

RESEARCH ARTICLE

Open Access



Metadata schema and ontology for capturing and processing of 3D cultural heritage objects

Timo Homburg^{1*} , Anja Cramer² , Laura Raddatz¹  and Hubert Mara¹ 

Abstract

Motivated by the increased use of 3D acquisition of objects by cultural heritage institutions, we were investigating ontologies and metadata schemes for the acquisition process to provide details about the 3D capturing, which can be combined with preexisting ontologies describing an object. Therefore we divided the 3D capturing workflow into common steps starting with the object being placed in front of a 3D scanner to preparation and publication of the 3D datasets and/or derived images. While the proposed ontology is well defined on a coarse level of detail for very different techniques, e.g. *Structure from Motion* and *LiDAR* we elaborated the metadata scheme in very fine detail for 3D scanners available at our institutions. This includes practical experiments with measurement data from past and current projects including datasets published at Zenodo as guiding examples and the source code for their computation. Additionally, the free and Open Source GigaMesh Software Framework's analysis and processing methods have been extended to provide metadata about the 3D processing steps like mesh cleaning as well as 2D image generation. Finally, we discuss the current limitations and give an outlook about future extensions.

Keywords: Ontologies, Semantic web, Data quality, Metadata, 3D scanning

Introduction

3D capturing of archaeological artifacts is steadily becoming a valued tool for artifact documentation. The reasons are the availability of reasonable precise entry-level i.e. low-cost capturing techniques typically using the principles of *Structure from Motion* (SfM) or *Structured Light Scanning* (SLS). Depending on the given task, we also see an increased use of industrial-grade high-resolution SLS, *Light Detection and Ranging* (LiDAR) and *Computed Tomography* (CT). For many years, the acquired 3D models were merely used to follow the rules of traditional documentation. So e.g. orthoimages or cross-sections were created, which enable certain compatibility with older publications. With the dawn of data science, machine learning, or the often

heavily emphasized artificial intelligence, we see larger 3D datasets having hundreds or thousands of 3D models published by modern projects¹ or open-minded institutions. These 3D datasets are typically created with a particular research question in mind or using a given hardware infrastructure. These common project-based approaches rarely provide metadata as this is commonly no requirement within these projects. However, there is an upcoming demand for mixing and merging collections of 3D datasets for reuse to answer research questions in larger contexts like machine learning of period classification of clay tablets [1] or enabling search capabilities of databases, i.e. shape based retrieval of 3D models [2]. Therefore, the interoperability of collections of 3D datasets becomes an emerging topic. While the focus is often on metadata about the objects themselves, little to no work has been done to define an ontology that enables interoperability at a technical level. For example, LiDAR

*Correspondence: timo.homburg@hs-mainz.de

¹ i3mainz, Hochschule Mainz, Mainz, Germany

Full list of author information is available at the end of the article

¹ Hilprecht Collection: <https://hilprecht.mpiwg-berlin.mpg.de> 3DHOP: <https://3dhop.net> Arachne: <https://arachne.dainst.org>.

devices have different metadata than the acquisition device for an SfM method, but the resulting data product, such as a point cloud or *Digital Elevation Model* (DEM), may be the same, so the detailed description of the origin of the 3D dataset is important. Therefore this publication proposes a metadata schema, a documentation process for metadata, and an accompanying ontology with a particular strong focus on 3D capturing and processing, which enables the adoption of our approach even beyond the heritage domain. Compatibility of our proposed ontology with CIDOC CRM² is possible.

While for 2D capturing and processing there are well-managed standards for describing metadata, such as the *Exchangeable Image File Format* (EXIF) [3] or the *Extensible Metadata Platform* (XMP)³, which are also available as linked data for the vast majority of imaging systems with optics, the situation for 3D capturing and processing is much more complex. The reason for this is a wider range of measuring techniques used by the capturing devices. The resulting outputs are also very diverse, e.g. regular grids such as voxels from tomographs, point clouds by *terrestrial laserscanning* (TLS), and irregular triangular grids as a result of optical systems such as SLS and SfM. There are some de-facto standards for file formats, which can hold very different resolution or topology datasets. Further confusion can be observed as synthetic datasets from Computer Graphics, and *Computer Aided Design* (CAD) used in architecture and manufacturing, in general, may use the same formats like 3D imaging. However, we will focus here on measurement data that documenting a real object and we will describe the workflow for creating metadata in reproducible steps. Using developed Python scripts, we will show how metadata can be retrieved in a structured way in exemplary acquisition and processing software.

The quality of 3D captured objects depends on capture methods, environmental conditions, object properties as well as individually chosen parameters during 3D capture and subsequent post-processing. Some properties can be taken directly from the final 3D model (e.g. the number of 3D points or the resolution of the 3D model), other parameters depend on the provision of metadata about the 3D model creation process. This includes metadata about the capturing device (e.g. technical properties, calibrations), geometric registration (e.g. global reference points, scales) as well as information about control elements (e.g. checkpoints, scales) [4, 5]. Information about the creation and post-processing of 3D models is

also included in the metadata, e.g. settings of filters. This information contributes to the trustworthiness of the data and thus also of the resulting 3D model. Examples of applications for metadata include the resolution of a 3D scan for pattern recognition tasks, post-processing such as closing holes to get the volume or for 3D printing. The parameters and settings for a 3D measurement and processing are always designed for a specific use case.

Therefore our metadata scheme contains three general categories of metadata: (i) basic information about the underlying project and the captured objects to be Incorporated from other metadata schemes like CIDOC CRM or Dublin Core. (ii) metadata captured during the measurement, which depends on the capturing device and its accompanying software. (iii) properties of the final 3D model, which can be computed for any given 3D dataset.

Besides the obvious example of quality concerns and acceptance criteria, many tasks are performed from capturing to publication, impacting the well-founded reusability of digital 3D models in future emerging research questions. With uniform metadata, one has the possibility to categorize or compare based on attributes. In analyses of 3D models, for example, the 3D point resolution of the scanners used or the prior application of filters can play a significant role in selecting the analysis parameters or in the interpretation of the results. Linked data as technology allows for the sharing of different such metadata definitions and to query a linked open data cloud of metadata for comparisons using, e.g., SPARQL [6]. A first and recent example for linked data assisting in 3D shape retrieval and therefore data re-use is [7]—a system including different semantic views on digitized archaeological objects such as vases, comprising a spatial, temporal, and shape similarity. With one or more of the above-mentioned metadata-enriched 3D captured objects, we will also enable or improve the anticipated merging or inter-linking processes of often decentralized 3D collections. Hence, by exposing metadata in a unified, standardized, interoperable, transparent and comprehensible way our work contributes a best practice of documenting scanning acquisition metadata as a prerequisite for a better quality assurance of 3D scans. In addition, as we show in our application cases, machines and people alike are enabled to assess and in the long run improve the confidence and reliability of 3D data using our unified metadata model.

Related work

While there are vast numbers of publications about capturing and analysis of 3D data and linked data concerning 3D datasets as the final product, there is little to no work about describing the process from data capture to final storage with all the intermediate steps with an ontology.

² <http://www.cidoc-crm.org>.

³ <https://www.adobe.io/open/standards/xmp.html#!adobe/xmp-docs/master/Namespaces.md>.

This is not to be confused with ontologies for 3D models - both captured and born-digital, as shown in the following subsections. Neither is it to be confused with norms and best practice guidelines for 3D capturing, which also exist in great quantities shown in this section. However, those guiding information is reflected on an abstract level in our suggested ontology and later on shown for specific examples acting as templates for other devices. This shall raise awareness to model preexisting procedures as linked data to integrate all those important details like 3D acquisition parameters to a digitized object's metadata. Most prominently, one can see the lack of metadata looking at file headers, e.g., from *Stanford Polygon* (PLY) [8], where typically only the last software package can be identified and all the other tools and especially applied processing settings are lost for the final 3D model.

We distinguish related work in the modeling of 3D data using ontological models, metadata vocabularies that describe the creation of objects, and finally, the representation of scanning environments, tools, and measurements in ontological models.

The CIDOC conceptual reference model

The CIDOC Documentation Standards Working Group's conceptual reference model [9] is a theoretical and practical tool for information integration in the field of cultural heritage.⁴ CIDOC extends the PROV-O ontology [10] and models properties of artifacts such as inscriptions, its temporal context, its spatial coordinates, its artifact classification, the artifact's condition, its curators, and current location (e.g. in a museum). The CIDOC reference model describes an artifact's provenance that enters a scanning process that is specialized for cultural heritage applications. Depending on the object's context, which is to be scanned, other ontology models to document this object may be of interest. In this sense, ontologies like CIDOC provide a basis of information which is available before a 3D scan is conducted. In addition, CIDOC provides an extension called CRMdig [11], which follows a similar approach to ours. CRMdig provides the classes to model a measurement device, a measurement, a person conducting the measurement, the digitization process and annotations added to the digital scan. Our approach uses similar classes to the CRMdig model, as clearly, this information is also relevant to be represented in our metadata model. However, we distinguish ourselves as follows: While the CRMdig model provides a class structure which may be used to model who used which equipment to scan what at which time and to describe which artifacts are being created, we want to find out

and capture which metadata is created in the different stages of data creation for the purposes of data quality assurance of 3D scans. Reference [12], while attempting a quality analysis on 3D cultural heritage replicas came to the conclusion that technicians conducting a 3D scan are often reluctant to manually collect metadata concerning the scanning process and that scanning equipment producers did at the time not provide sufficient opportunities to document the necessary metadata such as it is the case for 2D images using e.g. EXIF. However, this situation has changed in recent years, as scanning software developing companies allow to fetch metadata using an API. Furthermore 3D acquisition is increasingly used in mass acquisition of objects following pre-defined and quite well-defined guidelines, which were helpful for relevant definitions in our proposed ontology.

This allowed us to formalize the scanning metadata in our contribution similar to EXIF. It is important to formalize and capture these common metadata parameters per scanning device and its accompanying software so that comparisons between digital artifacts can be conducted on a unified basis. Although there are numerous manufacturers the scanning method itself follow well-known principles and parameters. We chose a bottom up approach to cover generic parameters by creating ontologies for devices using different methods. This allows to extend the ontology for different devices by adding only the device specific metadata fields.

Comparing sets of metadata from different devices, once formalized in a singular model can be conducted using reasoning approaches, similar to the ones suggested by [13] for metadata of a captured object and a part of its scanning process. Therefore, in this publication we provide the means of achieving this while at the same time not sacrificing compatibility to the CIDOC CRM and CRMdig model.

Representing 3D data using ontologies

Vasilakis et al. [14] defined the first ontology, including multidimensional shapes of 3D scans considering different types of scanning processes. They tested this ontology model and its adaptability to different knowledge domains using the two use cases of human shape capturing and during a product development process involving 3D modeling. Mi and Pollock [15] describe a case study at the University of South Florida to create a metadata schema for 3D cultural heritage objects. This endeavor aimed to gain global exposure to the 3D cultural heritage objects and connect them to other repositories of the same kind. We share this same goal of global exposure and interconnectivity, but our focus lies in providing information about the scanning and post-processing of the 3D scanning data, which can also be useful outside

⁴ <http://www.cidoc-crm.org>.

the cultural heritage domain. The CARARE 2.0 metadata schema [16] is another approach of providing information about digital representations of archaeological objects. However, as all previously stated approaches, it lacks essential information about the 3D acquisition process, which we tackle in this work.

Ontologies for 3D scanning tools

The *Colour & Space in Cultural Heritage* (COSCH) ontology [17] was built to give suggestions to archaeologists which scanning equipment could be used to produce an appropriate 3D representation for a given artifact. To achieve such an assessment, the ontology model included properties of artifacts that may hint towards the usage of one scanning technology over another. The intent of our work is to precisely capture metadata for validation, quality control and related matters, which can help in improving the recommendation system of COSCH-KR, which partially overlaps with our ontology as it has some rudimentary data fields to capture some information e.g. of a 3D scanners calibration. Within our approach we can incorporate information from COSCH-KR like artifact properties, which were of no concern of our ontology as we assume a proper choice of the used device. COSCH-KR also allows limited modeling of environmental conditions when a 3D scan is conducted. The motivation of COSCH-KR in doing so it to give a recommendation for the usage of certain measurement processes. For our work, this modelling of environmental conditions can be an input to our metadata model and a parameter in a data quality assessment approach.

Projects related to 3D objects in cultural heritage

Acquisition and analysis in *Cultural Heritage* (CH) is an ongoing research field of its own [18]. From this vast number of projects, the relatively closely related projects are found as European collaborations. The goal of ARIADNEplus as described on their homepage is “*developing a Linked Data approach to data discovery, making available to users innovative services, such as visualization, annotation, text mining and geo-temporal data management.*” Predictive digitization, restoration and degradation assessment of cultural heritage objects (PRESIOUS) developed predictive geometric augmentation technologies for auto-completion for 3D digitization, estimation and prediction of monument degradation and 3D CH object repair. PRESIOUS sparked national projects funded in Austria integrating linked data into respective machine learning approaches [7]. The project *Geometric Reconstruction And noVel semantIc reunificaTion of culturAl heriTage objEcts* (GRAVITATE)⁵ was an approach

to virtually unify collections using 3D models. They proposed an algorithm for making annotations of triangular algorithms robust against topology changes – typically reduction of resolution [19]. A more specific project focused on pottery fragments with overlaps to the previous projects was *ARCHaeological Automatic Interpretation and Documentation of cEramics* (ARCHAIDE), which aimed to create a new system for the automatic recognition of archaeological pottery from excavations around the world [20].

Arachne⁶ is a database by the *German Archaeological Institute* (DAI) and the *Cologne Digital Archaeology Laboratory* (CoDArchLab) which provides metadata on thousands of archaeological objects. This metadata is currently not available in a controlled vocabulary. Instead, only a gazetteer of archaeological terms⁷ is available. However, as being part of the internationally active DAI it is home to a wide variety of information of objects, including 3D datasets like fragments⁸ from the excavation in Honduras [21] having implemented a 3D acquisition workflow as outlined in the next Section. The DAI also hosts a national research data center providing recommendations for storage of 3D datasets.⁹

Metadata standards

The *Dublin Core Metadata Initiative* (DCMI) [22] defines a metadata vocabulary to describe arbitrary datasets. Dublin Core metadata distinguishes between technical metadata concerning the data format, type and language used in the dataset. It furthermore describes the content, creators, legal rights, its lifecycle and relates the respective dataset to other resources. Common usages of Dublin Core can be found in HTML metadata but also in the Semantic Web to describe metadata of RDF-encoded data [23]. Links to metadata described using the Dublin Core Vocabulary will be used in our ontology model.

Research software metadata

Software and, in particular, research software should be represented in the metadata scheme to the extent that the software can be identified, classified and judged according to their authors’ provenance information. First steps in this direction have been taken by Garijo et. al.¹⁰ who propose a vocabulary to represent research software metadata. Considering that a toolchain needs to be

⁵ European Commission Horizon 2020 project, Grant agreement ID: 665155 <https://cordis.europa.eu/project/id/665155>.

⁶ <https://arachne.uni-koeln.de>.

⁷ <https://gazetteer.dainst.org/>.

⁸ e.g. <https://arachne.dainst.org/entity/6886054>.

⁹ IANUS IT-Empfehlungen 3D – <https://www.ianus-fdz.de/3d>.

¹⁰ <https://sorser.github.io/programme/talks/event-010/>.

represented when assessing a 3D scan's provenance, this vocabulary will be linked to in our final ontology model.

Definition of measurements

Metadata, as defined in the previous section and captured by, e.g. Dublin Core, [22] provides a kind of metadata that cannot be derived from the processed 3D model. However, a lot of different information may be derived by analyzing the 3D scan itself either during the process of the 3D scan or in a post-processing step. To capture those measurements, vocabularies and ontologies like the Ontology of Units of Measurement [24] have been developed. Measurements provide the basis for a classification and data quality assessment of 3D objects which is one of the results one might obtain from a properly documented scanning process.

Data quality

Data quality is defined in many ways in the literature. We refer to the ISO8000 definition, [25] which defines data quality as

Quality is the degree to which a set of inherent characteristics fulfills requirements.

Data quality can be measured by data quality metrics which are grouped and described by data quality dimensions [26]. To define the quality of a given data set for a given application case, an often prioritized set of data quality metrics needs to be evaluated, a common data quality score needs to be defined, and finally interpreted to be able to result in an informative data quality statement. Reference [27] suggests that these data quality definitions and the requirements associated with a data quality statement may be modeled in an ontology model. Clearly, a goal of capturing metadata is to infer information about the quality of the respective 3D scan, therefore our ontology model should allow for the definition of data quality statements.

Acceptance tests

Validation of 3D models is, in general, a hard task as it is strongly coupled with the definition of quality. This topic was addressed in [28] to propose merely considerations on verifying 3D scans trying to define acceptance tests for valid 3D scans. However, we can derive that verification steps and acceptance tests rely heavily on provided metadata as we consider to provide as an ontology within this publication.

For specific application cases it is possible to state the required data quality and create acceptance tests for the state of 3D models. In this sense, acceptance can be

formalized as data quality definitions as defined in the previous section [27].

To enable acceptance tests and foster reproducibility of results a detailed and structured recording of metadata is crucial. Reference [29] discusses the reproducibility of recording and processing scenarios using a confocal laser scanning microscope with two different objectives with the same magnification for documenting surface traces on archeological tools. The influence of the setup of an acquisition device in relation to the results is reviewed to demonstrate the method's acceptability. The calculated surface parameters for the analysis of the results are taken from ISO 25178-2 [30] and statistically evaluated. Without an appropriate metadata acquisition analysis such as the one conducted in [29] are not possible.

Semantic Web technologies provide means of defining *Shape Constraint Language* (SHACL) constraints on RDF graphs - rules that may be evaluated automatically to verify if acceptance criteria have been met. Results of such verifications might be added to the Semantic Web graph as new statements to signify acceptance for certain application tasks [31]. In fact, many of the use cases proposed by the W3C Working Group on SHACL are concerned with the verification of data quality¹¹. Once sufficiently many acceptance criteria have been defined by research communities, the formalization in SHACL can lead to an automatic acceptance classification for several use cases, e.g., adding a statement to express the eligibility for 3D printing once data has been entered into a semantic database.

Processing stages and metadata collection

To be able to introduce our ontology model for the processing of capturing 3D models, we first have to define the processing steps beginning from the artifact which is to be scanned to different metadata sets which are to result from our documentation process. In general, there are different approaches to create a digital 3D model from an archaeological artifact. In this chapter, we focus on non-contact surface documentation with optical measuring devices. A survey of the state of the art of the relevant 3D scanning methods can be found in [5, 32]. To create metadata which is compatible with these state of the art methods we define processing phases in the metadata collection process. These processing phases exist in any of the aforementioned scanning methods, even though their content might vary according to the method of scanning being used. Figure 1 shows the proposed workflow, which we separate into:

¹¹ <https://www.w3.org/TR/shacl-ucr/>.

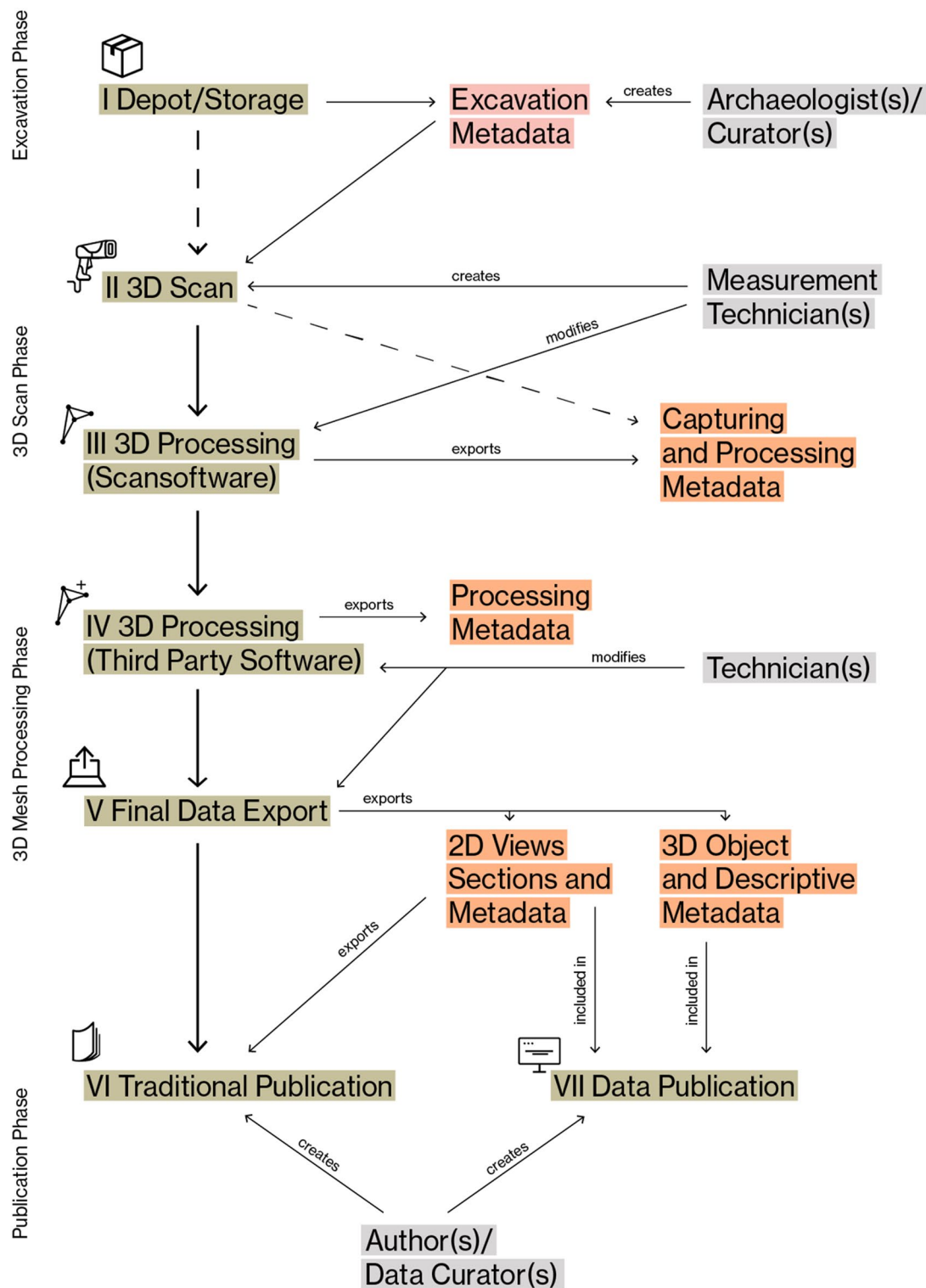


Fig. 1 Scanning workflow representing the respective stages, the involved persons and the metadata which is created at each of the seven stages during the four phases

- I Cultural heritage object from depot/storage
- II 3D Data capture
- III 3D Processing (Scan software)
- IV 3D Processing (Third-Party Software)
- V Final data export
- VI Traditional publication
- VII Data publication

Each of the processing phases produces a set of metadata that may be exported as a separate metadata file or handed to the next processing phase to be further enriched. It is not mandatory to start the proposed workflow at the first processing phase. The workflow can be entered at any step and exited at any following step to also achieve intermediate results. In simple terms, in every stage, data is imported, processed and may be exported. Processing should be described and stored in structured metadata vocabularies at each stage of the workflow. For optimal documentation, we recommend the execution of all steps in the workflow description if possible.

Cultural heritage object

In the first stage of the processing phase, the artifact to be scanned is to be documented. This in itself may be a process, including several stages. For example, an artifact might be obtained from an excavation, indexed, moved to a museum, curated, and finally, at some later point in time, prepared for the initial scan in a controlled environment. Metadata concerning the artifact might include information about its physical creation, such as its creation time, the creator, and the creation process, as well as its material, historical context, and possibly inscriptions or important features attached to the artifact. In our approach, it is important to incorporate such information - if available - however, our ontology model does not describe how this information should be described. We rely on previous publications such as CIDOC-CRM to describe artifact information to link them into our ontology model. However, information about size, material, and texture can be very relevant for selecting the appropriate scanning method or to retrospectively judge if the object has been scanned with an appropriate scanning method. The selection of an inappropriate scanning method is very likely to result in a bad quality of the 3D scan.

3D data capture

The next stage is the 3D data capturing of a cultural heritage object. This process is dependent on the technique of measurement being used, the measuring setup, the person operating and supervising the capturing process, and the capabilities of the scanning software, which digitizes the output of the scanner. Information about this

process should be stored as metadata to document the origin of 3D models. However, some of the information is influenced by external factors, which cannot be captured automatically in the metadata model. For example, the selection of the equipment for capturing 3D data depends on several factors, including the digitization purpose (how precisely do we need to scan?), which resources are available (scanning tools), and what the environmental conditions are like (e.g. laboratory conditions, outside scanning site). If this information is available, it needs to be manually added to the metadata given. If this information is not available, a user investigating the dataset needs to judge if the quality of the 3D scan is sufficient for his/her usecase to be tackled. Setting up a measurement concept is essential to achieve the goals. Examples of applications can be found here [33, 34]. Another factor is the person taking the measurements. Here individual decisions can be made on the settings that influence the measurement.

This processing stage's metadata is the information of the acquisition situation and cannot be changed in the 3D processing of the scan data. Examples of metadata captured here include: The sensor type, the acquisition time, the person who scanned, the focal length and the exposure time.

It might also be natural to capture information about the environment in which the 3D scan has been conducted. Has the scan been conducted in a laboratory under controlled conditions, in the context of an archaeological excavation or under other conditions? Automatically extracted metadata may only provide information captured by the scanning software, for example the light intensity used for scanning. This can only give indirect information about the scanning environment. However, the COSCH-KR ontology model can be used to model the scanning environment, so that the information can be injected into our ontology model covering the scanning process until publication.

3D Processing (Scan software)

The 3D processing as the second step in the 3D Scan Stage (Fig. 1) includes the processing of the raw capture data up to the creation of the 3D model. This includes cleaning the 3D point cloud from the single measurements, the alignment of the individual scans to each other, and the polygonization. No filtering or manipulation of the 3D mesh is applied in this stage. This process is also possible at a later time, e.g., in the office. The 3D model can only contain the areas that were captured by the 3D scan. Areas on the object that are not available for capture are not included in the 3D model, e.g., the floor may be covered by the object itself or the back by a wall. 3D processing is usually only possible in the scanning

software that comes with the scanner. However, for some acquisition methods, a variety of software can be used to calculate 3D models from the acquisition data. The metadata in this stage contains all the settings for the calculation of a 3D model. The metadata file and the 3D model are the results of this processing stage.

3D Processing (Third Party Software)

After scanning the artifact and the first processing using the respective scanning software, the resulting 3D model often undergoes various further processing steps. For example, 3D models may contain holes that are only closed afterwards. The 3D model can be realigned or reshaped. These possible further processing steps are carried out depending on the intended application. During all processing steps, further metadata is created, which is ideally also saved. The processing can be done with different software, for example, GigaMesh or Geomagic Wrap¹². The special feature of a currently in development version of GigaMesh¹³ is that the resulting metadata proposed in the publication can already be saved during processing. The result of this step is another 3D model and a metadata file containing the metadata of all the processing steps.

Final data export

The final data export includes a version of the 3D object, which has been optimized for a specific purpose. This means that the processing chain, which leads to creating this version of the mesh, has ended and the metadata associated with this given version of the mesh can be considered complete. At this stage, the metadata generated is ready for publication but may still undergo a stage of interpretation, i.e., a stage of reasoning to calculate data quality metric results and/or scores.

Traditional publication

After the final data export we expect a multitude of analysis methods applied to the 3D acquired object leading to an optional traditional publication. These analysis are depending on the object as well as driving research questions. So we expect according ontologies to be used. However, with tools like GigaMesh we can contribute technical metadata like rendering resolution and position of 2D screenshots making those images reproducible. Therefore our meta-data enriched renderings can be seen as an exported representation related to the 3D model and should be documented as such in an accompanying digital publication.

Data publication

Similar to the optional step of traditional publication is the data publication we provide detailed technical metadata fields helpful for future reuse of the enriched 3D models. Such publications of data [35] and software [36] have only recently gained more attention in the scientific community. This novel kind of publications should be in such a way that is uniquely retrievable and described using established vocabularies. This requires data publications to be documented by a unique identifier such as DOI [37] and to be documented with metadata such as we present in this publication. 3D models and their derivatives, such as screenshots, should be accompanied by the whole chain of metadata that we have described previously. Metadata of 2D images derived from 3D scans should contain all metadata of the previous stages of the 3D scan and the parameters used to create the 2D rendering as well as the 2D renderings metadata (e.g. EXIF or XMP).

Ontology model

We extend the W3C PROV-O Ontology Model in Fig. 2 to describe the relevant elements of a 3D scanning workflow. The workflow structure will stay the same, but depending on the type of scanning, metadata associated with the scanning workflow elements may vary. Metadata of each previously mentioned stage is stored in metadata graphs linked to the respective revisions of the different generated digital artifacts. We need to note that this workflow does not include an archiving and preservation phase. Archiving and preservation of objects and datasets are arguably highly relevant and long phases, which happen after and/or at the end of the outlined workflow for the digital objects we create and before our workflow for the physical objects to be documented. However, our model is focused almost solely on the data capturing process and therefore does not include this kind of metadata. Still, our model contributes to the long term matters of digital preservation of 3D scanning information by capturing this information in the first place. In the following, we describe the main elements of the ontology model to give an overview of the scope of the to-be-documented data. The complete ontology model until the time of writing this paper is published on Zenodo¹⁴, the in-development version of the ontology model can be seen here¹⁵.

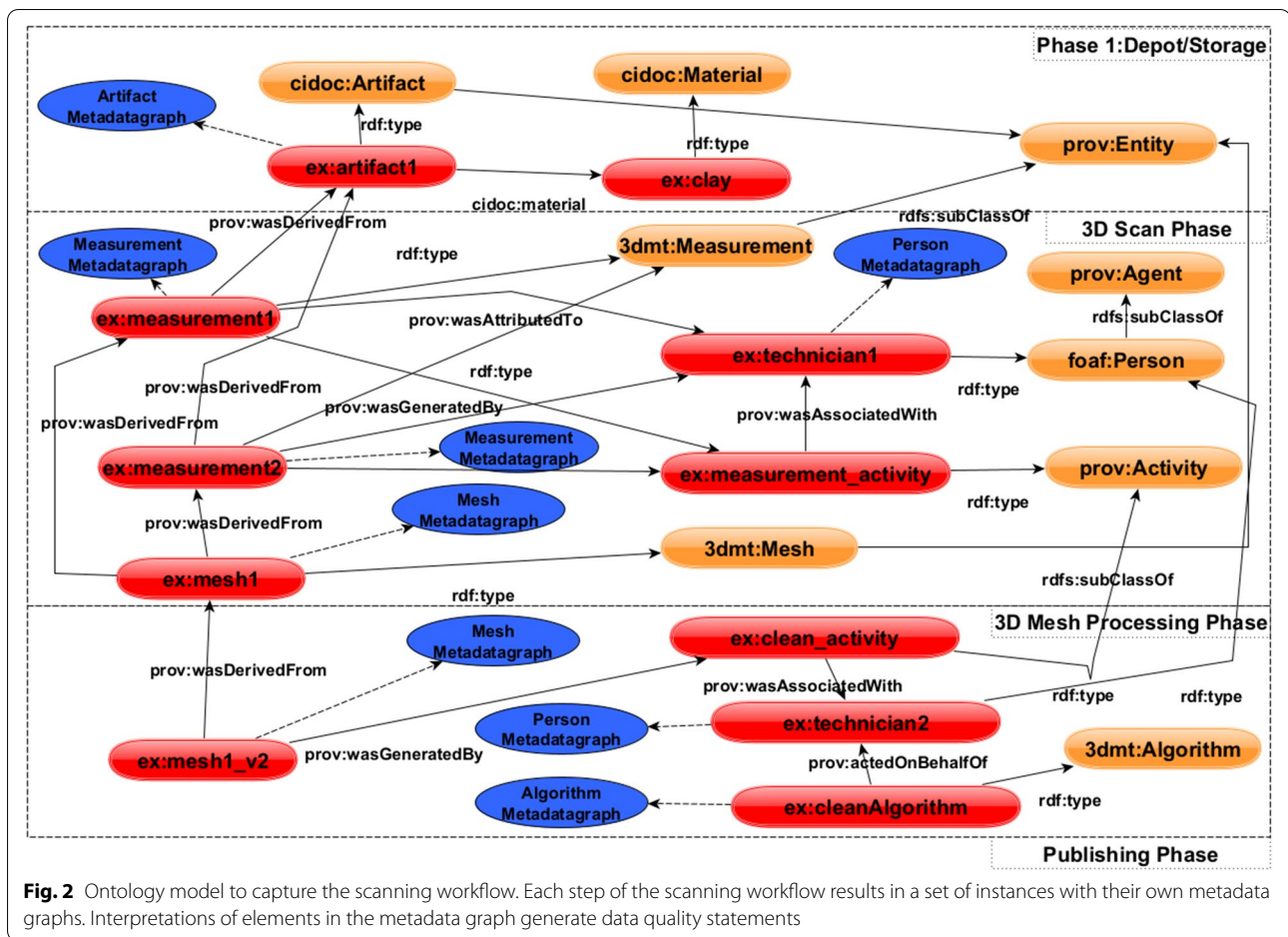
Entities We subclass the concept of a *prov:Entity* to describe results and intermediate results of the scanning process using a hierarchy of semantic concepts. In this way, we capture the algorithms, measurements, people,

¹² <https://de.3dsystems.com/software/geomagic-wrap>.

¹³ <https://gitlab.com/fcgl/GigaMesh/-/tree/develop>.

¹⁴ <https://github.com/mainzed/mainzedObjectsOntology>.

¹⁵ <https://mainzed.pages.gitlab.rlp.net/homepages/mainzedmetadata/index.html> or <https://mainzed.github.io/mainzedObjectsOntology/>.



tools and activities leading to each version of the final 3D model and its previous versions. Figure 2 shows an example of a scanning process involving two measurements in the 3D Scan Stage. The result of this stage is a mesh that is further cleaned in the 3D Mesh Processing Stage. Each subclassed *prov:Entity* instance is assigned a metadata graph that contains the captured metadata for the respective entity.

Agents Agents describe algorithms and persons that are involved in the creation and modification of the 3D model. We differentiate persons by the roles and responsibilities they take in creating the 3D model. At first, the artifact which is to be scanned needs to be documented by a specialist, e.g., an archaeologist or a curator in a museum. This documentation may also involve the people taking part in the process of artifact acquisition, such as an excavation at a particular site. Next, the artifacts are scanned by a measurement technician in a technical setup created by this technician and likely adjusted for every object to be scanned. One or many

possible post-processing steps might follow by the same or another technician before an eventual traditional and/or data publication by responsible people is conducted. In general, sufficiently classified agents may give information about the competence or suitability of people or algorithms for the scanning task, i.e., if the scan was conducted by a professional or a student and if the used algorithm is capable and/or has been rated as well-suited for the scanning task. Finally, agents may be linked to information concerning the context in which the scan was conducted, i.e., a research project or a research institution that may help assess the competence of said individuals.

Algorithms Algorithms describe processes used to manipulate or create an entity as described in “Ontology model” section. We distinguish algorithms for camera calibration, geometric modification algorithms such as merging of nodes or cleaning meshes, and measurement algorithms that produce measurements included in the metadata export of the created entity. All described

algorithms are modeled as instances of subclasses of *prov:Agent*. Often, algorithms are not described publicly, for example, if they are part of a proprietary software package. In that case, we may only refer to the respective software version and describe parameters used by the algorithm to retrieve the respective result in an execution description.

Activities Activities derived from *prov:Activity* are actions with an assigned start and end *xsd:datetime* which are initiated by a person or algorithm and produce one or many new or modified entities as a result. Activities in the scanning process are beside the scan, the scanner calibration, the scanning environment setup, and any further processing steps by a third-party software.

Tools As tools, we describe scanners, calibration equipment and software used to create, modify and export 3D scans and their metadata. Tools are connected to activities that make use of these tools to achieve the activity's goal.

Metrics Metrics are results of calculations or parameters of an algorithm considered adequate to be represented in the metadata description. Metrics may also be calculated from other metric results or be created using reasoning approaches inside the triple store. Metrics build the foundation of interpreting the metadata of the given scan under various user perspectives - to judge the fitness for use of the scan for given use cases.

Interpretations Interpretations of metrics describe the suitability of 3D models to certain fields of application. In essence, they are represented by additional statements about the 3D scan which are added to the ontology model. An interpretation consists of one or more prioritized metric results with an expected value range and one or more mathematical functions for calculating an interpretation score by aggregating the given metric results. Interpretations may exist explicitly in the given metadata due to the scanning software (e.g. if the object is waterproof). However, interpretations may also be derived from the ontology model by applying reasoning rules.

Implementation

We describe a toolchain that implements the creation of the metadata we described earlier. To illustrate the metadata collection, we focus on the data collections available at *Institute for Spatial Information and Surveying Technology, University of Applied Sciences Mainz* (i3mainz)¹⁶ and the *Römisch-Germanisches Zentralmuseum - Leibniz Research Institute for Archaeology* (RGZM)¹⁷. The focus

in the following section is on the data collected with Structured Light Scanners and software versions from GOM (a company of the ZEISS group). The goal is to create digital surrogates of real artifacts for object documentation and to conduct analysis on the object. The measurement projects' files are available in proprietary formats and can only be viewed with special software.

Metadata transfer

When defining a metadata schema such as the one in our publication, it should be ensured that this metadata can be displayed, extended and exported by various software implementations. While we cannot provide implementations for all concerned software, we can show how the data transfer between software can work.

At first, metadata could be provided as a separate file or in the header of the given 3D model. Both representations are not the norm at the moment and cannot be processed by commonplace 3D scanning software. Metadata provided as separate files might be accessed separately from the 3D model and be stored separately in open data portals. Metadata stored in the header of the 3D model is more likely to be used because one cannot forget to provide an additional metadata file alongside the 3D model file. Storing the metadata as a side-car file is an option, but changes to the filename or -location will break the bond between 3D- and meta-data. Of course, there is – in contrast to the older widely used formats – a novel XML-based X3D definition [38]. However, the comprehensive X3D standard is typically not fully implemented, making the embedded use of metadata equally prone to being discarded. Ultimately the role of a research data manager has to be assigned to a person responsible for the integrity of meta- and 3D-data. This also affects the metadata collected in the following two Sections.

Secondly, software needs to support the extension of 3D metadata by converting their processing information into the metadata representation we propose. This requires scanning software to provide data using an API that may be accessed by a customized script or the implementation of a customized export function to our metadata schema. We provide example implementations in this publication.

Finally, the question of the metadata format arises. Because our metadata schema should also be exportable as linked data and may form the basis of data in a triple store, we choose the TTL format [39] as linked data serialization. The format is text-based, can be easily extended and is convertible to a JSON representation such as JSON-LD [40]. Non-linked data-aware software implementations may also process JSON-LD.

¹⁶ <http://i3mainz.hs-mainz.de>.

¹⁷ <https://web.rgzm.de>.

Artifact information and provenance

While it is not part of this publication to describe a process to document the artifact itself, we can refer to best practices in the community to build such descriptions. [9] In general, we expect artifact information and provenance to be served as a linked data representation, e.g., TTL, in order to process it further using our toolchain. We are aware that this is currently not the standard in all areas and therefore refer to best practices of mapping non-semantic datasets to linked data and the respective tools to do so. [41–45] Besides, we define the minimum result of the first stage of our toolchain, describing an artifact by one artifact ID plus corresponding namespace, a label, and an optional OWL class [46] describing the type of the given artifact. This minimum representation needs to be represented as a TTL file but may be given as a CSV text file to the next tool in the processing chain. A mapping from this CSV to TTL is trivial because all relations (rdf:type, rdfs:label and owl:Class) are already predefined. This allows artifacts for which no metadata except for a unique identifier and classification has been collected to be sufficiently represented in linked data.

Metadata export scripts

The 3D Scan Stage of the workflow in Fig. 1 requires the collection of metadata from the respective scan software. This includes all settings and parameters applied for measurements and processing in the respective measurement project. Depending on the scanning software, customized scripts need to be developed which access the software's API or a software data export, including the necessary metadata. As proof of concept, we developed Python scripts for the scanning software GOM Professional 2016 and ATOS version 6.2. The scripts make use of the scan software's Python API and accesses the data elements stored in the measurement projects, such as individual scans, their alignment to each other, reference points used, and, if applicable, 3D models derived from them and their associated metadata.

Every script maps the metadata received by the scanning software to the unified vocabulary defined by the ontology model. The validity of these mappings has been verified by a group of experts in the field and by consulting the respective scanning software manuals. The storage of metadata is currently structured in two standardized, software-independent file formats and serves two purposes.

Firstly, they should serve as metadata representations that may be processed by third-party applications and possibly offered for download in open data portals. Secondly, they should be prepared in such a way that they

can be easily integrated into a semantic database and interconnected to other resources of a similar kind. This gives the possibility of querying metadata of 3D scans to select appropriate datasets for specific purposes. The first purpose is served by a JSON [47] representation of the given metadata. The second purpose is served by a *Turtle* (TTL) [39] representation or a possible JSON-LD [40] serialization of the given TTL data.

The scripts and related documentation are published online.¹⁸ An additional toolset for metadata extraction from the software Agisoft Metashape¹⁹ as prominent package for *Structure from Motion* (SfM) is work. Via the Metashape Python 3 module, the metadata from the recording and processing are extracted and exported in a structured way.

GigaMesh software framework

To highlight the post-processing in third-party software, we chose the GigaMesh Software Framework [48]. For this publication, we have extended GigaMesh to produce metadata according to stages 4 and 5 of the scanning workflow. In a next iteration, GigaMesh will be extended to also process previously created TTL files formatted according to the defined ontology model. In that way the software will be capable of continuing the given provenance history by logging the modifications to the given 3D model that are applied in the GigaMesh software. (stage 4) The exported 3D model or exported screenshots are accompanied by a TTL file encompassing all previous stages and the latest modifications applied in GigaMesh. They may be further processed in the publication stage.

Publication

In the data publication stage, 3D models, possible 3D model screenshots, and their metadata need to be published in, at best publicly accessible, persistent storages. The metadata reflects these changes by adding a *Digital Object Identifier* (DOI) and further bibliographical information to the previous stage's data. Further, the metadata may be published along with the 3D models, but in the case of linked data, representations might be better suited to be published in a thematically fitting SPARQL endpoint.

Application example

We illustrate the application of the ontology model using three examples of 3D scans of different scanning projects and show its usefulness using the following application cases:

¹⁸ <https://doi.org/10.5281/zenodo.4566044https://github.com/i3mainz/3dcap-md-gen>.

¹⁹ <https://www.agisoft.com/>.



Fig. 3 View of the largest 3D model with a length of 3 m (the wine ship); Structured light projector/sensor type: ATOS III Rev.02 with a theoretical point resolution of 0.4 mm. The original is in the RLM, inventory number 767

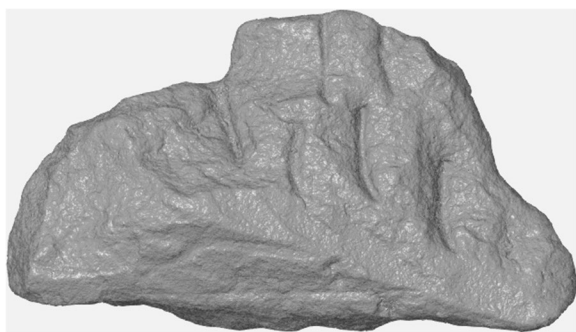


Fig. 4 View of a small 3D model with a length of 14 cm; Structured light projector/sensor type: ATOS III Rev.02 with a theoretical point resolution of 0.05 mm. The original is in the RLM, without inventory number

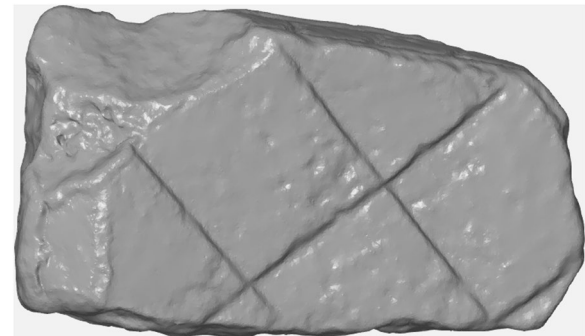


Fig. 5 View of a 3D model with a length of 50 cm captured with a lightweight hand-held scanner, the Faro Freestyle 3D X with a theoretical point resolution of 1 mm. The original is in a private collection in Neumagen

Application case 1: Documentation of scanning projects

The first application case illustrates the general usefulness of a uniform vocabulary for metadata using the example of two research projects for Roman burial monuments in Trier/Germany, with the 3D documentation of artifacts being one central part. The first project, “Römische Grabdenkmäler aus Augusta Treverorum im überregionalen Vergleich: mediale Strategien sozialer Repräsentation”²⁰, ran from 1 May 2016 to 31 March 2019. The second research project, “Grabdenkmäler aus Augusta Treverorum, digital vernetzt”²¹, ran from 1 May 2018 to 31 December 2020. The *Institute for Spatial Information and Surveying Technology, University of Applied Sciences Mainz (i3mainz)*, the *Goethe University Frankfurt/Main* and the *Rheinisches Landesmuseum Trier (RLM)* were involved in both projects.

A total of around 400 archaeological objects were 3D scanned, for which access must be guaranteed in order for them to be reused. In the above-mentioned projects, this means that the data is stored and published in Arachne²² (“[Projects related to 3D objects in cultural heritage](#)” section) and can be accessed via a web interface. In addition to the 3D models and their metadata, information on the measured object (“[Cultural heritage object](#)” section) is also stored in Arachne, such as the find location and the object’s material, and linked to the 3D models. Because of the size of the objects and the environmental conditions, different scanners with adapted settings were used for the 3D capture (“[3D data capture](#)” section). The information about the equipment and the capturing settings is especially indispensable for the evaluation of the 3D models. Figures 3 and 4 show views of 3D models captured with the same scanner but with different measurement volumes. The 3D model in Fig. 5, on the other hand, was captured with a hand-held scanner. Thus, each of these 3D models have a different resolution.

Much of the metadata for 3D acquisition is contained in the respective scan project files of the scanning software but not in the file of the 3D model. Not only during acquisition but also during post-processing, settings are made, and processing steps are carried out that would no longer be comprehensible without metadata, such as closing holes and cleaning up mesh errors. Therefore, it is important to include all metadata with the 3D models, which provides the user with valuable information for future applications. For example, when analyzing handling traces on the object, it is important to consider the resolution with which the object was captured and the settings with which the resulting 3D model was created and to check whether the 3D model has been

²⁰ DFG GEPRIS ID: 279232664 <https://gepris.dfg.de/gepris/projekt/279232664>.

²¹ <https://portal.wissenschaftliche-sammlungen.de/CollectionActivity/182765>.

²² <https://arachne.dainst.org/>.

Exploring the quality of cuneiform tablet 3D Scans with SPARQL!

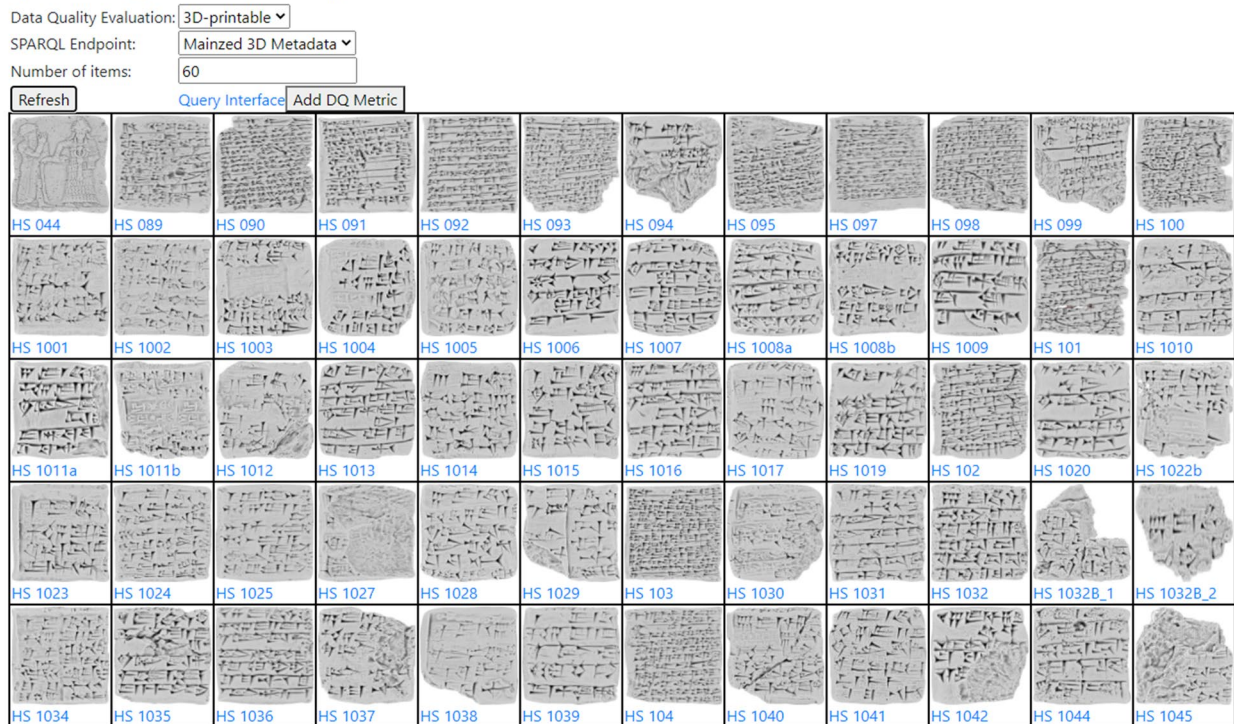


Fig. 6 HeiCuBeDa Demonstrator: Shows 3D printable 3D models out of the HeiCuBeDa dataset according to the metadata generated using the workflow presented in this publication

smoothed. By developing a structured metadata scheme that could be applied to different sensor types, metadata can be stored uniformly and compared more easily with each other.

The resulting metadata of this application case is accessible on Zenodo²³.

Application case 2: Quality of cuneiform tablet 3D scans

Our second application case is prototypic for the re-use of 3D models. It features 2.000 cuneiform tablets of the *Hilprecht Archive Online* (HAO)²⁴ in Jena and Berlin, which was transformed into the *Heidelberg Cuneiform Benchmark Dataset* (HeiCuBeDa)²⁵ for machine learning tasks to assist in the domain of Assyriology. Even though information about the 3D acquisition process is not provided and, therefore, to be considered lost, the workflow presented in this publication can be followed starting in stage 3. We could produce metadata using the previously described GigaMesh software [49], which showed that the overall quality is suitable for given tasks

in Assyriology as well as the related machine learning experiments. However, half of the tablets have only half of the spatial resolution possible. The most likely explanation is the use of two different export options provided by the software of the 3D scanner allowing to save 3D models either in full resolution or a so-called preview. While the preview is still sufficient to read the cuneiform texts, it may provide fewer details about sealings or fingerprints left on clay tablets.

We present two examples of such quality assessments. The first assessment indicates if a 3D model is eligible to be 3D-printed. The second assessment indicates if the 3D model is eligible to be processed by a machine learning algorithm for 3D cuneiform character recognition. Machine learning on 3D models of cuneiform tablets with the goal of cuneiform sign classification highly depends on the resolution of the given 3D model, which may be assessed with the metadata schema presented in this publication.

We created metadata according to our schema for the HeiCuBeDa dataset²⁶ to illustrate an application case and

²³ <https://doi.org/10.5281/zenodo.4428389>.

²⁴ <https://hilprecht.mpiwg-berlin.mpg.de/>.

²⁵ <https://doi.org/10.11588%2Fdata%2FIE8CCN>.

²⁶ <https://doi.org/10.5281/zenodo.4611087>.

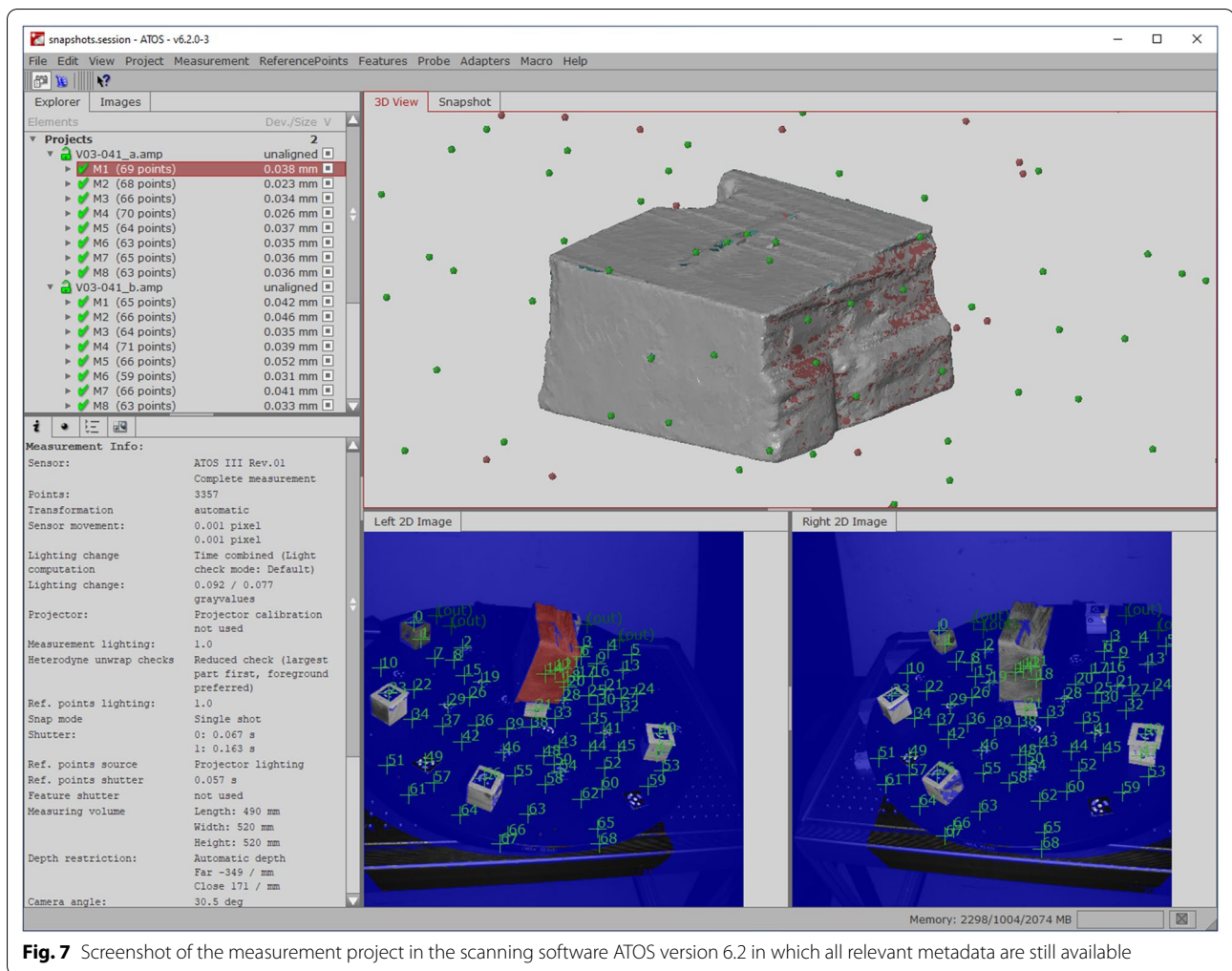


Fig. 7 Screenshot of the measurement project in the scanning software ATOS version 6.2 in which all relevant metadata are still available

created a demonstrator application to show the aforementioned two assessments. The demonstrator, shown in Fig. 6 queries a SPARQL endpoint into which the generated metadata has been entered. It retrieves exactly those 3D models which are printable. The creation of SPARQL queries for further data quality assessments is possible.

Application case 3: comparison of 3D models

The aim of the KUR project²⁷, which took place at the RGZM in cooperation with i3mainz, was to objectively and critically examine the influence of preservation methods on antique wet wood. One of the methods of investigation was the analysis of the change in shape caused by wet wood preservation. To assess this, 3D models were generated from 800 samples before and

after conservation using a structured light projector. In order to generate comparable 3D models, the samples were always captured and processed with the same sensor and the same parameters. The 3D models can be viewed on the project's website.²⁸ Metadata were not collected during the project period (2007–2011). Since the measurement projects are still available and can be opened, the metadata can now be created retrospectively with the scripts described in "Metadata export scripts". This allows the metadata for scanning and processing, described in "3D data capture" and "3D processing (scan software)", to be collected and exported in a structured way. Information on post-processing, such as filling holes or alignment, of the 3D model has unfortunately been lost (Fig. 7).

The reason to mention this research project here is the follow-up project CuTAWAY²⁹, which partly uses the

²⁷ https://www.hornemann-institut.de/german/epubl_projekte123.php.

²⁸ <https://www.rgzm.de/kur/>.

²⁹ <https://gepris.dfg.de/gepris/projekt/416877131>.

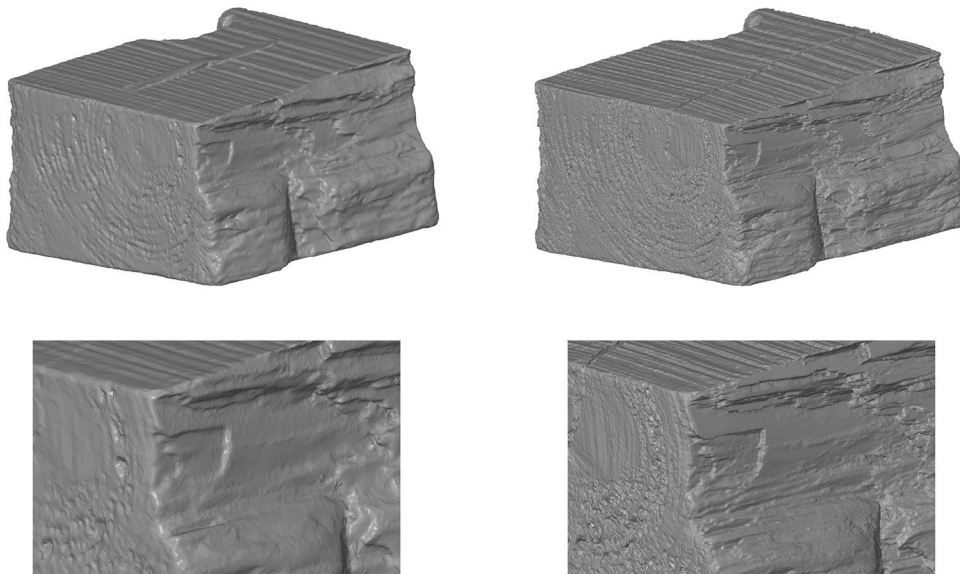


Fig. 8 Image of two 3D models of the same measurement object. The 3D model on the left was taken in 2012 with the available Structured Light Scanner and the one on the right eight years later, in 2020 with a newer Structured Light Scanner

wood samples and 3D models from the KUR project. New 3D models will be generated from the selected samples using Structured Light Scanning and compared with the existing datasets to quantify changes in the long-term storage of preserved wood. Since it is no longer possible to use the same scanner, new hardware and software will be used to create 3D models. For a comparison of the 3D models, the accuracy of the 3D scans should be considered.

Figure 8 shows two 3D models from the same sample. Eight years lie between the acquisitions and different Structured Light Scanners were used. In the illustration, no differences can be seen in the overall shape. Only the separately attached name badges on the top have changed and a zoom in reveals that the level of detail of the two models is indeed very different.

For a comparison, the different accuracy of the 3D scans and the settings used in the calculation and post-processing of the 3D models should be taken into account. The accuracy of a 3D model is influenced by many factors, including the scanner's resolution, the calibration accuracy, the alignment deviations of the individual measurements, the polygonization (computing the mesh) settings, and the postprocessing tools used.

In the completed KUR project, to compare the changes in two 3D models of a wood sample, distances, sections, areas and volumes were compared. The scanning accuracy and the post-processing of the 3D models were

taken into account and thus the results were rounded to one decimal place in millimetres. A more precise specification would have exceeded the accuracy of the 3D models.

We have published the two 3D models and the generated metadata of the 3D models on Zenodo to give the reader an impression of how the ontology model is used in practice.³⁰

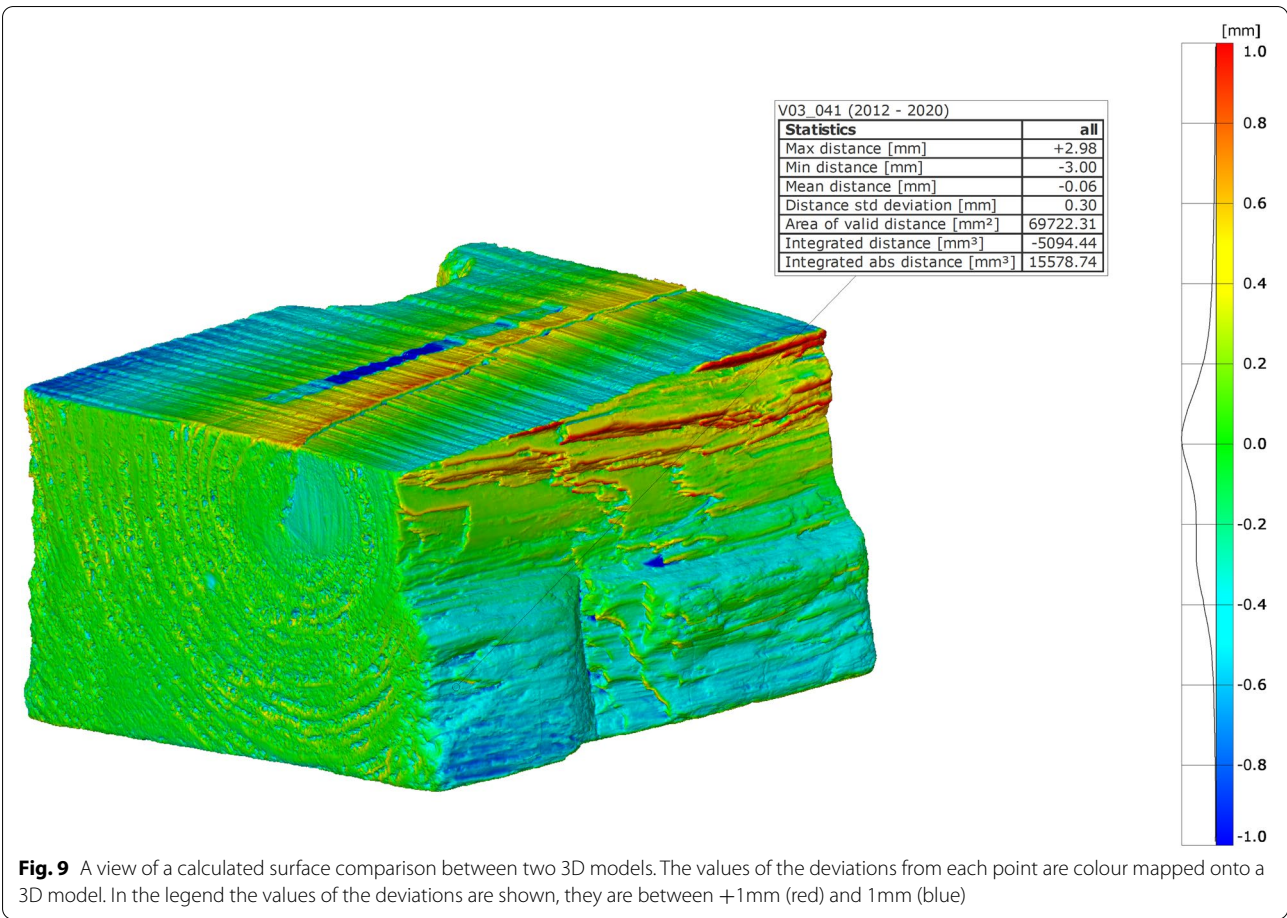
Limitations

The approach of capturing and saving metadata faces certain limitations, which we discuss as follows. At first, there is metadata that is not captured by the scanning software concerning the setup of the object and the setup of the environment used for scanning and its influence on the scanning result.

On the one hand, object-dependent properties such as the color, material, and structure of the object surface are not captured. Depending on what kind of object it is, certain settings are made in the scanning software. For example by photogrammetric techniques, a dark object must be exposed longer than a bright object. On the other hand, the surrounding situations can be different. Depending on how the ambient light is during the measurement, the settings must also be adjusted.

This means that even though the scanning process's metadata can be used for certain comparisons, the lack of the aforementioned information does not allow for a

³⁰ <https://www.doi.org/10.5281/zenodo.4428498>.



complete recreation of the circumstances related to the respective 3D scan. While such a recreation might be feasible if other means of capturing metadata are fulfilled (e.g., measuring the ambient light when the 3D scan is captured), it is currently not common practice to do so. Hence, we need to limit ourselves to metadata that can be captured, is reasonable to be captured and may give an added value compared to a previous situation without capturing metadata. Another limiting factor is the scanning software. Depending on the software’s capabilities, more or less metadata will be exposed and can be represented using the vocabulary defined in the ontology. However, the ontology we created mirrors the most common properties available in the set of software’s that we have considered. We expect this to be a minimal amount of metadata attributes to be available in comparable software implementations. Metadata dependent on human input, such as capturing people involved in the scanning process, has to be incorporated in a workflow, should be controlled by an auditing process, and is not always likely to be accomplished by the scanning software. This means that this kind of information is depending on the

workflow not always available. To change this circumstance, software support for these kinds of metadata is needed.

Conclusions

Our publication proposed and investigated the benefit of defining a metadata schema for capturing metadata about 3D scanning processes. We defined a typical workflow that is adaptable to various scanning processes and related this workflow to a corresponding metadata representation in linked data. The application cases we investigated showed that in this way, scanning processes become to a certain degree transparent and comparable across different capturing projects, and data quality metrics may harvest this new potential to rank or classify the fitness for the use of said 3D scans. In the future, it will be more and more likely to store 3D models and dependent digital artifacts such as 3D renderings in data repositories which have to expose among others also metadata about the capturing process which we propose here. Moreover, the combination of 3D models from different research projects and institutes and the sharing or accessing of these models in combined repositories is becoming a

possibility. It will then be crucial to evaluate the quality of the given scans to which we contribute by unifying the metadata scheme for describing the data origin. The informative value of analyses and interpretations is increased by existing metadata, which also includes the 3D acquisition.

Future work

In our future work, we would like to integrate additional 3D scanning systems and their outputs for compatibility with our ontology model, leading to an extended ontology model. The details of the presented ontology concern 3D scanners typically used to acquire cultural heritage objects with- and without a georeference.

However, there are further techniques to acquire 3D models, e.g., using SfM techniques or tomography systems. While the overall workflow is similar, there will be different means of metadata due to the nature of underlying measurement principles. For the utmost detailed coverage of the 3D model generation by metadata, several scanner types per principle have to be investigated.

Future implementations can be the (i) combination of datasets from different devices as shown in reference [50], (ii) inclusion of reference point coordinate systems and (iii) texture mapping using additional high-quality photographs. The combination of measurement methods can be seen in Fig. 1 both before and after the 3D scan phase. Another extension of our work could be the documentation of metadata for analysis methods. For example, the analysis of deformations between 3D models. Here the datasets must be aligned with each other, e.g. with the BestFit method or via reference points. For the interpretation of the deformation, not only the accuracy of the scans but also the transformation method and its results play an essential role [4]. In the example in Fig. 9, this was calculated with BestFit, the transformation deviation is 0.2 mm.

In addition, different metadata formats or even graph format exports are possible. Data might be analyzed using tools like Neo4J [51]. This would make further analysis and acceptance of formats possible for different research communities. Finally, metadata which can currently only be collected by manual inputs should be made either mandatory or more easily addable to the current database. Software needs to be enabled to automatically collect creator data, which we began to integrate in GigaMesh using for example an OAuth [52] login. In the same fashion standards and implementations to capture data about the scanning environment and other information about e.g. funding of the scanning endeavour among others.

Acknowledgements

This research was supported by experimental data created in the "Römische Grabdenkmäler aus Augusta Treverorum im überregionalen Vergleich: mediale Strategien sozialer Repräsentation" and the KUR project. We also thank the mainzed research network for enabling the cooperation which lead to this research effort. For fruitful discussions, we thank Kai-Christian Bruhn.

Authors' contributions

TH created the ontology model and worked on exporting metadata into linked data formats. AC created the scripts for the metadata export, captured and edited application case 3 and described the metadata for the ontology. LR created the scripts for the metadata export and captured and edited application case 1 and described the metadata for the ontology. HM is the author of the free and Open Source GigaMesh Software Framework, which was extended by linked-data exports and used for the practical examples. He created the first draft of the workflow schematic. All authors read and approved the final manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL. No funding was received for creating this publication.

Availability of data and materials

Ontology: <https://github.com/mainzed/mainzedObjectsOntology>. Export scripts: <https://doi.org/10.5281/zenodo.4566044>. GigaMesh Release with linked data support: <https://gitlab.com/fcgl/GigaMesh/-/tree/release2021.2>. Data for Application Case 1: <https://doi.org/10.5281/zenodo.4428389>. Data for Application Case 2: <https://doi.org/10.5281/zenodo.4611087>. Data for Application Case 3: <https://doi.org/10.5281/zenodo.4428498>.

Declarations

Competing interests

The authors declare that they have no competing interests.

Author details

¹i3mainz, Hochschule Mainz, Mainz, Germany. ²Römisch-Germanisches Zentralmuseum, Mainz, Germany.

Received: 4 May 2021 Accepted: 9 July 2021

Published online: 27 July 2021

References

- Bogacz, B., Mara, H.: Period classification of 3d cuneiform tablets with geometric neural networks. In: Proceedings of the 17th International Conference on Frontiers of Handwriting Recognition (ICFHR), Dortmund, Germany, pp. 31–43 (2020). <https://doi.org/10.1109/ICFHR2020.2020.00053>
- Yuan J, Abdul-Rashid H, Li B, Lu Y, Schreck T, Bai S, Bai X, Bui NM, Do MN, Do TL, Duong AD, He K, He X, Holenderski M, Jarnikov D, Le TK, Li W, Liu A, Liu X, Menkovski V, Nguyen KT, Nguyen TA, Nguyen VT, Nie W, Ninh VT, Rey P, Su Y, TonThat V, Tran MT, Wang T, Xiang S, Zhe S, Zhou H, Zhou Y, Zhou Z. A comparison of methods for 3d scene shape retrieval. *Comput Vision Image Understanding*. 2020;201: <https://doi.org/10.1016/j.cviu.2020.103070>.
- Tachibanaya, T.: Description of Exif file format (2001). <https://www.media.mit.edu/pia/Research/deepview/exif.html>
- Elkhrachy I. Modeling and visualization of three dimensional objects using low-cost terrestrial photogrammetry. *Int J Architectural Heritage*. 2020;14(10):1456–67. <https://doi.org/10.1080/15583058.2019.1613454>.
- Luhmann T. Nahbereichsphotogrammetrie: Grundlagen-Methoden-Beispiele. Offenbach, Germany: Wichmann; 2018.
- Prud'hommeaux, E., Seaborne, A.: SPARQL query language for RDF. W3C recommendation, W3C (January 2008). <https://www.w3.org/TR/2008/REC-rdf-sparql-query-20080115/>
- Lengauer, S., Komar, A., Karl, S., Trinkl, E., Preiner, R., Schreck, T.: Visual Exploration of Cultural Heritage Collections with Linked Spatiotemporal, Shape and Metadata Views. In: Krüer, J., Niessner, M., Stückler, J.

- (eds.) Vision, Modeling, and Visualization. The Eurographics Association, Geneva, Switzerland (2020). <https://doi.org/10.2312/vmw.20201196>
8. Turk, G.: The ply polygon file format. Recuperado de (1994)
 9. Cidoc, C.: The CIDOC Conceptual Reference Model (2003). <http://www.cidoc-crm.org>
 10. Sahoo, S., McGuinness, D., Lebo, T.: PROV-o: The PROV ontology. W3C recommendation, W3C (April 2013). <https://www.w3.org/TR/2013/REC-prov-o-20130430/>
 11. Doerr, M., Theodoridou, M.: Crmdig: A generic digital provenance model for scientific observation. In: 3rd USENIX Workshop on the Theory and Practice of Provenance (TaPP 11). USENIX Association, Heraklion, Crete Greece (2011). <https://www.usenix.org/conference/tapp11/crmdig-generic-digital-provenance-model-scientific-observation>
 12. Amico, N., Ronzino, P., Felicetti, A., Niccolucci, F.: Quality management of 3d cultural heritage replicas with cidoc-crm. In: CRMEX@TPDL, pp. 61–69 (2013). Citeseer
 13. Tzompanaki, K., Doerr, M., Theodoridou, M., Fundulaki, I.: Reasoning based on property propagation on cidoc-crm and crmdig based repositories. In: CRMEX@TPDL (2013)
 14. Vasilakis, G., Garcia-Rojas, A., Papaleo, L., Catalano, C.E., Spagnuolo, M., Robbiano, F., Vavalis, M., Pitikakis, M.: A common ontology for multi-dimensional shapes. In: SAMT 2007 2nd International Conference on Semantic and Digital Media, MARESO Workshop Proceedings, pp. 31–43 (2007)
 15. Mi X, Pollock BM. Metadata schema to facilitate linked data for 3d digital models of cultural heritage collections: A university of south florida libraries case study. *Cataloging & Classification Quarterly*. 2018;56(2–3):273–86. <https://doi.org/10.1080/01639374.2017.1388894>.
 16. D'Andrea, A., Fernie, K.: Carare 2.0: A metadata schema for 3d cultural objects. In: 2013 Digital Heritage International Congress (DigitalHeritage), vol. 2, pp. 137–143 (2013). <https://doi.org/10.1109/DigitalHeritage.2013.6744745>
 17. Karmacharya A, Wefers S, Boochs F. Knowledge based optimal recommendation of spatial technologies for documentation of buildings. In: Ioannides M, Martins J, Žarnić R, Lim V, editors. *Advances in Digital Cultural Heritage*. Cham: Springer; 2018. p. 65–77.
 18. Pintus R, Pal K, Yang Y, Weyrich T, Gobetti E, Rushmeier H. A survey of geometric analysis in cultural heritage. *Comput Graphics Forum*. 2016;35(1):4–31. <https://doi.org/10.1111/cgf.12668>.
 19. Scalas A, Mortara M, Spagnuolo M. A pipeline for the preparation of artefacts that provides annotations persistence. *J Cult Heritage*. 2020;41:113–24.
 20. Gualandi, M.L., Scopigno, R., Wolf, L., Richards, J., Garrigos, J.B.i., Heinzelmann, M., Hervas, M.A., Vila, L., Zallocco, M.: ArchAIDE - Archaeological Automatic Interpretation and Documentation of cEramics. In: Catalano, C.E., Luca, L.D. (eds.) *Eurographics Workshop on Graphics and Cultural Heritage*. The Eurographics Association, Graz, Austria (2016). <https://doi.org/10.2312/gch.20161408>
 21. Fecher F, Reindel M, Fux P, Gubler B, Mara H, Bayer P, Lyons M. The ceramic finds from guadalupe, honduras: Optimizing archaeological documentation with a combination of digital and analog techniques. *J Global Archaeol*. 2020;1:1–54.
 22. Initiative, D.C.M., et al.: Dublin core metadata element set, version 1.1. Technical report, Dublin Core Metadata Initiative (2012). <https://www.dublincore.org/specifications/dublin-core/dcmi-terms/>
 23. Lassila, O.: Resource description framework (RDF) model and syntax specification. W3C recommendation, W3C (February 1999). <https://www.w3.org/TR/1999/REC-rdf-syntax-19990222/>
 24. Rijgersberg H, van Assem M, Top J.L. Ontology of units of measure and related concepts. *Semantic Web*. 2013;4(1):3–13. <https://doi.org/10.3233/SW-2012-0069>.
 25. Secretary IC. Data quality - Part 8: Information and data quality: Concepts and measuring. International Organization for Standardization, Geneva, CH (November: Standard; 2015).
 26. Debatista, J., Lange, C., Auer, S.: daq, an ontology for dataset quality information. In: LDOW (2014)
 27. Homburg T, Boochs F. Situation-dependent data quality analysis for geo-spatial data using semantic technologies. In: Abramowicz W, Paschke A, editors. *Business Information Systems Workshops*. Cham: Springer; 2019. p. 566–78.
 28. Bathow, C., Breuckmann, B., Scopigno, R.: Verification and Acceptance Tests for High Definition 3D Surface Scanners. In: Artusi, A., Joly, M., Lucet, G., Pitzalis, D., Ribes, A. (eds.) *VAST: International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage*. The Eurographics Association, Paris, France (2010). <https://doi.org/10.2312/VAST/VAST10/009-016>
 29. Calandra I, Schunk L, Bob K, Gneisinger W, Pedergrana A, Paixao E, Hildebrandt A, Marreiros J. The effect of numerical aperture on quantitative use-wear studies and its implication on reproducibility. *Scientific Rep*. 2019;9(1):1–10. <https://doi.org/10.1038/s41598-019-42713-w>.
 30. ISO Central Secretary. Geometrical product specifications (GPS) - Surface texture: Areal - Part 2: Terms, definitions and surface texture parameters. International Organization for Standardization, Geneva, CH (March: Standard; 2012).
 31. Spahiu, B., Maurino, A., Palmonari, M.: Towards improving the quality of knowledge graphs with data-driven ontology patterns and shacl. In: ISWC (Best Workshop Papers), pp. 103–117. CEUR-WS, Aachen, Germany (2018)
 32. Daneshmand, M., Helmi, A., Avots, E., Noroozi, F., Alisananoglu, F., Arslan, H.S., Gorbova, J., Haamer, R.E., Ozcinar, C., Anbarjafari, G.: 3D Scanning: A Comprehensive Survey (2018). [arXiv:1801.08863](https://arxiv.org/abs/1801.08863)
 33. Michael Breuer, Rainer Maria Czichon, Marko Koch, Monika Lehmann, Dirk Paul Mielke: Photogrammetrische 3d-dokumentation von nassholz-funden aus oymaağaç höyük / nerik (provinz samsun / TR) *10*, 47–62 (2017). <https://doi.org/10.11588/ra.2017.0.75392>. Accessed 2021-01-20
 34. Anja Cramer, Heinz, G., Justus, C., Reich, T.: 3d-scanning und virtuelle 3d-modelle zur unterstützung der untersuchung und virtuellen rekonstruktionen des grabmals von frankfurt am main - zeilsheim. In: P. Fasold, A. Hampel, M. Scholz, M. Tabaczek, Der Römische Bestattungsort Von Frankfurt Am Main - Zeilsheim. Grabbau und Gräber der Provinzialen Oberschicht. *Schriften des Archäologischen Museums Frankfurt* 26, pp. 125–130. Schnell und Steiner, Frankfurt, Germany (2016). Backup Publisher: Schnell und Steiner ISSN: 978-3-7954-2974-4
 35. Klump J, Bertelmann R, Brase J, Diepenbroek M, Grobe H, Höck H, Lautenschlager M, Schindler U, Sens I, Wächter J. Data publication in the open access initiative. *Data Sci J*. 2006;5:79–83.
 36. Druskat S, Spaaks JH, Chue Hong N, Haines R, Baker J. Citation File Format (CFF) - Specifications. Zenodo. 2021. <https://doi.org/10.5281/zenodo.4813122>.
 37. Paskin N. Digital object identifier (doi®) system. *Encyclopedia of library and information sciences*. 2010;3:1586–92.
 38. Brutzman D, Daly L. X3D: Extensible 3D Graphics for Web Authors. Amsterdam, Netherlands: Elsevier; 2010.
 39. Beckett, D., Berners-Lee, T., Prud'hommeaux, E., Carothers, G.: Rdf 1.1 turtle. World Wide Web Consortium, 18–31 (2014)
 40. Champin, P.-A., Kellogg, G., Longley, D.: JSON-ld 1.1. W3C recommendation, W3C (July 2020). <https://www.w3.org/TR/2020/REC-json-ld11-20200716/>
 41. Arenas, M., Bertails, A., Prud'hommeaux, E., Sequeda, J.: A direct mapping of relational data to RDF. W3C recommendation, W3C (September 2012). <https://www.w3.org/TR/2012/REC-rdb-direct-mapping-20120927/>
 42. Kyzirakos, K., Vlachopoulos, I., Savva, D., Manegold, S., Koubarakis, M.: Geotriples: A tool for publishing geospatial data as rdf graphs using r2rml mappings. In: Proceedings of the 2014 International Conference on Posters & Demonstrations Track - Volume 1272. ISWC-PD'14, pp. 393–396. CEUR-WS.org, Aachen, DEU (2014)
 43. Das, S., Sundara, S., Cyganiak, R.: R2RML: RDB to RDF mapping language. W3C recommendation, W3C (September 2012). <https://www.w3.org/TR/2012/REC-r2rml-20120927/>
 44. Deursen, D.V., Poppe, C., Martens, G., Mannens, E., Walle, R.V.d.: Xml to rdf conversion: A generic approach. In: 2008 International Conference on Automated Solutions for Cross Media Content and Multi-Channel Distribution, pp. 138–144 (2008). <https://doi.org/10.1109/AXMEDIS.2008.17>
 45. Prudhomme C, Homburg T, Ponciano J-J, Boochs F, Cruz C, Roxin A-M. Interpretation and automatic integration of geospatial data into the semantic web. *Computing*. 2020;102(2):365–91.
 46. Antoniou, G., Van Harmelen, F.: Web ontology language: Owl. In: *Handbook on Ontologies*, pp. 67–92. Springer, Heidelberg, Germany (2004)

47. Bray, T.: The JavaScript Object Notation (JSON) Data Interchange Format. RFC Editor (2014). <https://doi.org/10.17487/RFC7159>. <https://rfc-editor.org/rfc/rfc7159.txt>
48. Mara, H., Krömker, S., Jakob, S., Breuckmann, B.: Gigamesh and gilgamesh: –3d multiscale integral invariant cuneiform character extraction. In: Proceedings of the 11th International Conference on Virtual Reality, Archaeology and Cultural Heritage, pp. 131–138 (2010)
49. Mara, H., Bogacz, B.: Breaking the code on broken tablets: The learning challenge for annotated cuneiform script in normalized 2d and 3d datasets. In: 2019 International Conference on Document Analysis and Recognition (ICDAR), Sidney, Australia, pp. 148–153 (2019). IEEE
50. Karl, S., Bayer, P., Mara, H., Márton, A.: Advanced documentation methods in studying corinthian black-figure vase painting. In: Digitize: Research–Record–Reactivate–Reconstruct. Proceedings of the 23rd International Conference on Cultural Heritage and New Technologies, Vienna, Austria, pp. 1–13 (2018)
51. Miller, J.J.: Graph database applications and concepts with neo4j. In: Proceedings of the Southern Association for Information Systems Conference, Atlanta, GA, USA, vol. 2324 (2013)
52. Hardt, D.: The OAuth 2.0 Authorization Framework. RFC Editor (2012). <https://doi.org/10.17487/RFC6749>. <https://rfc-editor.org/rfc/rfc6749.txt>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)