

OPENHERITAGE3D: BUILDING AN OPEN VISUAL ARCHIVE FOR SITE SCALE GIGA-RESOLUTION LIDAR AND PHOTOGRAMMETRY DATA

S. McAvoy^{1*}, J. Ristevski², D. Rissolo¹, F. Kuester¹

¹ CHEI, Calit2, University of California San Diego, USA – (smcavoy, drissolo, fkuester) @ucsd.edu

² CyArk, Oakland, California, USA – john.ristevski@cyark.org

KEY WORDS: Archive, Repository, Visualization, LiDAR, Photogrammetry, WebGL, Point cloud, Open Access.

ABSTRACT:

LiDAR and photogrammetry are common mechanisms for the documentation of cultural heritage sites. Their outputs provide foundational primary sources for research and important engineering decisions contributing to the on-going conservation of sites and structures. Unfortunately, due to the complex nature of these data, they are rarely shared with stakeholders in their full forms. Raw data is routinely transferred to portable hard drives and forgotten at a project's end. Digital heritage documentation, often acquired with great effort and cost, is at risk of loss (UNESCO 2003). When these data are shared, they are rarely presented or formatted in ways which enable their widespread re-use. As processing tools improve at a rapid pace, and offer new pipelines for improved reconstruction and fusion with additional data sources, it is important to preserve the data in a way which maintains all functionality and interoperability between processing platforms. It's not enough to share data or to simply make them "available." Data must be technologically and intellectually intelligible (and useful). A new system framework is required to ensure their integrity and widespread utility, and to democratize these data for re-use by the diverse communities of stakeholders with whose purposes fall outside the narrow scopes of a 3D documentation project's original goals. In this paper we present OpenHeritage3D.org as a platform and framework for an open visual archive seeking to provide authoritative and democratized access to site-scale lidar and photogrammetry through data curation, file-sharing, and web-visualization systems with granular segmentation and data conversion capabilities.

1. INTRODUCTION

OpenHeritage3D (OH3D) is a digital repository for large LiDAR and photogrammetry data related to cultural heritage sites. The project features open-source data formats, tailored metadata to ensure long term re-usability, and visualization as a service, building on a philosophy and framework established by the OpenTopography project (Krishnan et. All 2011) for geoscience aerial LiDAR data. By providing valuable functionality through web-based interactivity, rudimentary analysis, segmentation, and format conversion, OH3D seeks to encourage contributors to push past the barriers which prevent effective sharing of these data and offer immediate meaningful access to users of diverse interests and varying technical skill levels through the automatic generation of alternative media derivatives.

1.1 Challenges to Data Sharing

Field technicians and researchers are not adequately incentivized to properly share their data. The process is time consuming. Original files might need to be tracked down, reorganized, and converted. Data may be too large to transfer via conventional web-tools and require special arrangements. Metadata entry can be a tedious process, requiring lookup and multiple levels of review. Terms of use are often unclear, as collectors contracted/permitted by site authorities may not have arranged agreements for public dissemination of potentially sensitive datasets. To help persuade data contributors to overcome these obstacles value in long-term security, simplified access, in-browser visualization, simplified user tracking, and clear standardized license terms for re-use and academic citation, must be provided.

Though leading funding organizations in the United States, like the National Institute for Health (NIH) and National Science Foundation (NSF), increasingly require that data is made "open access" and require the creation of data management plans as part of their proposal processes, the result is often that data is shared

for data sharing's sake, without attention to ensuring integrity and re-use. Grantees can't be reasonably expected to be familiar with all potential use cases and workflows outside of the siloed environments in which they captured and processed their own results. They develop their own metadata structures, and systems, and may follow the data only to the publication of analytical results, relying on un-intelligible unstructured data-dumps on public repository platforms to full these grant obligations. A need exists for specialized data curation, with domain experts engaged in standardization efforts and archival platforms design to ensure that research outputs are repeatable, reproducible, trusted, accurate, and that users possess easy access to data assets.

Long-term preservation is a daunting challenge. The cyclic nature of grant and project-based funding does not provision for the future. Researchers are often left without resources for long-term management or storage for data assets. Institutional Academic and government libraries are the ideal long-term homes for data, but they have been slow to evolve to their patrons' needs regarding the evaluation, access, and re-use of complex visual data. A great many libraries are inundated with traditional data and media formats, underfunded, understaffed, and may lack expertise or incentive to cater to uncommon use cases within their diverse communities.

1.2 Research Domain-Specific Digital Archives

There exist multiple ongoing cross-institutional efforts within the GLAM sector (galleries, libraries, archives, and museums) to build national 3D repository systems and archival standards for 3D models. These efforts include the Community Standards for 3D Data Preservation (CS3DP) in United States (Moore et. all 2022), the Europeana 3D collection taskforce within European Union and U.K. (Europeana 2021), and global efforts by the International Image Interoperability Framework (2021). Unfortunately, these efforts lump sensor based 3D geospatial data with for 3D models as media and game engine assets, focusing largely on support for simplified meshes. These larger

institutional efforts to build an infrastructure to support 3D as a holistic domain, catering to both media and data, are still in a state of infancy and may fail to meet the specific needs of data domain experts.

As such, there is a present trend amongst primary U.S. research funding organizations to create free and open data format-specific digital repositories to be run by highly specialized subject matter experts who focus on the sharing, re-use, and (often) the visualization of big and complex datasets within their specific research-domain communities. U.S. Government funded open digital repositories in which meet this description include:

- **Opentopography.org** is funded by NSF to support large Aerial LiDAR surveys related to geosciences. OpenTopography provides streamlined academic access to the massive nationwide USGS 3DEP LiDAR program and other scientific studies related to geosciences or provided by other NSF funded partners. The platform includes a powerful web system allowing for on-demand segmentation, and generation of visualization outputs for web and GIS.
- **Morphosource.org** is funded by NSF to support collections related to animal skeletal imaging and natural history. Data is made available through a wide variety of licenses, and may still require special arrangements with license holders for re-use. Though built around volumetric medical-imaging modalities like MRI and CT, Morphosource has expanded to curate photogrammetric datasets of objects and animal samples. The platform provides web-based visualization for many volumetric and mesh models. Morphosource features a robust web-portal supporting user data submissions and metadata ingestion.
- **Designsafe-ci.org** is supported by the The Natural Hazards Engineering Research Infrastructure, (NHERI) and funded by NSF to support collections of scientific data related to natural hazards. This platform accepts a very wide variety of data types for simulation, event documentation, and structural analysis.
- **R2R Rolling Deck to Repository at Rvdata.us** (citation) is funded by U.S. National Oceanic and Atmospheric Administration (NOAA) Office of Naval Research (ONR) and NSF to support a variety of sensor data collected from sea-going research vessels.
- **Cancerimagingarchive.net** is funded by NIH to support anonymized volumetric human medical imaging data.

OH3D, with some seed funding from the U.S. National Endowment for the Humanities (NEH) seeks to fill a gap left by these efforts around site-scale cultural heritage data. Opentopography.org serves as the archetypal model for this effort, employing a scalable infrastructure to enable the segmentation, analysis, and re-use of aerial lidar data. The project even provides limited on demand differencing and raster transformation services to its users. Their automated visualization and on-demand analytical products do not scale well to OH3D's intended users and use cases, and thus underlying systems are altered and optimized for cultural heritage data.

1.3 Commercial Data Sharing Platforms

Self-supporting projects like the Digital Archaeological Record (tDAR) at tdar.org is a common self-supporting home for

archaeological data but may charge up to \$500 (USD) per gigabyte. A single LiDAR project may then cost tens of thousands of dollars to archive. In the case of the Mexico City Metropolitan Cathedral (Cyark 2019a) a total 463.19 GB of compressed photos and LiDAR scans would cost \$231,595 USD to archive. This may possibly be an appropriate cost representing for the expert preservation of critical data through decades and centuries, but it is nonetheless cost prohibitive for a majority of users working with big data.

Increasingly cultural heritage organizations rely on mainstream model sharing platforms, such as Sketchfab.com, to share their projects. There should be a clear distinction between media visualization platforms and data archives. Without the source data or a metadata schema which enables re-use. Many mainstream platforms do not natively support measurement or segmentation tools, required for the most basic analysis, and architectural level lidar or photogrammetry models must be drastically simplified to meet upload limits (Champion & Rahaman 2020).

The Scan the World project (MyMiniFactory, n.d.) represents a successful 3D cultural heritage repository project with singular focus on re-use for 3D printing. In the effort to provide 3D printable objects, source data are discarded. Museum objects are often reduced in quality and lose their original scales, making them useful as aesthetic representations only.

1.4 Project History

CyArk is a U.S. based non-profit specializing in 3D documentation of cultural heritage sites. In 2018, CyArk launched the Open Heritage pilot project, a showcase with Google Arts and Culture to make available a subset of the CyArk archive to a broader audience. The release of the data was well received by the cultural heritage and 3D community with hundreds of downloads and innovative re-use of the data. However, the initial pilot was limited, lacking a dedicated portal, search functionality, robust metadata, and no clear path to broaden the scope to include other potential publishers of 3D heritage data.

In January 2019, CyArk convened a workshop of organizations dedicated to documenting cultural heritage and formed a working group to determine an initial set of requirements for a web portal, file format standards and metadata schema to support the distribution and open access of 3D cultural heritage data. The group consisted of experts from CyArk, Historic Environment Scotland, and the University of South Florida Libraries, who together have significant collections of 3D cultural heritage data.

The workshop and subsequent work resulted in an initial specification for file formats, metadata schema and a web portal (openheritage3d.org) with an early prototype was launched in April 2019. Today, the portal provides a simple, easy to use interface with map-based search, project metadata, and interactive data inspection tool allowing users to preview datasets prior to download.

The number of contributing organizations has now increased to nine and includes University of California San Diego, the Oxford Centre for Islamic Studies, Texas Tech University, Stuttgart University, and University College London Institute of Archaeology. There are now 381 published projects with 664 total datasets on the platform representing locations in over 40 countries and a diverse collection of the world's cultural heritage.

In late 2022 the project was officially absorbed by UC San Diego's cultural Heritage Engineering Initiative (CHEI), which had built and hosted the site's visualization layer since 2020, and has committed to continued development of the platform. This concrete operational relationship with an academic institution is key to the long-term preservation and maintenance of the platform and all underlying datasets.

There has been broad sustained visitation and usage of the Open Heritage portal since launch with over 70,000 unique users from 183 countries generating over 300,000 page views. This usage has also resulted in 11,181 dataset download requests. Each dataset is licensed under a Creative Commons 4.0 (2013) license type chosen by the contributor.

	number
Projects (published)	381
datasets (total)	664
contributing organizations	357
countries	36
UNESCO sites	29
avg. monthly Visitors	1200
downloads	11,181

Table 1. OH3D descriptive statistics as of 31/12/2022.

	datasets	Size (TB)
terrestrial LiDAR	230	4.2
terrestrial photogrammetry	157	4.0
aerial LiDAR	2	0.1
aerial photogrammetry	142	1.7
data derivatives	113	0.3
totals	664	6.3

Table 2. breakdown of dataset data types.

2. ARCHIVAL SYSTEM STRUCTURE

The Openheritage3D archive is composed of 5 independent but inter-referenced systems or layers:

1. **The identity layer** providing a unique identity to each object, enabling a modular design and mobility as the other dependant layers are modified and preparing for the possibility that other systems may need to be re-homed or absorbed into other platforms over time.
2. **The metadata layer** is a multi-table database which stores key contextual information about the files, tying them to their creators, time, place, narrative details, and other related resources. Abstract (100 – 250 words)
3. **The storage layer** where files are kept in specific formats within compressed file archives. A long-term archival storage system, above all, is designed for preservation and addresses the problem of bit rot (degradation of lowest level physical storage medium) through backup copies and periodic comparison of files to an established baseline.
4. **The discovery layer**, (e.g. the main website openheritage3d.org), ties all other layers together for the web user, providing search functionality, and indexed by web-search platforms, and web viewers for visual derivatives (i.e. google, bing...)
5. **The visualization layer** serving derivative previews of the (point clouds, videos, unpacked images) and derivative visual metadata (polygons of geo-spatial boundaries). This layer is automatically generated from the storage and

metadata, serving intermediate files hosted formatted specifically for rapid on-demand access. The visualization layer, in this case, also provides tools to segment subsets of large datasets for download and re-use.

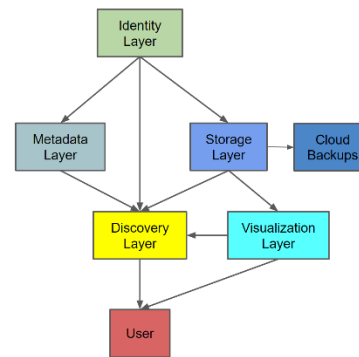


Figure 1. OH3D systems architecture diagram.

2.1 Identity Layer

Digital object identifiers (DOI) are now commonplace in academic publishing and digital repositories. Each OH3D project is assigned a DOI from datacite.org, providing a canonical reference that allows for easy citation and tracking of dataset re-use. For the archive, the DOI provides a means modify a systems back-end without any impact to users. to change metadata standard schema easing transfer to other systems, integration with search engines, and scholarly data aggregators. For the data contributor, an easily citeable source and automatic integration with a number of research profile systems like ORCID.

2.2 Metadata Layer and Metadata Ingestion

Raw data are useless to external expert users without key narrative details contextualizing the place, time, method, and goals of a site capture. These details must be curated separately from the files, as it is useful for populating other layers and identifying useful relationships between key information. One contributor may create many projects, many projects may exist within a single country, a number of different uploaded projects portraying the same site may have many variations of spelling or naming conventions, but all refer to the same geospatial location. The complexity of these relationships between files and narratives must be stored in a relational database, and the fields or descriptors in that database must conform to a standard which makes it widely useful by external systems. In addition to standard metadata fields, OH3D required specific descriptors to help users evaluate the data and potential for re-use at a glance. These fields include: capture device type, device model, size of the files submitted. This model is intended to evolve to support new imaging technologies, to encompass unique information related to other modalities.

2.2.1 Metadata Schema: OH3D descriptive metadata aims to describe 3 key relationships, between the project and the descriptive data, the multiple individuals and organizations (entities) involved in the capture, funding, or publication of a project, and the datasets related to each project which each possess a type and capture device. OH3D follows incorporates DataCite 4.4 metadata schema (DataCite, 2021) which promises simplified ingestion into external asset management systems and discovery systems.

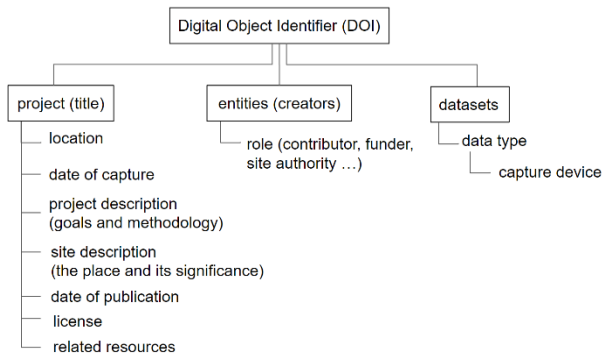


Figure 2. OH3D metadata architecture.

2.2.2 Metadata Ingestion is currently performed through the use of spreadsheet templates delivered through google sheets, which is then processed via a scripted system which assigns a DOI, injects the data into a DataCite record, and then replicates the record in OH3D's own local database. This method can be advantageous for batch ingestion of a great many objects, but many OH3D contributors find it to be cumbersome and confusing, and the system does not allow for simple text editing. OH3D will be investing in an account and form-based system which allows users to manage their own submissions, much like a system implemented by morphosource.org.

2.3 Storage Layer and File Structure

2.3.1 Archival Storage Layer: The storage layer is modular, easily moved between various cloud storage systems. Dropbox provides the live central archival storage for the OH3D project, though copies exist on Google Cloud storage, Seagate Lyve Cloud s3 storage, and at UC San Diego. Any copy can be toggled as the live storage layer should fail. Digital storage infrastructure is complex and, for the purposes of this project in its current phase, commercial cloud storage options offer the most attractive option. As the project grows and gains institutional footings it is very likely that the archive will be hosted on university managed storage servers, potentially split between multiple institutional systems. This is an attractive feature, recognizing that a number of partners have asserted a need to maintain some semblance of national sovereignty and responsibility for the original data, whatever the terms of the public facing user license.

For purposes of security the storage layer is not directly accessible by users. Instead, users request download through a very brief form on the discovery layer, which creates a request for download, and then sends temporary links to a user's email address.

2.3.2 Archival File Formats: LiDAR and photogrammetric processing tools are improving at a rapid pace. It is necessary to preserve projects in a way which enables future reprocessing with different tools. This can be a challenge as LiDAR systems take on more proprietary software-based profit models. The archival data must employ file standards which preserve the foundational nature of each data type, and promote interoperability between multiple proprietary processing ecosystems. All files are stored in compressed ZIP archives following the naming convention (DOI suffix)(data_type).zip. These formats and their key features are detailed below:

- **Terrestrial LiDAR (TLS)** is stored in the E57 file format (Huber, 2011). TLS is composed of individual spherical scans, aligned to each other via shared features. Each scan contains a centre point from which the rest of the scan was originally captured, the surrounding point-cloud, and may contain an associated image set or panoramic image used to pass RGB colour values onto the point-cloud. These images can be higher resolution than the associated laser scan and are key to processes enabling the merging of photogrammetric and terrestrial lidar data. The E57 point cloud format preserves this structure in its entirety, maintaining a hierarchy of scans, centres, points, and images within a single file.
- **Aerial LiDAR (ALS)** is stored in LAS file format, which is a longstanding geospatial standard for the datatype (American Society for Photogrammetry and Remote Sensing, 2019). The key features of ALS are the transects (lines) which are captured in a single sweep, and the "return" values which help differentiate ground points from other features. Each individual point in a LAS file contains fields with this information. Transects are akin to scans, preserved through a GPStime field, showing points captured in sequence. The alignment of individual lines may need to be manipulated, and any included feature classifications may need to be created or refined with a variety of external toolsets.
- **Terrestrial Photogrammetry and Aerial Photogrammetry** datasets are stored as JPEG, TIF, or DNG raw images. Raw images present a challenge, as most common formats (ARW, NEF, CR2) are proprietary. Though these formats are supported widely by open-source tools, the formats themselves are not open and their long-term utility are therefore suspect. Terrestrial photogrammetry datasets may also contain global control points (GCPs) in a simple CSV file and may contain GPS information within each file's embedded EXIF data. Aerial photogrammetry datasets, captured from drone or plane, are expected to always contain GPS EXIF data, and must be kept in formats which preserve that EXIF data schema. External documentation regarding flight plans and pilot logs are welcome additions.
- **Data Derivative** volumes are unstructured dumps of related files supporting a primary source dataset. They may contain reports, renderings, blueprints, and visualization derivatives which do not fit into the normal pipeline. It is expected that contributors organize and describe the contents of the data derivative volumes, but no rigid metadata model is currently applied to these data types. These volumes are, admittedly, a clumsy stopgap attempting to meet the need for wider process documentation, and will

hopefully evolve as paradata and reporting standards solidify (European Commission, 2022).

2.4 Discovery Layer

The landing page is a world map based on the Esri Leaflet plugin with location pins for each project. These pins direct to individual records pages. A text-based search option is available in another tab, displaying Project name, country, publication status, and Each project record page contains relevant metadata, a point cloud viewer (for all TLS datasets, and select photogrammetry datasets), a map viewer showing a simple polygon of dataset bounds overlaid on google maps. At the end of the page is a suggested citation populated from associated metadata.

2.5 Visualization and Segmentation Systems

The visualization system is a key motivating force for contributors and users alike, offering an immediate value through on-demand real-time access to otherwise unwieldy models. It is designed to provide quick web-streaming previews of the raw data, use of basic measurement tools, and allows the user to quickly identify and segment small component parts of a larger model. The visualizations themselves are derivative products and lose the structure and functionality available in raw data formats. Ideally a single file could serve all programs and needs through a single standardized and readable format, but the current state of viewer technologies and analytical platforms necessitate the creation of specialized derivatives which are optimized for visualization. As web visualization technology shifts, and underlying rendering engines, viewer software, and web-optimized data formats evolve, it is critical that these systems be rebuilt or expanded periodically, and therefore must be structured for automation and for the curation of multiple streaming products.

2.5.1 Point cloud Viewer, Formats, and Utilities: LiDAR and photogrammetry data are often difficult to utilize in their raw, native, full resolution formats. Much time is spent simplifying models for use in game engines or GIS. A user, unfamiliar with arcane point cloud workflows, may spend days downloading a huge dataset, only to discover that they are unable to open the file for lack of RAM on their personal machines. A streaming point cloud-viewer save all users time, frustration, and enables re-use of data by all stakeholders outside of the small community of technology-rich format savvy users. Point clouds are an ideal visualization format for the purpose of data-preview, as they benefit from a longstanding suite of tools built to process them at scale, can load quickly on web viewers, and show the data as-is (not interpolating sloppy geometry over empty spaces).

OH3D's point cloud viewer employs the Potree octree-based multi-resolution viewer (Schütz, 2016) built on the three.js (three.js authors,2020) WebGL game engine, with assets stored in the compressed Potree Converter 2 format (Schütz, 2023). The benefits of this pipeline are:

1. Throttled client-side rendering, enabling performance optimization on a variety of devices with varying network speeds and graphical capabilities.
2. Integration with a wide range of tools and active user community, including the Cesium.js base map, and over 40 other 3D formats.
3. Simple file structure (3 per point cloud) optimizing transfer between network locations.

4. Rapid octree structure format conversion for giga-resolution point clouds, saving days of conversion time over the Entwine Point Tile (EPT) archival octree-structured format (Manning, 2016) employed by opentopography.org.

OH3D automatically generates Potree viewers for all LiDAR datasets, and for select photogrammetric datasets Unfortunately, the Potree format requires a transitional conversion to the LAS/LAZ file format, losing the native spherical scan-based structure inherent to terrestrial lidar models. For this reason, the Potree viewer's point cloud must be considered a lossy derivative product and must therefore be maintained and described separately from the archival format files.



Figure 3. Mexico City Metropolitan Cathedral photogrammetric model.

2.5.2 Segmentation: Large datasets may contain many smaller features of interest to a diverse set of users who may not require the entire dataset. The visualization layer can support these users by enabling granular segmentation. A user may mark a region of interest (ROI) using a profile line or a clipping box and download only the points encapsulated within that ROI. The Metropolitan Cathedral in Mexico City (Figure 3.) and its ornate Altar of Kings (CyArk, 2019a) contains dozens of individual sculptures, paintings, and ornaments (Figure 4). A single statue (Figure 5.) can be selected and downloaded as a LAS file for external use, without requiring the download and external manipulation of the whole 400 GB model.

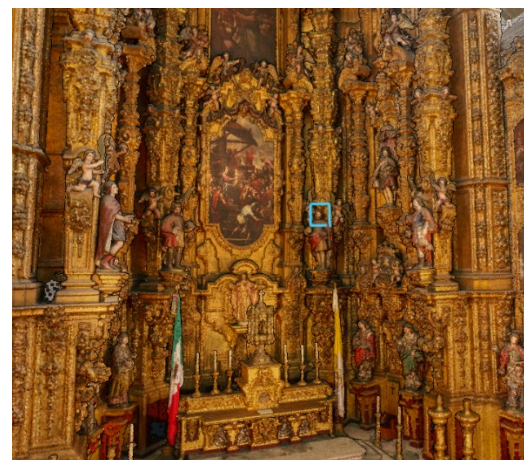


Figure 4. Mexico City Cathedral interior, the Altar of Kings, highlighting one statue's head.



Figure 5. Segmented statue from the greater Altar of Kings.

2.5.3 Oriented Image Viewer: Photogrammetric models are created through the matching and positioning of numerous photo inputs. A 3D model output may serve as a map for the locations of individual images. Instead of digging through a folder containing thousands of images, Potree's oriented image features allows the user to navigate each image by its spatial location, load that image and download it at full resolution. A sparse point cloud of the Tomb of Tu Duc (CyArk, 2019b) in Vietnam is ringed by its many photo inputs (Figure 6). A single photo of the tomb's ornate roof is previewed in space and downloaded (Figure 7).

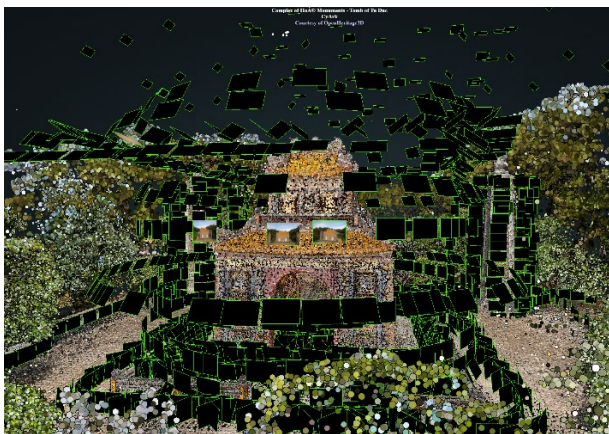


Figure 6. Tomb of Tu Duc sparse point cloud, with oriented images aligned in space.



Figure 7. Tomb of Tu Duc oriented image loaded in spatial context. Each image is individually downloadable.

2.5.4 Video: Point clouds are an arcane format, and 3D viewers are complex and require some time commitment to master. For these reasons OH3D generates flythrough videos for each of its point cloud visualizations, focusing on key features. These videos offer a preview backup should underlying point cloud viewer system malfunctions and are a more widely accessible means for stakeholders to engage with these data. The point cloud viewer template is programmed to detect whether a device is capable of running the base WebGL 2.0 library required for Potree and redirects to the video if the check fails. The Potree viewer offers built in tools for smooth flythrough planning, but it's just as easy to screencap any session within the viewer using any external screen recording software.

3. DISCUSSION

This paper reflects initial steps in the development of a larger system. OH3D will require support and feedback from the greater community and institutional partners.

3.1 Weaknesses of Research Domain and Data Format-specific Archives

UNESCO (2003) recommends that cultural heritage documentation is done with institutional partners whenever possible, to ensure long term security and maintenance of supporting infrastructure. Library collections offer a much greater diversity of content, and as a single archaeological expedition or engineering survey may result in a wide variety of data formats and analytical derivatives which do not fit into any one of these pots, the narrow focus on particular data outputs can result in a body of work being split across a half dozen different platforms. Ideally a single archive collection would be able to encompass all materials involved in the initial planning of documentation efforts, annotations, the onsite data acquisition including the paradata (European Commission, 2022) describing the site conditions at the of the moment of capture and various skill levels of human data collectors, along with post campaign data staging and analysis.

Libraries can also ensure long term data preservation beyond the lifespan of any one project or person. The lives of OH3D collections are much more precarious than those hosted at institutional repositories. With that in mind, all OH3D records and data are designed to be transferable, with global licenses, with standardized metadata schema, and web identifiers. This structure supports a potential migration to an entirely different

context, leaving a pathway to streamline future mergers with institutional library preservation/discovery systems.

3.2 Populating the Metaverse

One of OH3D's eventual goal is to provide a suite of outputs which can easily integrate into multiple metaverses, requiring specific proprietary file outputs for various platforms and updating as necessary. For example, a single photogrammetric dataset could be converted into web-optimized point clouds, Cesium.js GLTF tiles, Unreal 5 nanite assets, videos, orthophotos, DEMs... each streaming into a different system and within a customized context. This effort requires a continued commitment to maintain interdisciplinary visualization pipeline expertise, for both analytical and media-driven platforms.

3.3 Annotation

Centrally hosted 3D visualizations built on geospatial coordinate systems offer an opportunity for layering of expert and community annotations, a sort of 3D Wikipedia. An engineer may add commentary concerning the structural integrity of a beam, an art historian may add a layer of notes describing a mural, an epigrapher may contribute translations to all inscriptions. With a translatable geo-spatial coordinate system in place, these same simple annotations can be loaded in a 3D or GIS context, stored as separate layers, and attributed to expert authors. That same infrastructure can then be leveraged towards a more open community annotation space.

3.4 Supporting Additional Under-Served Sensor Formats

Though the current implementation of OH3D is structured for aerial/terrestrial LiDAR and Photogrammetry, there is a need for new data preservation standards for mobile SLAM-based lidar systems, short-range structured light scans, and of various maritime 3D/volumetric data formats like sonar multibeam portraying cultural heritage data. OH3D is actively evaluating these formats for potential inclusion within the larger system framework.

4. CONCLUSION

The digital heritage community has made great strides in documentation (data acquisition) and the dissemination of more easily sharable derivatives and visualizations. These "lightweight" products, however, have limited utility (beyond enhancing interpretive or educational efforts). The value of an online digital asset (to researchers and conservators, in particular) is diminished when they exist only as an OBJ file, for example, spinnable in a web browser. Most challenging is providing communities of practice with ready access to original digital assets and more robust datasets – across resolutions and modalities – for the purpose of facilitating analysis. Often researchers require only a portion of a larger digitization or virtualization project. Ideally, such components – from artifacts to sculpture to architectural elements – should be findable, viewable, and retrievable within their larger (and meaningful) spatiotemporal context. And once retrieved, with provenance intact, they can be integrated into external workflows or platforms. Designed to be responsive and resilient, OH3D strives to meet the demands of contributors and users across the expertise spectrum.

ACKNOWLEDGEMENTS

The authors would like to thank Scott Lee, Ilayda Bozhurt and Avidan Fernandez for their support and vision in designing and implementing facets of the OH3D platform. We would also like to acknowledge University of California San Diego Library staff Maria R. Adams, Daniel LaSusa, Thor Brickman, Michael Stuart, Ho Jung Yoo and Abigail Pennington, for their work, support, and guidance in the building of the OH3D platform visual systems.

REFERENCES

- American Society for Photogrammetry and Remote Sensing, 2019. LAS Specification v.1.4 - R15. www.asprs.org/wp-content/uploads/2019/07/LAS_1_4_r15.pdf (8 February 2023)
- Champion, E., Rahaman, H., 2020. Survey of 3D digital heritage repositories and platforms. *Virtual Archaeology Review*, 11(23), doi.org/10.4995/var.2020.13226
- Creative Commons 4.0, 2013. <https://creativecommons.org/licenses/>. (8 February 2023)
- CyArk, 2019a. Mexico City Metropolitan Cathedral - LiDAR - Terrestrial , Photogrammetry - Terrestrial, Photogrammetry - Aerial. doi.org/10.26301/8j7j-6m43
- CyArk, 2019b. Complex of Hué Monuments - Tomb of Tu Duc - LiDAR - Terrestrial, Photogrammetry - Terrestrial, Photogrammetry - Aerial. doi.org/10.26301/n06n-qa49
- DataCite, 2021, DataCite Metadata schema 4.4. schema.datacite.org. (8 February 2023).
- Europeana, 2020. 3D Task Force Report V1. Europeana Network Association Members Council <https://pro.europeana.eu/project/3d-content-in-europeana> (8 February 2023).
- European Commission, 2022. Study on quality in 3D digitisation of tangible cultural heritage. – VIGIE 2020/654. digital-strategy.ec.europa.eu/en/library/study-quality-3d-digitisation-tangible-cultural-heritage (8 February 2023).
- Huber, D., 2011. The ASTM E57 file format for 3D imaging data exchange. Proceedings of SPIE - The International Society for Optical Engineering. 10.1117/12.876555
- International Image Interoperability Framework, 2021. IIIF 3D Community Group Charter. iiif.io/community/groups/3d/charter (8 February 2023).
- Krishnan S, Crosby CJ, Nandigam V, Phan M, Cowart C, Baru C, Arrowsmith RJ. 2011. OpenTopography: A Services Oriented Architecture for Community Access to LIDAR Topography. Proceedings of the 2nd International Conference on Computing for Geospatial Research & Applications.
- Manning, C., 2016. Entwine Software, Hobu Inc. entwine.io (8 February 2023).

Moore, J., Rountrey, A., Kettler, H. S., 2022. *3D Data Creation to Curation: Community Standards for 3D Preservation* (16-85), Chicago, Illinois: Association of College and Research Libraries.

MyMiniFactory, n.d. Scan the World. myminifactory.com/scantheworld (5 February 2023).

Schütz, M., 2016. Potree: Rendering Large Point Clouds in Web Browsers. (unpublished master's thesis), Vienna University of Technology, Vienna, Austria. cg.tuwien.ac.at/research/publications/2016/SCHUETZ-2016-POT (5 February 2023).

Schütz, M., 2023. (PotreeConverter) Software, Version 2.1. github.com/potree/PotreeConverter/releases/tag/2.1.1 (5 February 2023).

Three.js Authors, 2020. Three.js Software, Version r126. github.com/mrdoob/three.js (8 February 2023).

UNESCO, 2003. Guidelines for the Preservation of Digital Heritage. unesdoc.unesco.org/ark:/48223/pf0000130071 (8 February 2023).