determines what information is necessary for understanding and reproducibility. Our insects do not require an entire Smithsonian 3D Metadata Model to be successfully accessible in 10 years; our insects will also not require the same metadata as a CT scan of a bone or a virtual map, not only because of differences in discipline, but because of difference in technique. In choosing multiple metadata schemas that best suit our content for its unique descriptive and technical needs, we were able to fully capture our 3D content's integrity for both the end users and for preservation purposes.

C. *Preservation Criteria*

In addition to collecting provenance and technical granular information, our strategy for preserving these 3D objects is reconstruction. 3D creation and opening software evolves quickly so there is no guarantee that a software we predict will be able to open a dataset accurately in the medium to long-term will be successful. The goal of supporting reconstruction involves creating folders for each object containing their original raw images,

[7] Entomo3D preliminary JSON file: https://drive.google.com/file/d/1JG9DL7ugnrpWkO3u1xNAKQjczpcoDU_z/view

descriptive metadata, technical metadata (which includes hardware and software for capture, processing, and access) and masks. The mask is a small PNG file with just an alpha channel that isolates the object from the background. Using masks speeds up processing time and results in a cleaner model. This package will be ingested into our preservation system and should support the recreation of the current model in the short-term, and the reconstruction of a model in the long-term.

PREMIS metadata will also be applied per the Virginia Tech Digital Libraries Platform Preservation Events Audit Policy.[8]

D. 3D Models Presentation

In order to provide users with an interactive 3D model viewing experience on the Web, we use Extensible 3D (X3D) open standard to represent our 3D model datasets. The most challenging task we have faced is the file size of the X3D. Originally the X3D files we created were 200-300 MB, which is too large for end-users to view through a Web browser. Thus, we need to find a balance between the size of the X3D file and the acceptable resolution to present our 3D models. Currently, we use low resolution for the fastest upload speed, and Agisoft Metashape to generate a very light X3D file so we can generate each file at around 20 MB. Finally, we will display our 3D collections through the VTDLP. This platform is a serverless architecture and is built entirely in AWS. We utilize the edge computing provided by AWS with our open-source software to provide a low lentancy, fast, and highly responsive website. Users will be able to view and interact with the 3D models and see the detailed metadata information.

VI. CONCLUSION

This paper explored the case studies performed on evaluating and preserving 3D content at Virginia Tech and the results of a web survey distributed to the digital preservation and 3D community. The findings shaped specific interventions and design considerations we applied to the 3D models derived from the Entomology Department Insect Collection. The result is a customized, preservation-centric metadata schema designed for reproduction or recreation of our 3D objects.

The immediate next step for modeling the Virginia Tech Entomology Department specimens is to finalize our documentation and implement the metadata schema. We will be sharing any documentation and schemas produced, and the anonymized data from the web survey is openly accessible in the Virginia Tech Data Repository.

Another option for expanding our program in the medium-term is to communicate more with the Smithsonian. Currently the Smithsonian 3D Metadata Model is at Version 0.6 with various elements undergoing edits and discussion. Once the Model is finalized, we will evaluate the value of and method for adapting our technical metadata to Version 1 of the Model. Additionally, this work specifically supports the Entomology Department Insect Collection. We will broaden our methods to support multiple projects, project types, and methodologies for 3D capture and datasets.

We will also stay apprised of new workflows and new technologies in 3D and photogrammetry. 3D methods and technology is a quickly evolving field. There is recent research and technology in automation of capture, including faster automation [12], and more in-depth automation [13]. New open-source technology like scAnt: An Open Source 3D Scanner For Ants are already improving the capture of small specimens, such as insects, to increase clarity. There is also work to improve the workflow for attaining the natural color of insect 3D models to enhance the quality of iridescent coloring [14]. Ultimately we are taking a preservation-centric approach to migrate and adapt to the community best practices as we expand our 3D program.

14) REFERENCES

[1] M. Wilkinson, "The FAIR guiding principles for scientific data management and stewardship," Scientific Data, 3, 2016. doi.org/10.1038/sdata.2016.18

[2] J. Blundell, Smithsonian Institute's 3D Metadata Model, v 0.6, Smithsonian Digital Program Office, 2011. https://dpo.si.edu/blog/smithsonian-3d-metadata-model

[8] VTDLP Preservation Events Audit Policy: https://apps.es.vt.edu/confluence/display/LIBDPLD/Preservation+Events+Audit+Policy

[3] "Reference model for an Open Archival Information System," Consultative Committee for Space Data Systems, CCSDS 650.0-B-1, 2012. https://public.ccsds.org/Pubs/650x0m2.pdf

[4] B. F. Lavoie, "The Open Archival Information System reference model: Introductory guide (2nd edition)," DPC Technology Watch Series Report 14-02, Digital Preservation Coalition, 2004. dx.doi.org/10.7207/twr14-02

[5] D. G. Davies, "Open data and digital morphology," The Royal Society, 248:1852, 2017. doi.org/10.1098/rspb.2017.0194

[6] H. Richards-Rissetto and J. Schwerin, "A catch 22 of 3D data sustainability: Lessons in 3D archaeological data management & accessibility," Digital Applications of Archaeology and Cultural Heritage, 6, 2017. doi.org/10.1016/j.daach.2017.04.005

[7] D. M. Boyer, G. F. Gunnell, S. Kaufman, and T. McGeary, "Morphosource: Archiving and sharing 3-D digital specimen data," The Paleontological Society Papers, 22, 2017. doi.org/10.1017/scs.2017.13

[8] A. Kinnaman, "Preserving 3D data in the OAIS reference model," unpublished. http://doi.org/10.5281/zenodo.3674769

[9] A. Kinnaman, "Building a 3D data preservation model for cross-disciplinary use - web survey," v 1, University Libraries, Virginia Tech, 2021. https://doi.org/10.7294/14501640.v1

[10] SO/IEC 19775-1:2013, Extensible 3D (X3D), "Information technology — Computer graphics, image processing and environmental data representation — Extensible 3D (X3D) — Part 1: Architecture and base components," 2013. https://www.web3d.org/documents/specifications/19775-1/V3.3/Part01/Architecture.html

[11] K. Stevenson, "An open source 3D scanner for ants," Fabbaloo, 2021. https://www.fabbaloo.com/news/scant-an-open-source-3d-scanner-for-ants

[12] J. J. Medina, J. M. Maley, S. Sannapareddy, N. N. Medina, C. M. Gilman, and J. E. McCormack, "A rapid and cost-effective pipeline for digitization of museum specimens with 3D photogrammetry," PLoS ONE 15(8): e0236417, 2020. https://doi.org/10.1371/journal.pone.0236417

[13] B. Ströbel, S. Schmelzle, N. Blüthgen, and M. Heethoff, "An automated device for the digitization and 3D modelling of insects, combining extended-depth-of-field and all-side multi-view imaging," ZooKeys 759: 1-27, 2018. https://doi.org/10.3897/zookeys.759.24584

[14] C. V. Nguyen, D. R. Lovell, M. Adcock, and J. La Salle, "Capturing Natural-Colour 3D Models of Insects for Species Discovery and Diagnostics," PLoS ONE 9(4): e94346, 2014. https://doi.org/10.1371/journal.pone.0094346