An Archival Framework for Sharing of Cultural Heritage 3D Survey Data: OpenHeritage3D.org

Scott McAvoy [1*], Beatrice Tanduo [2], Alessandra Spreafico [2], Filiberto Chiabrando [2], Dominique Rissolo [1], John Ristevski [3], Falko Kuester [1]

[1] Cultural Heritage Engineering Initiative, University of California San Diego, USA (smcavoy, drissolo, fkuester)@ucsd.edu
[2] Laboratory of Geomatics for Cultural Heritage (LabG4CH), Department of Achitecture and Design (DAD), Polytechnic University of Turin, Viale Pier Andrea Mattioli, 39, Torino (TO), Italy – (beatrice.tanduo, alessandra.spreafico, filiberto.chiabrando)@polito.it
[3] CyArk, Oakland USA – john.ristevski@cyark.org

Keywords: Archives, 3D-survey, Metadata, Paradata, Reuse, Reproducibility, Data Quality.

Abstract

Photogrammetry and LiDAR have become increasingly accessible methods for documentation of Cultural Heritage sites. Academic and government agencies recognize the utility of high-resolution 3D models supporting long-term asset management through visualization, conservation planning, and change detection. Though detailed models can be created with increasing ease, their potential for future use can be constrained by a lack of accompanying topographic data, data collector skill level, and incomplete recording of the key metadata and paradata which make such survey data useful to future endeavors. In this paper, informed by various international survey organizations and data archives, we present a framework to record and communicate Cultural Heritage - focusing on architectures based on 3D metric survey - to first describe the data and metadata which should be included by surveyors to enable data usage and to communicate the expected utility of this data.

1. Introduction

In recent years, the proliferation of low-cost and easy-to-use sensors and software for the 3D documentation of Cultural Heritage (CH) assets has democratized the process, enabling even non-technical stakeholders to engage in 3D model creation and sharing. Alongside this democratization, significant knowledge gaps have emerged regarding the quality and utility of the generated data. Numerous commercial and institutional platforms for online sharing of CH replica have arisen in recent years, still, the literature points to a scarcity of shared digital models within scientific publications (Champion & Rahaman, 2019). There is often little consideration given to providing metadata and paradata describing how these replicas are generated and evaluating their metric quality, making difficult data visualization and reuse, especially for scientific purposes, like change detection or restoration efforts.
The sharing of digital CH replicas has stimulated debate within the scientific community concerning interoperability, metadata supply, and quality assessment (Champion & Rahaman, 2019; Homburg et al, 2022; Pamart et al, 2023). 2D and 3D products can be the result of invention or can be based on graphical/textual documention or on reality. Reality-based models can rely on manual or digital (instrumental) measurements, generating digital models with different levels of accuracy, detail, and reliability. Focusing on digital surveys, object size, geometry, materials, climate conditions, and documentation scope influence the choice of a survey methodology and sensor over another. Movable objects, architecture and landscape cover different scales and related levels of detail, requiring a range of specialized survey expertise and metric products as outputs. More-or-less-consolidated techniques and sensors generate different raw data types (images, trajectories, point clouds, etc.) in open or proprietary file formats, while derivative products (point clouds, meshes, orthophotos, DTM, DEM, CAD and BIM files, etc.) can be independently from the sensor used for the acquisition. The complexity of these sensors and functional outputs result in a diversity of digital models. There is no single unified metadata schema and file format able to reflect the complexity of these data, or anticipate future technological, software, and hardware changes. Data provenance is essential, it can be impossible to reconstruct critical details if no metadata and paradata are provided, depriving the raw and derivate products of understandability, searching, reliability, trustworthiness, reusability, and reproducibility. However, as Champion (Champion & Rahaman, 2019) stated, if properly documented, CH 2D and 3D data can have as much scholarly value as textual publications; hence, they should be treated as scholarly sources. Survey data, especially those performed with great attention to accuracy and precision, can be held to a scientific standard. This stance, however, is rarely adopted within the field of CH. It can be difficult for creators/surveyors to anticipate the wide range of user needs and knowledge levels beyond the immediate scope of their projects, and there are many potential data points which might facilitate reuse by external users. It is important to weigh these potential uses against the creators' ability or willingness to build additional documentation in support of hypotheticals.
With this paper we propose a new framework within the OpenHeritage3D platform to help users and contributors evaluate data quality and utility, to record, and maintain key metadata enabling the dissemination and possible future reuse of CH data. This evaluation considers only 3D metric surveys conducted with digital sensors for CH corresponding to an architectural scale, using Salvation Mountain (California, USA) as case study. By incorporating key indicators of metric data quality (such as resolution, density, accuracy, precision, representational scale) and completeness (including availability of raw/processed data, image/range-based data, topographic data, interior/exterior geometry), we provide an exhaustive documentation of the available data to serve as an example for future reuse, and to provide critical information enabling further study of sites represented in these data.

2. 3D Data Archives and Cultural Heritage

2.1 Related works

The clear popularity of 3D model – inside and outside of CH - has captured the attention of institutional and commercial organizations, replying to different needs and providing

diversified options. For CH, often institutional platforms - like *Europeana* and *Googles Arts and Culture* –were built to support a variety of more traditional CH documentation formats, like books, images, videos, sounds, and have later extended their existing systems to integrate digital models of architecture (Statham, 2019; Fernie, 2024), making it difficult to provide adequate support to these complex data. Some institutions and EU- or national-funded projects (like *INSEPTION*[1], *ROCK*[2], *3DHOP* by Italian CNR[3]) are developing web platforms to store, visualize, and share interoperable 3D models of CH in multiple formats. Commercial platforms (such as *Sketchfab, Scan the World*[4] and *TurboSquid*[5]) - enabling licensing of 3D content of not only CH - are proliferating, but often focus on derivative models, and do not archive and disseminate raw data. Though useful and popular, these platforms are unstable and may not be relied upon for long-term storage, as was the case with the *Google Poly* platform, discontinued in 2021. 3D models find more success in repositories for commercial purposes, however metadata and paradata are almost always missing. *Sketchfab* is becoming very popular as visualization tool in institutional platforms, *Eurepana* portal and UNESCO *Dive into Heritage*[6] (a dedicated web portal for exploring in 3D UNESCO sites models that is currently under development) are using *Sketchfab*. It should be noted that certain institutional platforms for CH – like *Europeana* - are aggregators of digital models uploaded to other platforms, like *Sketchfab*. On almost any platform – commercial or not - the digital replica is shown in a gallery and the search can be performed and refined according to metadata (such as category, license, downloadability, format). In many cases the 3D models float in a blank 3D space without context. Some platforms are only for visualization denying data download and not permitting simple interactions like measuring and sectioning (such as *TurboSquid* and *Skecthfab*). Limitations are connected also to file weight limits and file formats, requesting decimation and or converting of original data to fit the requirements and upload the model onto the web (for example the *Sketchfab* allows 500MB maximum for a single file).
As 3D documentation tools for CH assets become increasingly accessible, so does the need for effective data management practices. While these tools empower diverse stakeholders in preservation efforts, it is crucial to establish standardized protocols for archiving data and metadata, and display the contents. Despite the efforts of numerous funding agencies, which allocate significant resources to support these initiatives (Champion et al., 2020), a critical aspect is the incorporation of raw data and metadata describing the acquisition and processing methods. This is particularly evident in the field of metric quality control, a key component for ensuring the reliability and accuracy of 3D data in CH preservation projects. Notable examples, such as the institutional *Smithsonian 3D digitization initiative*[7], *Europeana,* and commercial platforms like *Sketchfab* (Statham et al., 2019), underscore the importance of integrating comprehensive metadata protocols into archival practices.
To address these challenges, it can be helpful to look beyond the domain of CH and draw inspiration from existing models in other fields. Platforms like *OpenTopography.org*, *Morphosource*, and *OpenAerialMap*, gathering geospatial and scientific data formats, offer valuable insights into effective data management strategies and standardized protocols for metadata documentation. Through the utilization of these platforms' established best practices, the CH domain can improve the 3D data quality, understating, accessibility, reuse, and interoperability.

2.2 Existing Standards for Metadata and Paradata

Data, by themselves, often lack the crucial contexts which give them meaning. Almost all disciplines produce data and need to rely on metadata and paradata to describe them. Over the years, many institutions provided frameworks for metadata and paradata like ICOMOS, Data Documentation Initiative (DDI), National Information Standards Organization (NISO), World Wide Web Consortium (W3C), International Organization for Standardization (ISO), Dublin Core Metadata Initiative (DCMI), International Federation of Library Associations and Institutions (IFLA), Getty Research Institute, International Council on Archives (ICA). Among these, ISO 19115 – which codifies geographical information and services - simply defines metadata " as data about data". Regarding paradata, literature demonstrates significant disagreement regarding the definition and intended purpose of paradata. In this research we refer to paradata definition provided by Sköld (Sköld et al., 2022): "information phenomenon that describes processes that put into existence some scholarly product, e.g., scientific publications, 3D heritage models, or museum artefacts." As recognized by the Seville Principles (2017), paradata are fundamental to guarantee scientific transparency of data.
In the field of CH, the relevance of metadata and paradata is stressed by international documents like the London Charter for the computer-based visualisation of Cultural Heritage (2009), UNESCO Vancouver Declaration - The Memory of the World in the Digital Age: Digitization and Preservation (2012), and ICOMOS Principles of Seville for Virtual Archaeology (2017). As called for by Seville Principles, metadata and paradata should be concise, clear, and easily available. In fact, metadata and paradata are crucial for enabling the widely accepted FAIR Principles of Findability, Accessibility, Interoperability, and Reuse of digital data (Wilkinson et al., 2016).
Many schemas for metadata and paradata are conceived for different purposes and data type (D'Andrea & Fernie, 2013), such as Metadata Object Description Schema (MODS) and Metadata Encoding and Transmission Standard (METS) by the Library of Congress, Data Documentation Initiative (DDI), Lightweight Information Describing Objects (LIDO) and Europeana Data Model (EDM) by Europeana, CARARE, CRMdig extension of CIDOC CRM, and Dublin Core Metadata Initiative (DCMI). However, these schemas are sophisticated and hard to understand from non-experts.
Evaluation about the quality of metadata and paradata of shared data is a discussed issue (Király et al, 2019), but it is based on completeness, consistency, accessibility, and conformity, but not on the reliability of the data itself, and specifically not on the metric control of 3D models related to their possible reuse. Furthermore, metadata and paradata do not specify for which use the data are good. A synthetic grade is auspicial to help more-or-less expert users in understanding a given dataset and for which purposes it is suitable.

[1] http://www.inceptionhbim.eu/Platform/ (Retrieved 26th April 2024)
[2] https://opendata.rockproject.eu/rock/#/home (Retrieved 26th April 2024)
[3] https://3dhop.net/ (Retrieved 26th April 2024)
[4] https://www.myminifactory.com/scantheworld/about (Retrieved 23rd April 2024)
[5] https://www.turbosquid.com/ (Retrieved 26th April 2024)
[6] https://whc.unesco.org/en/dive-into-heritage/ (Retrieved 26th April 2024)
[7] https://3d.si.edu/ (Retrieved 26th April 2024)

https://doi.org/10.5194/isprs-archives-XLVIII-2-2024-241-2024

2.3 Use Purpose, Audience and Interactions

Use purposes can be manifold and depend on the users. Fernie (Fernie, 2024) identified five types of audiences: scholars and researchers, educators and students, museums, experts and professionals, and general users. Each category requests the use of digital models in several manners and for different purposes, like researchers, professionals, and museum experts working on restoration operations - or other works like 3D printing - needs models with a metric accuracy suitable to the needs and other metadata and paradata to understand how the data are obtained and especially in which day, allowing spatial-temporal analysis. Museum experts working on exhibitions, educators and students, and general users probably do not need accurate measures, requesting only to visualize a digital model to observe the shape and colors of the CH site. These uses influence the manners in which data are captured and shared. The online visualization of digital CH replica allows a preview of the downloadable data, removing software and knowledge related barriers to user access. The online viewer might be considered the sole final use, hence, a series of functions (zoom, query, pan, rotate, light control) are relevant to fulfil this need. Other tools (measuring, sectioning, clickable bookmarks, hide/show elements, annotation) can be integrated to let people work online to perform basic analysis, without requesting them to download and install additional software packages.

2.4 Data Types and File Formats

To determine the most appropriate format for the full description of 3D CH assets, it was required to identify and analyze the existing 2D and 3D data types and formats, including their pertinent metadata (Table 1). The analysis was conducted from three different points of view: i) universality, which refers to their ability to be utilized across various software without requiring conversions; ii) interoperability, which pertains to their ability to be imported and utilized without loss of data; iii) flexibility, which refers to their capability to accommodate new features (European Commission, 2022). In this summary, derivates products aim to display technical drawing - like BIM and CAD files - are not considered.

Data type		Data formats
3D data	Point clouds/scans	LAS/LAZ, E57
	Mesh, 3D models	OBJ, PLY
2D data	Images	JPG, PNG, TIFF, GeoTIFF
	Raw images	DNG
	Vector files	DXF
GIS data		Shapefile, GML, GeoPackage
Topographic Data	Total Station Data	CSV, TXT, ASCII
	GNSS data	RINEX
Other	Documents	ODF, PDF

Table 1. Existing open data formats for different data types.

3. OpenHeritage3D.org

The OpenHeritage3D.org (OH3D) project is an open-source scholarly archival platform, first released in 2019, designed in collaboration with architects, engineers, experts in Geomatics, digital archivists, and 3D data processing and visualization specialists. It is a joint project supported by the University of California San Diego, CyArk, and Politecnico di Torino (Italy). OH3D partners with academics, non-profits, individuals, and government entities that collect 3D surveys of important CH sites but do not otherwise possess the infrastructure to store or publish these data. Laser scanners have been commercially available for almost two decades, but there has been little perceived benefit in sharing the raw data externally. The increased graphical computing capabilities of consumer-grade computer hardware, powerful open-source visualization tools, and improved processing software have encouraged the re-processing and reuse of these old data, which often contain useful detailed models. OH3D seeks to enable the open publication, dissemination, visualization, and reuse of over 700 complex architectural scale LiDAR and photogrammetric raw datasets. These datasets span 36 countries, including 30 UNESCO World Heritage sites, contributed by nearly 400 private, government, and academic entities. The platform incorporates carefully structured open data format, authorship, and metadata frameworks reflecting the data structures of 3D capture modalities to establish a canonical provenance for survey data, and to ensure the broadest utility of these data for posterity, anticipating potential opportunities beyond the scope of their initial intended purposes (McAvoy et al., 2023). OH3D uses the Potree which enables visualization of models of unlimited size (Schuetz 2016). Datasets are organized into projects and are searchable by various metadata or by navigating a map. The archive was built around terrestrial laser scanner (TLS) and aerial laser scanner (ALS) data, which are saved as point clouds in within the standardized formats (.E57, .LAS/.LAZ formats) and in metric units. Aerial and terrestrial photogrammetric datasets are also supported as key survey methodologies (images in .JPG format and open RAW formats such as .DNG), as are data from various Mobile Mapping Systems (MMS) which contain point clouds, images, and trajectories (in .TXT format). All datasets must be made available under a Creative Commons license, explicitly enabling free non-commercial use. For each project submitted to OH3D, several information are provided by the submitter to complete the dataset, while the Digital Object Identifier (DOI) is provided by OH3D team (Table 2).

General attributes	Mandatory	Provided by
DOI	Yes	OH3D
Project Name	Yes	Contributor
Publication Status	Yes	Contributor
License Type	Yes	Contributor
Country	Yes	Contributor
Lat Long Center	Yes	Contributor
Country	Yes	Contributor
Collection Start Date	Yes	Contributor
Collection End Date	Yes	Contributor
Date of Publication	Yes	OH3D
Data Type	Yes	Contributor
Data Bounds Lat/Long	Yes	Contributor
Device Type	Yes	Contributor
Device Model	Yes	Contributor
Data Size	Yes	OH3D
Project Background	Yes	Contributor
Site Description	Yes	Contributor
Keywords	No	Contributor and OH3D
External Project Link	No	Contributor
Contributors	Yes	Contributor

Table 2. Information collected by OH3D for each project.

Efforts to revive and reorganize legacy data into a structured data schema are not generally funded and can require significant

https://doi.org/10.5194/isprs-archives-XLVIII-2-2024-241-2024 | © Author(s) 2024. CC BY 4.0 License.

commitments of extra time on the part of data contributors. As such, OH3D has implemented a number of incentives to provide immediate value for this effort:
i) scholarly publishing- each dataset is granted a DOI, enabling easy citation and tracking of reuse. This non-traditional publishing method enables data collectors and other stakeholders to be designated as “authors”, an attractive proposition for individual participants and organizations who would otherwise be excluded from related peer-reviewed works. There are no limits on these attributions;
ii) embedded web-streamable 3D point clouds of unlimited size, enabling contributors to access and share their full resolution models on the fly without specialty software;
iii) creation of simple web-based video previews on OH3D and youtube, a simple and accessible media output to help describe the contents and character of the datasets to various stakeholders.

3.1 Proposed Metadata and Paradata Implementation

OH3D is designed with the intention of providing a platform for scientific archiving and sharing of 3D data for CH, but it lacks a synthetic indicator to help communicate data quality and completeness. When performing fieldwork with complex equipment and data processing pipelines, there is often a great deal of potentially important information concerning the surveyed object, who acquired the data, what and how is acquired, and by whom and how the collected data are processed (European Commission, 2022.). There are two strategies to cope with these issues. One is to attempt to build sufficiently thorough metadata models to incorporate as many globally relevant identifiable variables as possible (Pamart, et al. 2022), but a result of this method is a data contributor presented with long forms listing of fields which are not of any immediately apparent value. This can be a significant demotivating factor, causing significant disruption to archival efforts. The second strategy is to provide a separate, less demanding, a short list of metadata, file formats which are easily queried, abbreviated fields for information which can be referenced externally, and an expedition report which attempts to describe relevant descriptors. The latter strategy is the selected one for implementing OH3D, keeping only a few entries as mandatory, the ones that are almost impossible to retrieve if Exif files do not exist or if the creator does not provide them. Other metadata and paradata are listed to provide a comprehensive understanding of the datasets, considering separately raw and processed files, and including specific entries according to the sensor type. Since OH3D is mainly focused on visualization of un-interpolated point clouds, we are more focused on describing this kind of data, leaving space for further expansion to other data like meshes for the future. Collecting metadata and paradata can be achieved manually – requesting the creator to collect the needed information and correctly fill out different fields - or automatically – letting software and codes to fill out a predefined schema (Homburg et al., 2021; Sköld et al., 2022). Even if the automatic method seems more reliable and less time-demanding for operators, it works only with software and schema for which it has been implemented. Since OH3D collects datasets worldwide and leaves creators free to propose datasets collected with several different sensors and processed by many software, we opted for a manual inserting. Based on the questionnaire entitled *“Survey on quality in digitisation of tangible cultural heritage”* promoted within the project “*Study on quality in 3D digitisation of tangible cultural heritage*” (European Commission, 2022), we have defined a series of metadata and paradata to take into account and provide to the users for each dataset published in to the OH3D platform (Table 3).

Site/object general parameters		**Mandatory**
Site access (e.g. official permission, remote areas, ...)		No
Site dimension		No
Surface conditions (e. g. material, texture,…)		No
Environmental conditions (e. g. sunny, cloudy, …)		No
Acquisition phase parameters		No
Photogrammetry	Number of images	No
	Quality of images	No
	GSD	Yes
	Sensor Name	Yes
	Mean acquisition distance	Yes
TLS	Acquisition parameters	Yes
	Number of scans	No
	Colored scans (yes/no)	Yes
	GNSS data (yes/no)	Yes
MMS[8]	Number of scans	No
	Acquisition time	No
	Colored scans (yes/no)	Yes
	Acquisition of GCPs within the performed trajectory (yes/no)	Yes
	GNSS data (yes/no)	Yes
Survey completeness (e.g., interior and exterior environment, just the exterior, just a façade, just the roof, etc.)		No
Record of other metadata		
Record of paradata (e. g. pictures, sketches, ...)		Yes
Processing phase parameters		
Accuracy (RMS)		Yes
Performed editing (decimation, filtering, color balance, exposure balance,…)		Yes
Used software (name and version)		Yes
Point clouds properties	Total number of points	No
	Density	No
Other		No
Data storage		Yes
Operators’ level of experience		No
Connection to a wider survey network		No

Table 3. Information to be stored in OH3D for each project.

This table provides a comprehensive overview of key parameters essential to assess the quality and completeness of the archived 3D data, considering every step of the survey process, from the survey planning phase to the data processing phase. Parameters related to site access and environmental conditions provide context on the data origin, while parameters regarding the acquisition phase include technical and operational details about the field work. The processing phase's parameters provide insights into the accuracy of the obtained dataset and the editing techniques employed. Furthermore, information related to the software used for processing and the density of point clouds improves understanding regarding the dataset quality. Additional parameters consider other aspects of each dataset: understanding each dataset’s size beforehand is essential for improved data management; additionally, information about the dataset’s connection to larger survey networks can broaden its context and potentially increase its interoperability, allowing for more meaningful and extensive use of the data in the context of CH

[8] An MMS is defined as a mobile survey platform, based on SLAM (Simultaneous Localization And Mapping) algorithms, that integrates both mapping sensors (LiDAR scanner, cameras) and positioning (GNSS receiver, IMU platform) (Puente et al, 2013)

https://doi.org/10.5194/isprs-archives-XLVIII-2-2024-241-2024

documentation and preservation. Special attention is paid to the formatting of raw data inputs, with the understanding that these data can be reprocessed, fused, and reused in myriad unanticipated ways if their key attributes are properly preserved within common non-proprietary file formats. These variables can be used to evaluate the overall quality of datasets and assist users in deciding if a dataset is suitable for their needs.

3.2 Proposed Evaluation

OH3D is not an institutional platform for providing certified datasets, such as Cadastral or Cartographic institutions, but we endeavor to propose a scientific platform to share datasets and information enabling various kind of users to visualize and reuse the published datasets. This can be achieved thanks to a synthetic grade. The evaluation is conceived to prioritize datasets with metric-controlled data, and we envisage usage by professionals and researchers, but OH3D is open to other types of uses that do not request metric control of data, leaving low grades for them. Paradata and metadata are provided by the individual who propose the dataset and, though reviewed by OH3D curators, are self-certified. The OH3D team will control the proposed datasets and attribute the grades. Incorporating existing work models for the description of 3D processing workflows (Homburg et al,. 2021), the aim of the proposed research is the definition of best practices and operational methodologies that allow both to obtain complete and accurate documentation of CH sites and to share, within OH3D, 3D survey data for archiving and future reuse. The proposed evaluation here presented is a test and is not yet implemented on OH3D platform. The following basic classification will be employed to communicate data quality and potential for reuse to the users:

A – High-Quality Model with Accurate Georeferencing
These kinds of datasets meet every standard of a reliable metric survey. The scope of the project and the ownership of the data are outlined, and the data is made available within a unified global reference system. Georeferencing information is paramount to combine 3D models in the same reference system and to guarantee interoperability between surveyors and users. Considering that every data type of A-grade dataset must be georeferenced within a standard cartographic reference system and geodetic elevation, the EPSG code has also to be provided. The georeferencing is considered accurate only if is obtained with Real Time Kinematic (RTK) or Network Realt Time Kinematik (NRTK) Global Navigation Satellite Systems (GNSS) on board of the used sensor or if the data are processed using Ground Control Points (GCPs) measured with a total station and/or a GNSS receiver. All raw and processed data are included in common non-proprietary formats maintaining key data attributes, imaging and mapping methodologies are described in detail, coordinates of the GCPs and information enabling their location and reuse are provided. Detailed information is given concerning the calibration of cameras and other equipment. The resolution, precision and accuracy of the survey are declared and the representational scale (such as 1:100, 1:200, etc.) for reuse is provided with reference to the following formula (1).

$$\boldsymbol{Representational\ scale} = \mathbf{1}\text{: } \frac{\boldsymbol{precision}}{\mathbf{0.2\ mm}} \qquad (1)$$

All provided data, metadata, and paradata are certified by experienced individuals. This information is the most complete and has the potential to inform complex future expert users.

B – High-Quality Model without Georeferencing
The scope of the project and the ownership or the data are outlined, raw data are included in appropriate formats without additional corroborating data. Information regarding the dataset scale is provided, but in a local coordinate system without any georeferencing approach. Each dataset is to be evaluated within an isolated context, within the limitations inherent to the equipment and software used. Measurements performed across a large scale may possess significant inaccuracies and should not be used for further studies requesting metric control. However, these data can be used for visualization of the CH shape, dissemination, artistic uses, and any kind of work that do not require metric control.

C – Non-Metric Model
Raw data may or may not be included, perhaps in formats which lose key attributes enabling re-processing and fusion with additional data, provenance is poorly described, scale and positioning information is not included, and the accuracy of measurements cannot be relied upon. While the object is of interest and deemed worthy of preservation, any reuse of these data will require significant research effort on the part of the individual user.

The presence of a synthetic quality and completeness indicator is crucial for facilitating communication and dataset selection within the OH3D platform, emphasizing its role as a scientific platform for sharing 3D data in the context of CH. It is our hope that this framework will help OH3D users understand complex concepts related to 3D survey and help data contributors to optimize survey data utility for posterity.
We decided to keep the mandatory fields as lowest as possible, and provide a separate evaluation assigning a “+” to the datasets which have all the paradata and metatada fields filled out.

4. Case study and survey process overview

4.1 Salvation Mountain

In this paper we present a case study involving an important American folk-art site called Salvation Mountain (Figure 1), an area of approximately 9.5 square kilometers located in the middle of the Colorado Desert (California, USA).

Figure 1. UAV picture of Salvation Mountain.

An artificial slope with a large cross on top and several interior chambers are the core of the CH site (Salvation Mountain Inc, 2022). The entire complex is composed of natural and recycled materials (including bales of straw, pipes, telephone poles, ladders, branches, and full trees) covered in locally sourced clay and latex paint, which is periodically reapplied. In January 2024, Salvation Mountain was recognized as a place of historical significance by the California’s Imperial County (Matus and Brown, 2024). The site is open to the public and currently preserved by Salvation Mountain Inc, a non-profit entity that supports the compound's restoration and conservation. A rare tropical storm event in August 2023, Hurricane Hillary, caused significant flooding and erosion to the dry and dusty region. Though the artwork and main painted facade of the mountain suffered little damage, surrounding earthen support structures and access ramps suffered significant erosion. The site requires daily restoration efforts and monitoring to maintain safety, otherwise it risks public closure.

4.2 Data acquisition and processing

Survey details are described within the survey report (McAvoy et al., 2024) and here summarized. The survey campaign has been carried out on October 25th, 2023. The entire area has been documented integrating traditional and experimental Geomatics techniques. Employing a DJI Mavic 3 Pro, a multi-scale Uncrewed Aerial Vehicle (UAV) survey has been performed, acquiring both nadiral and oblique images to collect data regarding the entire site and the surrounding area. Terrestrial photogrammetry has been employed to document the indoor environment of the “Hogan” (a Navajo name for certain adobe structures), the northern nook, where the slope had collapsed, and all the trucks, tractors, and cars placed in the Salvation Mountain area. Also, both a TLS and a MMS approach has been adopted to acquire data concerning the overall complex, including the indoor part of the Museum. A total of 47 static scans have been acquired using a Leica RTC360 system, and as regards the MMS survey, 3 scans of 15 minutes each have been performed employing a Stonex 120^{GO} platform. A topographic and a GNSS survey have been performed for the metric control and georeferencing of the final 3D point clouds (Figure 2).

Figure 2. Topographic and GNSS survey.

The processing of the collected data has been done following the typical workflow provided for every type of technique, and all the final products have been integrated and georeferenced in the same reference system (NAD83(2011)/UTM Zone11, EPSG 6340). UAV dataset accuracy has been evaluated using GCPs and Check Points (CPs), while the RMS error on the Iterative Closest Point (ICP) alignment has been assessed for TLS and MMS datasets. The main processing results are reported in Table 4.

UAV survey	N° of images	N° of GCPs	N° of CPs	Mean RMS Error [cm]	
				GCPs	CPs
	1701	8	10	2.2	2.3
TLS survey	N° of scans	ICP RMS Error [cm]		Target Mean Error [cm]	
	47	0.5		0.022	
MMS survey	N° of scans	ICP RMS Error [cm]		Target Mean Error [cm]	
	3	1.2		-	

Table 4. Main post-processing results.

5. Results

The multiscale and multi-sensor survey of Salvation Mountain provided an opportunity to consider a complex dataset. An in-depth analysis and consideration of each facet of the data collection and processing pipeline has provided key examples upon which to build a meta/paradata schema to promote and describe the potential for data sharing and reuse. The final dataset related to this survey has been published on OH3D (Chiabrando et al., 2023) (Figure 3).

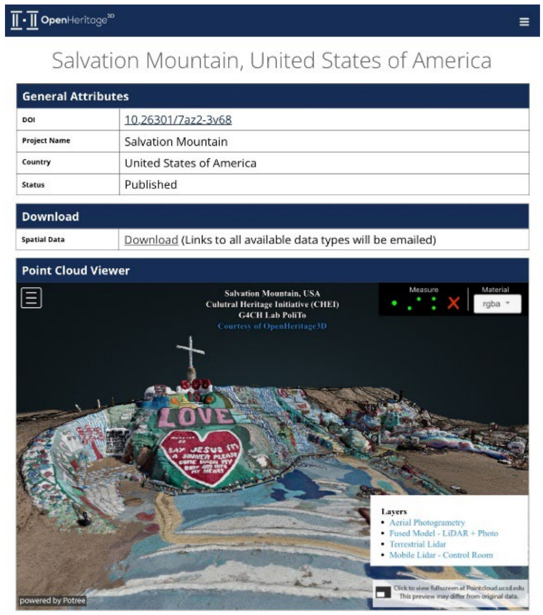

Figure 3. 3D visualization of the processed data within the OH3D platform.

The dataset includes both processed and raw data in commonly accepted open formats, accompanied by detailed paradata such as the documentation outlining the positioning of the control points. Furthermore, the dataset includes comprehensive information regarding the utilized equipment and the acquisition methodology. The processing phase followed the standardized and well-known workflow intended for each of the employed techniques: the Structure from Motion (SfM) software Agisoft Metashape has been used to process the UAV and the terrestrial photogrammetric data; the TLS scans have been co-registered and georeferenced within the proprietary software Leica Cyclone Register 360; and the MMS datasets have been post-processed in the GoPost software developed by Stonex. Each data type's accuracy has been assessed through a topographic network of control points, and every obtained Root Mean Square (RMS) error is provided to the users. Additionally, to improve the validation process, each data type is compared to others to prove every dataset reliability and to identify any potential errors and discrepancies. Also, a photogrammetric fused model is included (obtained by processing the photogrammetric dataset together with the TLS one in Reality Capture), flattening and interpolating reliable TLS data with high-resolution images from drone and terrestrial cameras. This approach ensures the synthesis of different data sources, enhancing the overall richness of the final model. By adhering to standardized protocols and embracing open data principles, these datasets exemplify a commitment to promoting transparency, accessibility, and interoperability within the CH community. It not only serves as a valuable resource for ongoing research and conservation efforts but also sets a precedent for future data sharing and collaborative initiatives within the CH field. Table 5 provides a description of the site, the environmental conditions, and general information regarding the survey. Within Table 6, Table 7, and Table 8 the description of the various datasets composing the final 3D model of Salvation Mountain that can be implemented on OH3D platform in the next future are summarized. After examining all the proposed parameters, an A grade has been assigned to each of the three datasets. According to the completeness of the information reported in the following tables, in this case, a + can be added to the assigned grade. As a consequence, we could assign Grade A+, this criteria will be followed for the other grades as well. According to the data acquired during the survey campaign Grade B and Grade C datasets were acquired and processed as well and are displayed on OpenHeritage3D.org[9].

Site/object general parameters	Value/description
Site access (e.g. official permission, remote areas, ...)	Site access is free, arranging with Salvation Mountain Inc has been necessary for the survey operations
Site dimension	9.5 km^2
Surface conditions (e. g. material, texture,…)	Adobe and latex paint
Environmental conditions (e. g. sunny, cloudy, …)	Mostly sunny
Operators' level of experience	Certified professional surveyors
Connection to a wider survey network	No connection to broader networks

Table 5. Site and survey general informations.

Dataset name	Parameters	Value/description
UAV photogrammetry	Sensor Name	DJI Mavic 3 Pro
	Number of images	1701
	Quality of images	5280 x 3956 pixels
	GSD	1.01 cm
	Mean acquisition distance	38 m
	Survey completeness	Exterior environment, surrounding area
	Metadata	EXIF files, report
	Paradata	GCPs position sketches
	Accuracy	Mean GCPs' RMS: 2.2 cm Mean CPs' RMS: 2.3 cm
	Software	Agisoft Metashape (v 2.0)
	Point cloud number of points	483,218,386 (filtered)
	Point cloud density	15,214 points/m^2
	Data storage (uncompressed)	38.1 GB

Table 6. UAV photogrammetric dataset evaluation.

Dataset name	Parameters	Value/description
TLS	Instrument name	Leica RTC 360
	Acquisition parameters	Medium resolution, 6mm @ 10m, HDR enabled
	Number of scans	47
	Colored scans (yes/no)	yes
	GNSS data (yes/no)	No onboard receiver, georeference provided by the Emlid Reach RS2 GNSS receiver
	Survey completeness	Most of the site, two chambers missing
	Metadata	E57 file
	Paradata	Target monography, sketch with each scanning position
	Accuracy	ICP RMS: 0.5 cm Target RMS: 0.022 cm
	Software	Leica CYCLONE REGISTER 360
	Point cloud number of points	937,039,369
	Point cloud density	300,116 points/m^2
	Data storage (uncompressed)	26.7 GB

Table 7. TLS dataset evaluation.

Dataset name	Parameters	Value/description
MMS	Instrument name	Stonex X120^{GO} SLAM Laser Scanner
	Number of scans	3
	Acquisition time	15 minutes each scan
	Colored scans (yes/no)	yes
	Acquisition of GCPs within the performed trajectory (yes/no)	no
	GNSS data (yes/no)	No onboard receiver, georeference provided by the Emlid Reach RS2 GNSS receiver
	Survey completeness	Entire site (excluded the parts that were not accessible)
	Metadata	Odometry, Trajectory
	Paradata	Target monography
	Accuracy	ICP RMS: 1.2 cm
	Software	Stonex GoPost
	Point cloud number of points	51,499,437
	Point cloud density	7,990 points/m^2
	Data storage (uncompressed)	24.3 GB

Table 8. MMS dataset evaluation.

[9] https://openheritage3d.org/project.php?id=pdsg-cb91 (retrieved 30th April 2024)

6. Discussion and Conclusion

Digital models are generated and shared online every day and today these data can reach new audience for entertainment, research, and education. However, often these data are not properly accompanied by documentation describing them in a useful and simple way, leading to improper use or avoiding their usage because are not reliable. Providing synthetic descriptors and evaluators for digital products of CH and bringing them together into a single platform provides a unique point of access, saving time for users. To address these challenges, this paper proposes a structure for archiving and sharing CH data, advocating for the adoption of standard protocols for data acquisition, processing, and archiving. By categorizing datasets into different grades based on the reliability of measurements, the framework aims to guide users in assessing the utility of CH data for researchers and professionals who need metrically controlled data (grade A), favoring the one with georeferencing. However, users that do not request high metric control or no metric control – like visualization for educational or museum purposes or for making digital arts – , can also use data with grade B and C. Overall, the proposed framework represents a significant step towards enhancing the documentation and communication of CH sites through 3D metric survey data. By promoting standardized protocols, comprehensive metadata and paradata documentation, and open data formats, the framework seeks to facilitate data use, support conservation efforts, and advance research in the field of CH preservation.

Acknowledgements

OH3D project has been made possible in part by the National Endowment for the Humanities: Democracy demands wisdom. Special thanks to: Kaycra Family Foundation; CyArk staff Scott Lee, Ilayda Bozhurt and Avidan Fernandez; Salvation Mointain INC, Bob Luvesque and Ron; AlertCalifornia project.
This publication has been possible partially thanks to the project PNRR-NGEU which has received funding from the MUR-DM 351/2022.

References

Chiabrando, F., Clark, L., Drisocll, J., McAvoy, S., Rissolo, D., Spreafico, A., Tanduo, B. (2023) Salvation Mountain - Photogrammetry - Terrestrial , Photogrammetry - Aerial , LiDAR - Terrestrial , LiDAR - Mobile , Survey Data . Collected by. Distributed by Open Heritage 3D https://openheritage3d.org/project.php?id=7az2-3v68

Champion, E.; Rahaman, H. 3D Digital Heritage Models as Sustainable Scholarly Resources. Sustainability 2019, 11, 2425. https://doi.org/10.3390/su11082425

Champion, E., Rahaman, H., 2020. Survey of 3D digital heritage repositories and platforms. *Virtual Archaeology Review, 11(23),* https://doi.org/10.4995/var.2020.13226

D'Andrea, A., Fernie, K. CARARE 2.0: A metadata schema for 3D cultural objects, 2013 Digital Heritage International Congress (DigitalHeritage), Marseille, France, 2013, pp. 137-143, doi: 10.1109/DigitalHeritage.2013.6744745.

European Commission, Directorate-General for Communications Networks, Content and Technology, *Study on quality in 3D digitisation of tangible cultural heritage – Mapping parameters, formats, standards, benchmarks, methodologies, and guidelines* – Executive summary, Publications Office of the European Union, 2022, https://data.europa.eu/doi/10.2759/581678z

Fernie, K. (2024). 3D Content in Europeana: The Challenges of Providing Access. In: Hostettler, M., Buhlke, A., Drummer, C., Emmenegger, L., Reich, J., Stäheli, C. (eds) The 3 Dimensions of Digitalised Archaeology. Springer, Cham. https://doi.org/10.1007/978-3-031-53032-6_10

Király, P., Stiller, J., Charles, V., Bailer, W., Freire, N. (2019). Evaluating Data Quality in Europeana: Metrics for Multilinguality. In: Garoufallou, E., Sartori, F., Siatri, R., Zervas, M. (eds) *Metadata and Semantic Research*. MTSR 2018. Communications in Computer and Information Science, vol 846. Springer, Cham. https://doi.org/10.1007/978-3-030-14401-2_19

Homburg, T., Cramer, A., Raddatz, L., & Mara, H. (2021). Metadata schema and ontology for capturing and processing of 3D cultural heritage objects. Heritage Science, 9(1), 1-19. https://heritagesciencejournal.springeropen.com/articles/10.1186/s40494-021-00561-w

Matus, D., Brown, M.B., (2024) Salvation Mountain is of 'Historical Significance,' a 'Historical Resource', Calexico Chronicle. https://calexicochronicle.com/2024/01/11/salvation-mountain-is-of-historical-significance-a-historical-resource/(4 April 2024)

Mcavoy, S. P., Tanduo, B., & Driscoll, J. (2024). Salvation Mountain 3D Survey - October 25th 2023. UC San Diego: Cultural Heritage Engineering Initiative (CHEI) at Calit2. Retrieved from https://escholarship.org/uc/item/96z5t8vc?

McAvoy, S., Ristevski, J., Rissolo, D., Kuester, F., (23 June 2023) OpenHeritage3D: Building an Open Visual Archive for Site Scale Giga-Resolution LiDAR and Photogrammetry, ISPRS Annals of Photogrammetry, Remote Sensing, and Spatial Information Sciences. X-M-1-2023 https://doi.org/10.5194/isprs-annals-X-M-1-2023-215-2023

Pamart, A., De Luca, L., & Véron, P. (2023). A metadata enriched system for the documentation of multi-modal digital imaging surveys. Studies in Digital Heritage, 6(1), 1–24. https://doi.org/10.14434/sdh.v6i1.33767

Puente, I., Arias, P. Review of mobile mapping and surveying technologies. Measurement, 2013, 46, 2127–2145, https://doi.org/10.1016/j.measurement.2013.03.006

Salvation Mountain INC (2022) https://salvationmountain.org (10 April 2024)

Sköld, O., Börjesson, L., & Huvila, I. (2022). Interrogating paradata. In Proceedings of CoLIS, the 11th International Conference on Conceptions of Library and Information Science, Oslo, Norway, May 29 - June 1, 2022. Information Research, 27(Special issue), paper colis2206. Retrieved from https://doi.org/10.47989/colis2206

Statham, N. (2019). Scientific rigour of online platforms for 3D visualization of heritage. Virtual Archaeology Review, 10(20), 1. https://doi.org/10.4995/var.2019.9715.

Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018 (2016). https://doi.org/10.1038/sdata.2016.18