1 Preface

1.1 Aim of the specification

This E-ARK specification is part of a family of specifications that provide a common set of requirements for packaging digital information. These specifications are based on common, international standards for transmitting, describing and preserving digital data. They have been produced to help data creators, software developers and digital archives tackle the challenge of short-, medium- and long-term data management and reuse in a sustainable, authentic, cost-efficient, manageable and interoperable way.

The foundation for these specifications is the Reference Model for an Open Archival Information System (OAIS) which has Information Packages at its core. Familiarity with the core functional entities of OAIS is a prerequisite for understanding the specifications. A visualisation of the current specification network can be seen here:

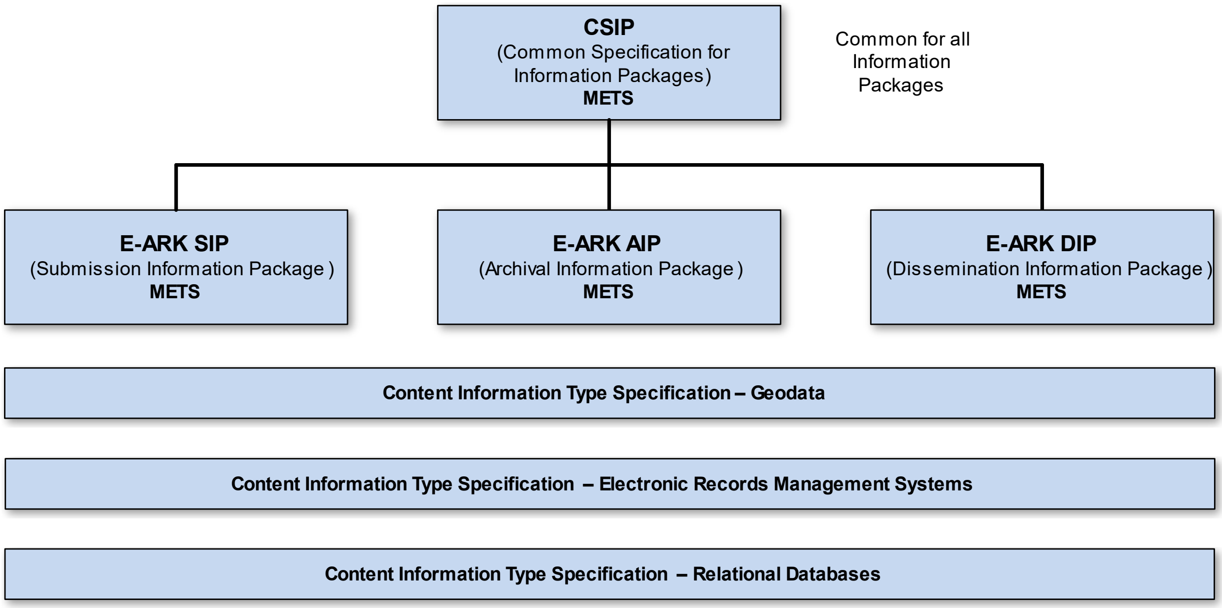

Figure 1: E-ARK specification dependency hierarchy

Table 1: Specification hierarchy aims and goals

Specification	Aim and Goals
Common Specification for Information Packages	This document introduces the concept of a Common Specification for Information Packages (CSIP). Its three main purposes are to: • Establish a common understanding of the requirements which need to be met in order to achieve interoperability of Information Packages.

Specification	Aim and Goals
	• Establish a common base for the development of more specific Information Package definitions and tools within the digital preservation community. • Propose the details of an XML-based implementation of the requirements using, to the largest possible extent, standards which are widely used in international digital preservation. Ultimately the goal of the Common Specification is to reach a level of interoperability between all Information Packages so that tools implementing the Common Specification can be adopted by institutions without the need for further modifications or adaptations.
E-ARK SIP	The main aims of this specification are to: • Define a general structure for a Submission Information Package format suitable for a wide variety of archival scenarios, e.g. document and image collections, databases or geographical data. • Enhance interoperability between Producers and Archives. • Recommend best practices regarding metadata, content and structure of Submission Information Packages.
E-ARK AIP	The main aims of this specification are to: • Define a generic structure of the AIP format suitable for a wide variety of data types, such as document and image collections, archival records, databases or geographical data. • Recommend a set of metadata related to the structural and the preservation aspects of the AIP as implemented by the reference implementation eArchiving ToolBox (formerly earkweb). • Ensure the format is suitable to store large quantities of data.
E-ARK DIP	The main aims of this specification are to: • Define a generic structure of the DIP format suitable for a wide variety of archival records, such as document and image collections, databases or geographical data. • Recommend a set of metadata related to the structural and access aspects of the DIP.
Content Information Type Specifications	The main aim and goal of a Content Information Type Specification is to: • Define, in technical terms, how data and metadata must be formatted and placed within a CSIP Information Package in

Specification	Aim and Goals
	order to achieve interoperability in exchanging specific Content Information. The number of possible Content Information Type Specifications is unlimited.

1.2 Organisational support

This specification is maintained by the Digital Information LifeCycle Interoperability Standards Board (DILCIS Board). The DILCIS Board (http://dilcis.eu/) was created to enhance and maintain the draft specifications developed in the European Archival Records and Knowledge Preservation Project (E-ARK project) which concluded in January 2017 (http://eark-project.com/). The Board consists of eight members, but there is no restriction on the number of participants in the work. All Board documents and specifications are stored in GitHub (https://github.com/DILCISBoard) while published versions are made available on the Board webpage. Since 2018 the DILCIS Board has been responsible for the core specifications in the Connecting Europe Facility eArchiving Building Block (https://ec.europa.eu/cefdigital/wiki/display/CEFDIGITAL/eArchiving).

1.3 Authors

A full list of contributors to this specification, as well as the revision history can be found in Appendix 1.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

2 Introduction

2.1 Purpose

The purpose of this document is to accompany the Content Information Type Specification for 3D Heritage Model data (CITS 3D HM) and to provider context and rationale for its principles and requirements. The Specification is designed to be used for the transfer to archives as well as for records exchange between different organisations and repositories. The Specification is supported by METS profiles for the Root and Representation METS files, an example package and this accompanying Guideline document.

2.2 Scope

Use of 3D data is widespread across many domains, with a plethora of applications and data formats. This 3D HM content specification limits its scope to the area of 3D models of cultural heritage artefacts, buildings and sites that have been produced as:

- Point Cloud models (for example by: photogrammetry, laser scanning, structured light);
- CAD based models (including building information models, BIM amd heritage building information models, HBIM);
- Volumetric models (computer tomography);
- GIS models; and
- Procedural models (reconstructions).

The scope of the CITS may evolve in future to accommodate other production technologies. The 3D HM CITS has a related specification for 3D Product Model data, CITS 3D PM.

2.3 Types of 3D Models

3D models used for heritage may be created to document existing arefacts and their status (heritage documentation) or to represent a reconstruction of their assumed status in the past (virtual reconstruction). 3D models may be created using various techniques, producing a variety of model types and file formats. For the models produced with any of the techniques described below, the model quality is measured by its resolution (the minimum distance of two points on the target to be recognized as distinct), accuracy (how close the measurement is to its true value) and precision (how close different measurements of the same target are to each other).

2.3.1 3D Point Cloud data

3D Point Cloud models are realised as masses of digital references in the 3D space (equivalent to a 2D pixel image), which can be further processed to produce a more solid representation by meshing, i.e. generating a surface reconstruction based on a collection of vertices, edges, and faces (usually triangles), built by interpolation. The mesh defines the shape of a polyhedral object approximating the surface of the real one. Several interpolation methods are available to do this.

Point Cloud models can represent the present or past shape of heritage artefacts with variable fidelity depending on the precision of the capture or creation process. There are several techniques for their acquisition:

2.3.1.1 Photogrammetry

Photogrammetry involves the extraction of three-dimensional measurements from two-dimensional data (such as images). Developments in GPU-based processing have allowed rapid reconstruction of 3D surface meshes from sets of conventional photographs of physical objects or environments. The mesh output from this technique can be enhanced by colour information at vertices (vertex colour) or by using associated 2D images representing surface colour, mapped to the mesh (texture map).

2.3.1.2 Laser Scanning

Laser Scanning is the process of recording three-dimensional information about real-world objects or environments by rapidly sampling or scanning an object's surface with lasers. The information is returned as a collection of x, y, z coordinates. Laser scanning devices may use time-of-flight, phase or triangulation methods and the Point Cloud or the mesh output from this technique may be enhanced by colour information at points/vertices (vertex colour) or with an associated 2D image of the surface, which is mapped onto the mesh (texture map). Scanning results can be influenced by the reflective properties of the surface of the object to be scanned.

2.3.1.3 Structured Light

Structured Light is a method of 3D capture that relies on the distortion of projected light to calculate surface form. A known pattern (often a grid or horizontal lines) of light when projected onto a surface appears distorted from perspectives other than that of the projector and observation of this distortion can be used for geometric reconstruction of the surface shape. The mesh output from this technique may be enhanced by colour information at points/vertices (i.e., vertex colour) or an associated 2D image representing surface colour, which is mapped to the mesh (texture map).

The three methods produce substantially similar 3D models. The choice of one of them depends on the object to be scanned, on its position (e.g. for the roof of a building, laser scanning can be difficult) and on the availability of the capture device. Using any of the above techniques there may exist parts of the object to be scanned which are not visible, e.g. undercuts. These have to be completed by hand drawing.

2.3.2 CAD-based modelling

In Computer Aided Design (CAD) modelling, 3D models are built using 3D CAD systems by defining vertices, edges and faces. Curves may further be defined with various methods, e.g. as splines. The model may also incorporate information about the corresponding parts of the object. In the specialied case of construction design, the technique takes the name of BIM (Building Information Modelling) where additional information is organized according to the IFC (Industry Foundation Classes) standard. If this standard is extended to include information that is relevant for cultural heritage the approach takes the name of HBIM (Heritage BIM).

2.3.3 Volumetric modelling

Volumetric 3D models are obtained as the result of a volumetric data acquisition based on layers, i.e. slices of the object automatically assembled to build the 3D model. A typical technique is Computed Tomography (CT), borrowed from medicine, sometimes used to explore the interior of objects. CT directly produces a 3D model of the object including its interior.

2.3.4 GIS 3D modelling

The Geophysical Information System (GIS) modelling technique is used to generate the environment in which a heritage asset is placed. It uses the same techniques used in GIS to generate land surfaces. It is used to study the relationship of a monument with its geographical position (e.g. a castle in relation to its position) or the mutual relationship of different assets (e.g. a system of sighting towers).

2.3.5 Procedural Modelling

Procedural Modelling produces models generated by the application of a mathematical model that generates the model components. It is very rarely used for heritage documentation but may be used for reconstructions, especially to produce films. It enables the fast production of large models, e.g. a whole town with its inhabitants.

2.3.6 180° Panoramas

180° Panoramas consist of a 180° circular photographs and are typically used for interiors. Although they are not proper 3D models, their use is similar. Relevant parameters are the same as those of photographs.

2.4 Sources of 3D models

3D models may be created by direct automatic acquisition, CAD, based on measures of the object to be modelled, bibliography and source-based modelling. The latter is a method of model production based on documents, reference photographs, or other sources of information about the present shape of a real-world object or place or about a past one. They are mainly created as CAD-based models, or more rarely using other digital design techniques (e.g. 3D sculpture).

2.5 Data collected or produced during the creation of a 3D model

The 3D model data sets extend beyond that which simply defines the model geometry (the model). Such data can correlate to the model but also to the conditions of the whole production process and to its intended use. This additional data can determine the suitability and quality of the model for its intended use. Such additional data can include:

2.5.1 Metadata

In addition to the metadata requirements of an information package as described in the common specification (CSIP) and categories of metadata typically found in archival systems (descriptive, structural, administrative, provenance, technical and preservation) there may be additional metadata requirements for models concerning the tools used for the digitization, information about the hardware and software used for the acquisition or creation of the model and metadata that forms part of the data model itself (model metadata). Information can also be recorded concerning the resolution, accuracy, and precision of the digitization process.

Metadata is by definition structured data which can be serialised by a variety of methods (xml, json, rdf, turtle etc) and defined by a standardised or localised schema.

2.5.2 Paradata

Paradata concerns the conditions under which the digitization or model creation took place (who did it, where, what were the environmental conditions, what was the equipment used), the intended purpose of the model (documentation, research, communication), rationales for decisions taken in the digitization process leading to the final result, and for reconstructions, references to the documentation supporting the model creation process (images, drawings, reports, archaeological/historical interpretation).

Standards do not exist for Paradata although much academic work has been done on describing best practice for its acquisition and recording. Paradata is mainly realised as unstructured text.

2.5.3 Geometry information

Geometry information is direcftly associated with the model and is generated as a result of the creation of the model or from further processing, for example the original point cloud, the meshed object and so on. Geometry information is typically structured data values and specific to a model format.

2.5.4 Processing information

Processing information includes information about the processes used to produce any intermediate and the final data model. For example the software used, the interpolation type used for meshing, the standards used (if any) for the production of geometry data and so on. Prcessing information is typically detailed textual (unstructured) data, but could be encoded to a defined schema.

2.5.5 Quality and Authentication information

Quality information can be expressed as estimated resolution, accuracy and precision, and suitability for the models intended purpose. Authentication information can include model validation rules and results (is the model a good representation of the original) and verification quality rules and results (is the model within defined quality rules). Authentication information can also include Digital Signatures that guarantee the quality/authenticity of the model at the time it is archived. Quality and authentication information can be unstructured or structured data and can include recording of events within preservation metadata. More information on Authentication can be found in the CITS 3D PM Guideline at: https://github.com/DILCISBoard/CITS-3DPM/tree/rel/v1.0.0.

2.6 File Formats

3D file formats contain data representing the model type and the necessary information for its display. Use of file formats that offer the best long-term prospects in terms of usability, accessibility, and sustainability is recommended and is generally achieved with formats that have wide usage, have open specifications, and are independent of specific software or developers. In practice it is not always possible to meet all three criteria above, and a balance has to be found. Open file formats are available, and there may be options to convert proprietary file formats to them; but there is risk which should be assessed for information loss in the transformation. Keeping original, proprietary formats is not ideal, but with the risk of loss of information in transforming file formats, the inclusion of original and format derivations in archival packages is recommended.

No single format is ideal for preserving 3D data. The choice of preservation file format depends on the format type, the specific features and functions that need preserving (significant properties) and the intended future applications of the model. As with all preservation format transformation decisions, the choice can limit future wide reusability and the planned future users or designated community must be kept in mind.

Recommendations of preferred file formats have been compiled by for example, the Expert Group on Digital Cultural Heritage and Europeana[1], Data Archiving and Networked Services (DANS)[2] and the UK Data Service (UKDS)[3]. Preferred formats specific to cultural heritage types of data have been compiled by the Archaeological Data Service (ADS)[4], with extensive documentation40 on their use. A summary of preferred formats for 3D data can also be found in the excellent paperon long-term preservation of 3D models by Nicola Amico and Achille Felicetti[5].

Interoperability is key. The European Commission has compiled a comprehensive list of current 3D formats, covering both raster and vector types[6]. This compilation not only serves as a valuable resource but also highlights the concerted efforts of international standardisation bodies, such as the European Committee for Standardization, the International Organization for Standardization (ISO), and the Web 3D Consortium.

2.7 Layered Data Model

This section introduces the role of the CITS 3D HM and its dependencies on the basic structures of the Information Package.

> The specification is created based on the requirements of the Common Specification for Information Packages (CSIP), the specification for Submission Information Packages (E-ARK SIP) and the specification for Archival Information Packages (E-ARK AIP). To fully understand its requirements, we highly recommend that users review the requirements and the terminology of the source documents, before using the specification.

[1] https://digital-strategy.ec.europa.eu/en/library/basic-principles-and-tips-3d-digitisationcultural-heritage
[2] https://dans.knaw.nl/en/file-formats/

[3] https://ukdataservice.ac.uk/learning-hub/research-data-management/format-your-data/recommended-formats/
[4] https://archaeologydataservice.ac.uk/help-guidance/guides-to-good-practice/data-analysis-and-visualisation/3d-models/creating-3d-data/file-formats/
[5] https://archaeologydataservice.ac.uk/help-guidance/guides-to-good-practice/data-analysis-and-visualisation/3d-models/creating-3d-data/file-formats/
[6] https://digital-strategy.ec.europa.eu/en/library/study-quality-3d-digitisation-tangible-cultural-heritage

The data model structure is based on a layered approach for information package definitions (Figure 2). The Common Specification for Information Packages (CSIP) forms the outermost layer. The general SIP, AIP and DIP specifications add respectively, submission, archiving and dissemination information to the CSIP specification. The third layer of the model represents specific content information type specifications, such as this 3D HM specification. Additional layers for business-specific specifications and local variant implementations of any specification can be added to suit the needs of the organisation.

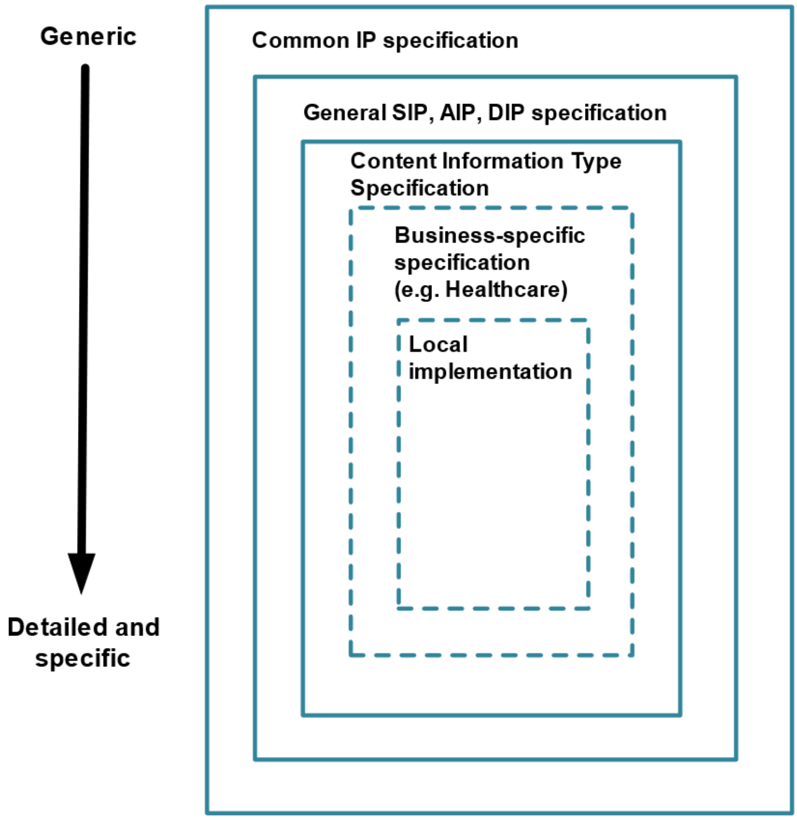

Figure 2: Data Model Structure

Every level in the data model structure inherits metadata entities and elements from the higher levels. In order to increase adoption, a flexible schema has been developed. This will allow for extension points where the schema in each layer can be extended to accommodate additional information on the next specific layer until, finally, the local implementation can add specific entities or metadata elements to satisfy specific local needs. Extension points can be implemented by:

- Embedding foreign extension schemas (in the same way as supported by METS [http://www.loc.gov/standards/mets/] and PREMIS [http://www.loc.gov/standards/premis/]). These both support increasing the granularity of existing metadata elements by using more detailed data structures as well as adding new types of metadata.
- Substituting metadata schemas for standards more appropriate for the local implementation.

The structure allows the addition of more detailed requirements for metadata entities, for example, by:

- Increasing the granularity of metadata elements by using more detailed data structures, or
- Adding local controlled vocabularies.

For consistency, design principles are reused between layers as much as possible.

3 Specification

3.1 Requirements Structure

The Content Information Type Specification for 3D Heritage Model data aims to define the necessary elements required to preserve the accessibility and authenticity of 3D Heritage Model data over time and across changing technical environments. The specification elevates the level (and adjusts the cardinality) of some of the requirements set out in the Common Specification (CSIP) and package specifications (namely SIP and AIP) and adds new requirements for the package structure, descriptive metadata, preservation metadata and accompanying METS files. The specification sets out general Principles that underpin the specific requirements and further context for the requirements and Principles can be found in this accompanying guideline.

3.2 Principles

3.2.1 Principle– data formats and representations

The E-ARK 3D Heritage Model content specification should be data agnostic, i.e. any type of data or data format for 3D models can be packaged for long-term archiving and none is excluded; however the use of open-standards and the inclusion in packages of multiple representations is encouraged to reduce the risk of data format obsolescence and of information loss. Where open 3D model data formats exist (e.g. STEP-IFC for BIM) their use alongside original, proprietary data formats is strongly encouraged.

3.2.2 Principle– use of PREMIS

From the CSIP and PREMIS specification:

- PREMIS should be used to record detailed technical metadata;
- Technical information should be included in PREMIS metadata by using extension schemas;
- Information about agents carrying out preservation actions must be recorded in the PREMIS metadata (this is because METS agents describe agents relevant for generic IP level events such as the creation or submission of the package, not the preservation of the data);
- Event descriptions should be included in PREMIS metadata. Use of the official PREMIS event vocabulary (https://id.loc.gov/vocabulary/preservation/eventType.html) is recommended;

- Detailed Rights information should be included in PREMIS. High-level Rights information in METS indicates restrictions. Detailed, object-specific Rights information should be included in the PREMIS metadata;
- File format information for all files should be included as Persistent Unique Identifier (PUID) values in the appropriate PREMIS semantic units.

Desirable technical and preservation metadata in the context of the 3D HM CITS can include:

- Reference to the content information standard (e.g. STEP-IFC);
- Information about the generating system (Creating Agent software);
- A description of the environment used to render the model.

Event descriptions can include:

- Creation events;
- Conversion or change events;
- Model authentication events.

Detailed technical metadata in the context of the 3D Heritage CITS can include:

- File format, characterisation, checksums;
- Relationships (is part, contains parts).

Rights information can include:

- Access rights;
- License restrictions;
- Copyright owner;

Use of PREMIS must conform to the requirements of the CITS Preservation Metadata.

(CSPM).

3.3 Use cases for archiving of 3D Heritage Model Data

3D Heritage Models are captured or created as representations of physical objects, buildings or sites as so called ‘Digital Twins’ when they are combined with comprehensive contextual information.[7] According to Niccollucci “A digital twin is the digital replica of a real-world object. It includes all the necessary information and is able to simulate - in a digital environment - the characteristics and the behaviour of its real counterpart”.

[7] Niccolucci F, Markhoff B, Theodoridou M *et al.* **The Heritage Digital Twin: a bicycle made for two. The integration of digital methodologies into cultural heritage research [version 1; peer review: awaiting peer review]** Open Research
Europe 2023, **3**:64 https://doi.org/10.12688/openreseurope.15496.1

The use of 3D in Heritage is being driven by needs of conservation, academic research, public access demonstrated by projects such as the European Commission 'TwinIt'[8] which states " Making cultural heritage available for future generations to enjoy and be inspired by is a major public policy goal in the EU. 3D technologies offer unprecedented opportunities to advance this objective, widening access to culture, supporting digital preservation and fostering the reuse of Europe's cultural assets."

Academic work such as by Niccolucci and in the Parthenos[9] report "Digital 3D Objects in Arts and Humanities: challenges of creation, interoperability and preservation"[10] aim to continue work to establish best practice and standards for the creation and long-term preservation of 3D models started by the seminal "London Charter" in 2006[11].

The use cases for creation of 3D Heritage Models then are wide ranging, but the use cases for the archiving of 3D models remain fairly straightforward while serving a wide community:

1. To enable long-term archiving of 3D Heritage Model data whilst preserving the usability, authenticity and accessibility of the data over time;
2. To enable inter-organisational exchange of 3D Heritage Model data whilst facilitating the understanding of the provenance, context and means to render;
3. To enable acceptance of 3D Heritage Model data submission packages (SIPs) at a central repository.

4 Implementation

4.1 Package Structure

The CITS 3D HM CITS inherits its package structure from the E-ARK Common Specification for Information Packages and is shown in Figure 3. It can be seen that additional folders have been added for Paradata, Other and optionally Authentication Documentation at root and Representation level but otherwise the structure is identical.

[8] https://pro.europeana.eu/page/twin-it-3d-for-europe-s-culture
[9] https://www.parthenos-project.eu/about-the-project-2
[10] https://www.academia.edu/120516742/Digital_3D_Objects_in_Art_and_Humanities_challenges_of_creation_interoperability_and_preservation_White_paper
[11] https://londoncharter.org/

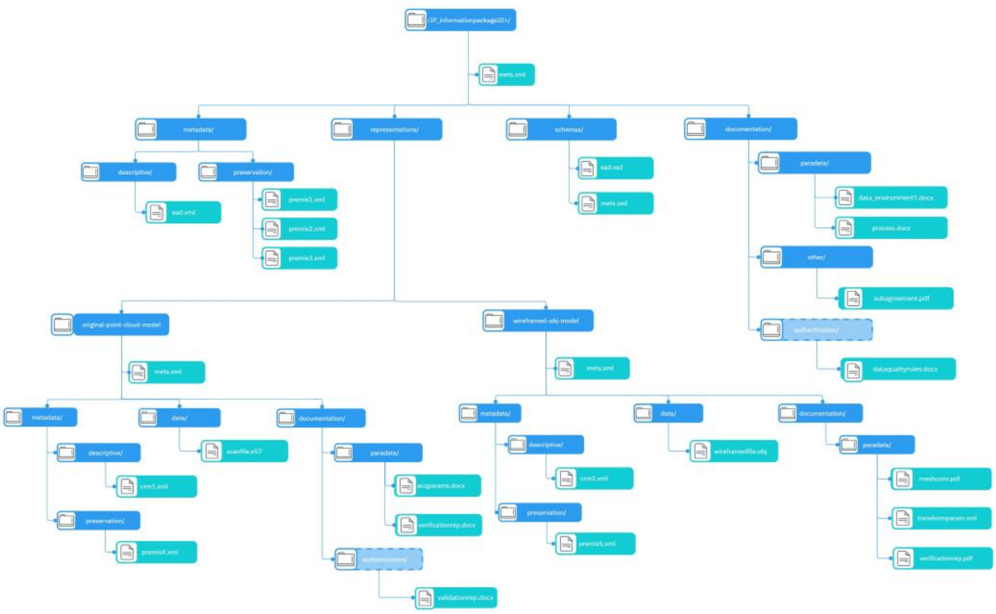

Figure 3: Example Information Package Folder Structure

Multiple formats of the same 3D model which have been created through transformation events may be included in a single package and held in individual Representations (for example proprietary BIM models and STEP-IFC derivations). Information related to a transformation event (software used, parameters etc) should be included in the relevant documentation/paradata folder and details of the event recorded in PREMIS.

3D models of the same physical object can be created via different means (e.g point cloud vs CAD based) and it is a matter for the repository to determine if it considers these models to be the same or a different intellectual entity and thus whether they should be in the same or different information packages. Should models created via different methods be included in the same package then Rights information and Paradata may differ for each model and should therefore be held at Representation level. Submission Agreements between Producer and Archive should detail whether different capture/creation versions of the same physical object are permitted in a single information package.

It is strongly recommended that as much documentation as possible related to the creation, context, provenance and terms of use of the model is included in the information package such as for example: licence terms, project overview, model creation and transformation information. The CITS 3D HM recommends the inclusion of such information related to the production of the model in a documentation/paradata sub-folder and related to archiving and use of the model in a documentation/other sub-folder. If model Authentication information (such as data quality rules) is included with the model then this can also be included in the optional documentation/authentication folder. Further information on model Authentication can be found in the CITS 3D PM at:

https://github.com/DILCISBoard/CITS-3DPM/tree/rel/v1.0.0. The inclusion of a Submission Agreement in the information package is recommended and should be located in the documentation/other sub-folder.

Individual folder structures are recommended within each representation/data folder to provide structure appropriate to the model format used.

4.2 Metadata and Supporting Information

4.2.1 Descriptive Metadata

The package should contain general Descriptive Metadata about the digitised object and may contain specific Descriptive Metadata about the individual 3D model Representations. According to the CITS Archival Description "Descriptive metadata can be expressed according to many current standards (i.e., MARC, MODS, Dublin Core, TEI Header, EAD, VRA, FGDC, DDI) or a locally produced XML schema".

Users may find that existing standards do not provide sufficient elements to adequately describe the creation of 3D models. This can be overcome through the use of specialised 3D Model schemas (i.e. CARARE, LIDO) and creation of locally produced, well formed schema as described in the CSIP and CITS Archival Information. The use of standards however is encouraged and greater extensibility can be achieved through use of for example the CIDOC-CRM (Common Resource Model) and extensions. Serialisation methods for metadata are not prescribed in the CSIP but use of METS and PREMIS which are only serialised in XML is mandatory, and hence encoding of additional, descriptive metadata in XML is preferable. Although CIDOC-CRM is primarily encoded in RDF, a schema is available for serialisation in XML.

4.2.2 Preservation Metadata

Good practice in the management of 3D Heritage Model data requires that Information Packages include Preservation Metadata, specifically information related to:

- Provenance
- Reference
- Fixity
- Context

According to the CITS PREMIS: "When using preservation metadata together with the Common Specification for Information Packages (CSIP) (http://earkcsip.dilcis.eu), it is recommended that these are included in the information package in PREMIS format. Although this is not mandatory, all tools claiming to be able to validate CSIP compliant Information Packages must also be able to validate PREMIS metadata once it exists within the package. The two high-level requirements for the use of PREMIS in Common Specification IPs are that:

- All preservation metadata is created according to official PREMIS guidelines;
- All PREMIS metadata is referenced from the amdSec/digiprovMD element of the appropriate METS file.

It is recommended that users review the CSPM.

CITS 3D HM recommends the inclusion of PREMIS files at Representation level to include provenance, rendering environment, origination information and if it exists Authentication information.

4.2.3 Rights Metadata

From the Principle above, detailed Rights information should be recorded within PREMIS. As such a PREMIS file should be included at either package or in each Representation to contain this.

4.2.4 Paradata

Paradata is defined as supporting textual material which concerns the conditions under which the model digitization (creation) took place, (who did it, where, what were the environmental conditions), the intended purpose of the model (documentation, research, communication), the reasons of decisions taken in the digitization process leading to the final result, and for reconstructions, the references to the documentation supporting the reconstruction (images, drawings, reports, archaeological/historical interpretation).

Standards do not exist for the recording of Paradata but considerable work has been done in recommending best practice in the documentation of the process of creation of 3D models. Further information is included in the accompanying guideline to this specification.

Within this 3D Heritage Model specification a specific folder for Paradata is defined as a sub-folder of the Documentation folder at root or Representation level. Within the CITS the concept of Paradata is extended to include any unstructured textual material related to the creation, processing, transformation, rendering or quality or the model(s).

4.2.5 Creating, Transforming and Rendering Hardware and Software

Details of the hardware and software used in the creation and transformation and required for the rendering of Heritage 3D Models can be captured in some specialised Descriptive Metadata schemas, described in event based ontologies such as the CRMdig extension of CIDOC-CRM, captured in PREMIS preservation metadata within object (creating application and extensions), environment function (rendering environment), event and agent entity semantic units or can be recorded as simple, unstructured text.

4.2.6 Quality Information

Quality information may be included in some specialised metadata schemas or if unstructured can be included in the documentation/paradata folder of each representation.

4.2.7 Authentication Information

If Authentication documentation exists then the optional documentation/authentication subfolder can be used at both root and representation level. Authentication events can be recorded within PREMIS and examples are given in the CITS 3D PM specification and guideline as to how this can be done (see https://github.com/DILCISBoard/CITS-3DPM/tree/rel/v1.0.0).

4.2.8 Processing Information

Processing information unless specifically encoded according to specialised schemas or ontologies can be included in the documentation/paradata sub-folder either at root or Representation level.

4.2.9 Geometry Data

Geometry data is closely aligned with the data model itself and should be included in the data folder within each Representation alongside the model data. Specific folder structures can be defined within the data folder for different model types or derivations.

4.3 Standards

4.3.1 METS

The main requirement for METS files in a CSIP Information Package is that they follow the official METS Schema version 1.12 http://www.loc.gov/standards/mets/mets-schemadocs.html (used by CSIP version 2.2) and the extension schema developed for CSIP and published by the DILCIS Board. As new versions of METS Schema become available the DILCIS Board will evaluate these and, if necessary, update the CSIP.

The METS specification requires a METS profile document which describes the use of METS and the METS elements. All the requirements described in this specification are also expressed with METS profiles for the CITS 3D HM root and representation METS which can be found at: https://github.com/DILCISBoard/CITS-3DHM/tree/rel/v1.0.0.

CSIP specifies that METS files be located at the root of the package folder structure (Root METS) and optionally in each of the Representations within its respective root folder (Representation METS). CITS 3D HM has a requirement to contain at least one Representation and so will contain at minimum a root METS and a Representation METS file.

4.3.2 PREMIS

From the EARK Common Specification for Preservation Metadata:

"When using preservation metadata together with the Common Specification for Information Packages (CSIP) (https://dilcis.eu/specifications/common-specification), it is recommended that these are included in the information package in PREMIS format. Although this is not mandatory, all tools claiming to be able to validate CSIP compliant Information Packages must also be able to validate PREMIS metadata once it exists within the package". The two high-level requirements for the use of PREMIS in Common Specification IPs are that:

- All preservation metadata is created according to official PREMIS guidelines;
- All PREMIS metadata is referenced from the amdSec/digiprovMD element of the appropriate METS file.

Further, to enhance the interoperability of the CSIP and to strengthen the management of information packages (IPs) in an archive, this specification imposes additional requirements regarding the use of PREMIS for describing IPs. The principles adopted in the CSIP for deciding which additional PREMIS semantic units are required are:

- PREMIS should be used to record detailed technical metadata;
- Technical information should be included in PREMIS metadata by using extension schemas;
- Information about agents carrying out preservation actions must be recorded in the PREMIS metadata (this is because the METS agents describe agents relevant for generic IP level events, such as the creation or submission of the package, not the preservation of the data);
- Event descriptions should be included in PREMIS metadata. Use of the official PREMIS event vocabulary (https://id.loc.gov/vocabulary/preservation/eventType.html) is recommended (note that more elaborate descriptions can be made than are made in this CITS);
- Detailed rights information should be included in PREMIS. High-level rights information in METS indicates restrictions. Detailed, object-specific rights information will be included in the PREMIS metadata;
- File format information for all files should be included as Persistent Unique Identifier (PUID) values in the appropriate PREMIS semantic units.

Desirable technical and preservation metadata in the context of the 3D HM CITS can include:

- Reference to the content information standard (e.g. STEP-IFC);
- Information about the generating system (Creating Agent software);
- A description of the environment used to render the model.
- Model creation, authentication and transformation events
- Softwarem hardware, individual and organisational agents
- Rights statements including: copyright, licenses and access restrictions

The use of PREMIS within the 3D HM CITS requires that information packages also conform to the CSPM (use of PREMIS is not mandatory within CSIP, but is a SHOULD requirement in CITS 3D HM). If used, the following PREMIS semantic units should be considered:

Characteristic	PREMIS Semantic Unit or Container
File format	object/objectCharacteristics
File fixity	object/objectCharacteristics
Generating software	object/creatingApplication
Content information standard	object/significantProperties
Rendering environment	object/environmentDesignation

Model creation, authentication or transformatiuon events	event
Software, hardware, individual or organizational agents	agent
Copyright	rights/copyrightInformation
License	rights/licenseInformation
Access	rights/otherRightsInformation

5 Glossary

Term	Description
Archival Creator	Organisation unit or individual that creates records and/or manages records during their active use.
Archival Information Package (AIP)	An information package, consisting of the Content Information and the associated Preservation Description Information (PDI), which is preserved within an Open Archival Information System (OAIS).
Authentication	Process of establishing the authenticity or quality of a digital object either through tests against given rules or through validation of Digital Signatures.
Building Information Modeling (BIM)	A process involving the generation and management of digital representations of the physical and functional characteristics of buildings and other physical assets.
Cardinality	The term describes the possible number of occurrences for elements in a set. The numbers have the following meanings: (1..1) – in each set, there is exactly 1 such element present (0..1) – the set can contain from 0 to 1 of such elements (1..n) – the set contains at least one element (0..n) – the set can contain up to n of such elements, but it is not mandatory

	(0..0) – the element is prohibited to use
Computer Aided Design (CAD)	A way to digitally model 2D drawings and 3D models of real-world objects (mobile and immobile).
Content Data Object	The Data Object, that together with associated Representation Information comprises the Content Informartion (Source OAISA – ISO 14721:2012)
Content Information	A set of information that is the original target of preservation or includes part or all of that information. It is an Information Object composed of its Content Data Object and its Representation Information. (Source OAIS – ISO 14721:2012)
Data File	A component which contains data and has an associated MIME file type. A Data File can encapsulate multiple bit streams and metadata according to a standard such as a DICOM but must have a recognised MIME file type. A Data File may comprise one or more subsidiary Byte Streams; for example, an MP4 file might contain separate audio and video streams, each of which has its own associated metadata.
Data Quality Rules	Data Quality or Verification rules ensure that a representation of a Model meets quality requirements within defined tolerances, i.e. that a specific representation represents the Model with sufficient accuracy.
Derived Representation	A transformation of the native data, which may be based on a Native Format or a Standardised Format, e.g. an html version may be derived from a text document as an alternative representation
Digital Signature	An Digital Signature is a defined method to sign an premis:object in electronic environments; it provides means to authenticate the signatory and the signed premis:object in an unambiguous and safe way by attaching to or logically associating data in electronic form to other electronic premis:objects
Dissemenation Information Package (DIP)	An Information Package, derived from one or more AIPs and sent by Archives to the Consumer in response to a request to the OAIS.
Document	A single or group of related Data Files with common metadata. For example, a Document may consist of a PDF file together with associated attachments or a word file with a separate image signature sheet. A document can be

	considered to be an entity that is approved/signed as a whole by a practitioner.
Heritage Model	A 3D digital model produced by any means of a cultural heritage asset such as an object, building, or archeological site.
Information Package	A logical container composed of optional Content Information and optional associated Preservation Description Information used to delimit and identify the Content Information and Package Description information used to facilitate searches for the Content Information.
Internal Archival Long Term Preservation guidelines	This type of guideline can have different names depending on the creator. Generally, archives specify technical guidelines and/or regulations for formats, specifying what they will accept and maintain for the long term/ Depending on the archive and available technical resources, the criteria for the selected formats can differ from archive to archive.
Level	The level of requirements of the element following RFC 2119 http://www.ietf.org/rfc/rfc2119.txt **MUST –** this means that the definition is an absolute requirement **SHOULD** – this means that in particular circumstances, valid reasons may exist to ignore the requirement, but the full implications must be understood and carefully weighed before choosing a different course. http://www.ietf.org/rfc/rfc2119.txt **MUST NOT** – this means that the prohibition described in the requirement is an absolute prohibition of the use of the element. **SHOULD NOT** – this means that in particular circumstances, violating the prohibition described in the requirement is acceptable or even useful, but the full implications should be understood and the case carefully weighed before doing so. The requirement text should clarify such circumstances. **MAY** – means that a requirement is entirely optional.
Open Archival Information System (OAIS)	An Archive consisting of an organisation, which may be part of a larger organisation, of people and systems, that has

	accepted the responsibility to preserve information and make it available for a Designated Community. It meets a set of responsibilities that allows an OAIS Archive to be distinguished from other uses of the term 'Archive'.
Original Model or Native Representation	Used specially to keep the design intent for long term archiving in the contact of certification and legal requirements for proof. It can be stored in native or standardised formats.
Paradata	Documentation of the evaluative, analytical, deductive, interpretative and creative decisions made in the course of computer-based recorded in such a way that the relationship between research sources, implicit knowledge, explicit reasoning, and visualisation-based outcomes can be understood.
Preservation Description Information (PDI)	The information which is necessary for adequate preservation of the Content Information and which can be categorised as Provenance, Reference, Fixity and Access Rights Information.
Producer	The producer is an organisation, person, or client system, which provides the information to be preserved. This can include other archives or internal archive personnel or system components. Typical roles of type "producer" may be. System Designers, Design Engineers, Subcontractors, Manufactures or Test Engineers. (Source EN 9300)
RDBMS	Relational Database Management System
Representation	A Representation within an Infiormation Package contains archival data. If an Information Package contains the same data in two or more different formats (i.e. an original and a long term preservation format) or in different types of organisations (arrangements), the are placed within two or more separate Representations within the Representations folder of the Information Package of the Information Package.
Representation Information	The Representatiuon Information must enable or allow the re-creation of the significant properties of the original data object.
Standardised Machine-readable Documentation	A standardised machine-readable document is a document whose content can be readily processed by computers and is based on a commonly accepted standard. Such documents are distinguished from machine-readable data

	by virtue of having sufficient structure to provide the necessary context to support the business processes for which they are created.
Standardised Open Format	A format of data in a syntax which is derived by as broad community, such as ISO and which is independent of a specific system or interface. “Open” means completely and precisely documented in syntax and semantics and is applicable for free. In addition, standardisation processes regulate the change process for the standard. (Source EN9300)
Submission Agreement	The agreement reached between an archive and the submission producer that specifies a submission format (eHealth1 CITS), and any other arrangements needed, for the data submission session. Any special conditions on patient confidentiality could be specified in the submission agreement.
Submission Information Package (SIP)	An Information Package that is delivered by the Producer to the OAIS for use in the construction or update of one or more AIPs and/or the associated Descriptive Information.
Submitting Organisation	Name of the organisation submitting the package to the archive.
Validation	Validation is applied to guarantee the integrity of the content of a Document throughout the entire process of long-term archiving. Validation is typically performed by calculation and comparison of Validation Properties which are like a fingerprint for the content of the Document. Each change of the content changes one or more attributes of the Validation Properties.
Verification	A process to ensure that data is correctly represented (e.g in a package representation). Verification rules ensure that a data representation meet he quality requirements within defined tolerances. Verification rules at domain specific (CAD, PDM, Electronic Assembly, Fluid Dynamics) and are defined within EN 9300. (Source EN 9300)

Table 2: Glossary

6 Vocabularies

6.1 Vocabulary3D HM

Value	Description
cits3D HM_v1_0	Content Information Type : `mets/@csip:CONTENTINFORMATIONTYPE`
Heritage Model Data	Other Content Category: `mets/@csip:OTHERTYPE`
Other Documentation	Other Documentation fileGrp and division label: `mets/structMap[@LABEL='CSIP']/div/div/div[@LABEL='Other Documentation']`
Paradata Documentation	Paradata documentation fileGrp and division label `mets/structMap[@LABEL='CSIP']/div/div/div[@LABEL=Paradata Documentation']`
Authentication Documentation	Authentication documentation fileGrp and division label `mets/structMap[@LABEL='CSIP']/div/div/div[@LABEL=Paradata Documentation']`

Table 3: Vocabularies

7 Postface

AUTHOR(S)	
Name(s)	Organisation(s)
Stephen Mackey	Penwern Limited, E-ARK Consortium

REVIEWER(S)	
Name(s)	Organisation(s)
[Name]	[Affiliation]
[Name]	[Affiliation]
[Name]	[Affiliation]

Project co-funded by the European Commission within the ICT Policy Support Programme Dissemination Level		
P	**Public**	x
C	**Confidential, only for members of the Consortium and the Commission Services**	

REVISION HISTORY AND STATEMENT OF ORIGINALITY

Submitted Revisions History

Revision No.	Date	Authors(s)	Organisation	Description
V1.0.0	30/08/2024	Stephen Mackey	Penwern Limited, E-ARK Consortium	Review version

Statement of originality:
This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.