










Preflight Summary Report for: book-a2a-fix.pdf

Profile: Convert to PDF/A-2a (Processed pages 1 to 17)



Processed by simon, Date: 01/12/2023 17:37

Fixups

-  Convert to PDF/A-2a (1 object)
-  Remove, apply or adjust object compression (1 object)
-  Compress all uncompressed objects using lossless ZIP compression (1 object)
-  Recompress LZW as ZIP (1 object)
-  Fix font encoding (CIDToGIDMap) (15 objects)
-  Insert missing Type entry in StructElem objects (530 objects)
-  Make document XMP Metadata compliant with PDF/A-2 (1 object)
-  Remove CIDSet if incomplete (15 objects)
-  Repair invalid bookmark hierarchies (1 object)

Results (Summary)

Error

-  Implementation limit: Name object with a length greater than 127 byte
-  ToUnicode CMap contains a Unicode value of zero (112 matches on 11 pages) - 4, 7, 8, 9, 10, 11, 12, 13, 14, 16, 17

Document information

File name: "book-a2a-fix.pdf"

Path: "C:\Users\simon\OneDrive\Documents\GitHub\fsci\3d-digitisation-and-preservation"

PDF version number: "1.4"

File size (MB): 1.2

Title: "3D Digitisation and Preservation – Survey Results from the NFDI4Culture Community"

Author: "Jörg Heseler"

Creator: "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/119.0.0.0 Safari/537.36"

Producer: "Skia/PDF m119"

Created: "01/12/2023 14:10"

Modified: "01/12/2023 17:37"

Keywords: "NFDI4Culture; survey; 3d model; 3d object; digital preservation; digitisation; research data; metadata scheme; file format"

Subject: "One of the project goals of NFDI4Culture – Consortium for Research Data on Material and Immaterial Cultural Heritage and specifically the Task Area Data Publication and Availability (TA4) is the development of concepts for central long term preservation services to ensure a sustainable infrastructure. In particular, publications from cultural disciplines with 3D data are in the foreground. The requirements for developing a concept are the needs of producers, providers and administrators of research data. One measure of this needs assessment was the implementation of a survey on the creation, use and handling of 3D models, the results of which are discussed below."

Trapping: "Unknown"

Additional actions

Page open action

Type: Go to page in current document

Page open action

Type: Go to page in current document

Page open action

Type: Go to page in current document

Page open action

Type: Go to page in current document

Number of plates: 4

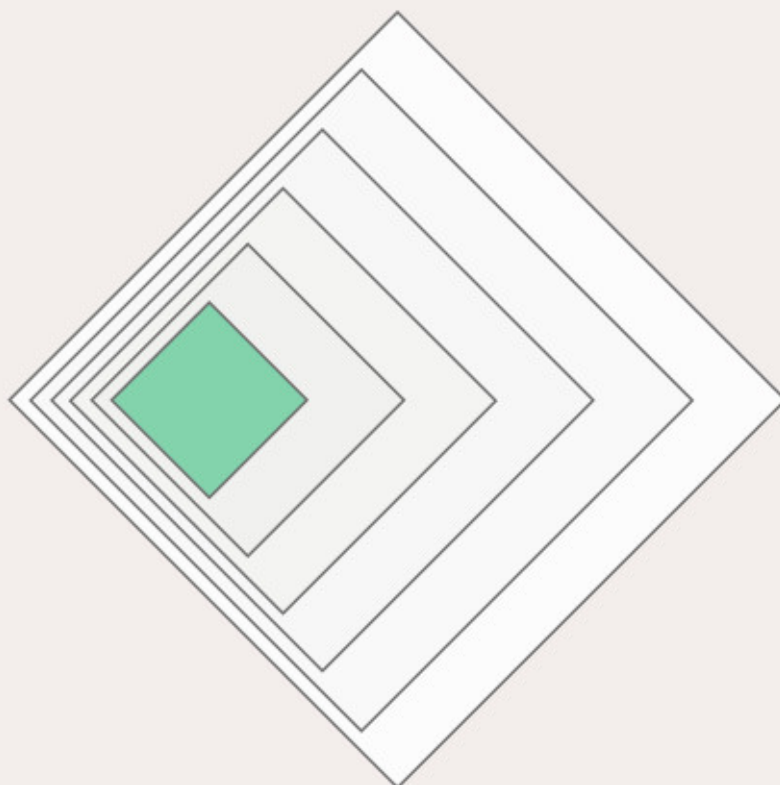
Names of plates: "(Cyan) (Magenta) (Yellow) (Black) "

Environment

Preflight, 18.6.0 (268)

Acrobat version: 23.006

Operating system: Microsoft Windows 10 Home Edition (Build 22000)



3D Digitisation and Preservation

Survey Results from the
NFDI4Culture Community

Heseler, Jörg
Arnold, Matthias
Büttner, Alexandra
Worthington, Simon



3D Digitisation and Preservation

Survey Results from the NFDI4Culture Community

Jörg Heseler, Matthias Arnold, Alexandra Büttner and Simon
Worthington

Version: 1.0.0, December 2023

Published by: NFDI4Culture, Digital Publications and Data Working Group

Last updated: 2023-11-29

Created: 2023-10-04

Language: English (United Kingdom)

Created by: Editorial Team

Contents

Survey results	7
Introduction	7
Analysis	7
1. Participants and disciplines	7
2. Digitisation methods	8
3. Digitised objects	9
4. Metadata schemas	9
5. Digitisation software	10
6. Capture hardware	11
7. Display devices for 3D data	12
8. 3D formats	12
9. 3D data viewing software	13
Conclusion and outlook	13
Appendix: Survey data	14
Acknowledgements	14
Imprint	16

All problems according to Preflight profile
Convert to PDF/A-2a

Survey results

Introduction

One of the project goals of NFDI4Culture – Consortium for Research Data on Material and Immaterial Cultural Heritage and specifically the Task Area Data Publication and Availability **TA4** is the development of concepts for central long term preservation services to ensure a sustainable infrastructure. In particular, publications from cultural disciplines with 3D data are in the foreground. The requirements for developing a concept are the needs of producers, providers and administrators of research data. One measure of this needs assessment was the implementation of a survey on the creation, use and handling of 3D models, the results of which are discussed below.

Analysis

1. Participants and disciplines

This survey on 3D digitisation and long-term preservation was conducted during the NFDI4Culture Forum "3D objects: digitisation, presentation and preservation" in May 2022. This forum was jointly organized by the task areas Data capture and enrichment of digital cultural assets **TA1** and Data publication and data availability **TA4**. The aim of the event was to support the exchange between 3D data providers and repository operators in order to identify needs and problems and to work together to solve them in the future. The main goal of this survey was to identify used file formats and metadata schemas in the 3D area used by the NFDI4Culture communities, i.e. media and music studies as well as architecture, art history and performing arts.

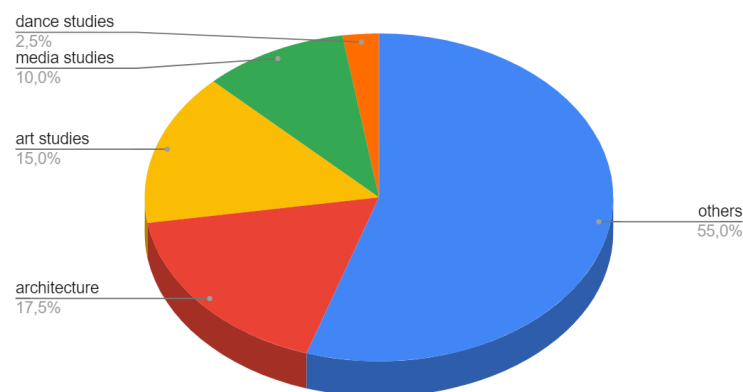


Figure 1 Which discipline would you classify yourself as?

First, we asked about the disciplines to which the respondents belong. Nine out of twenty participants (45%) assigned themselves to NFDI4Culture disciplines. Of these, architecture was most frequently represented with 17.5%, followed by art studies with 15%, media studies with 10% and dance studies with 2.5%. Participants were able to assign themselves to several specialist areas. No other personal data was collected.

2. Digitisation methods

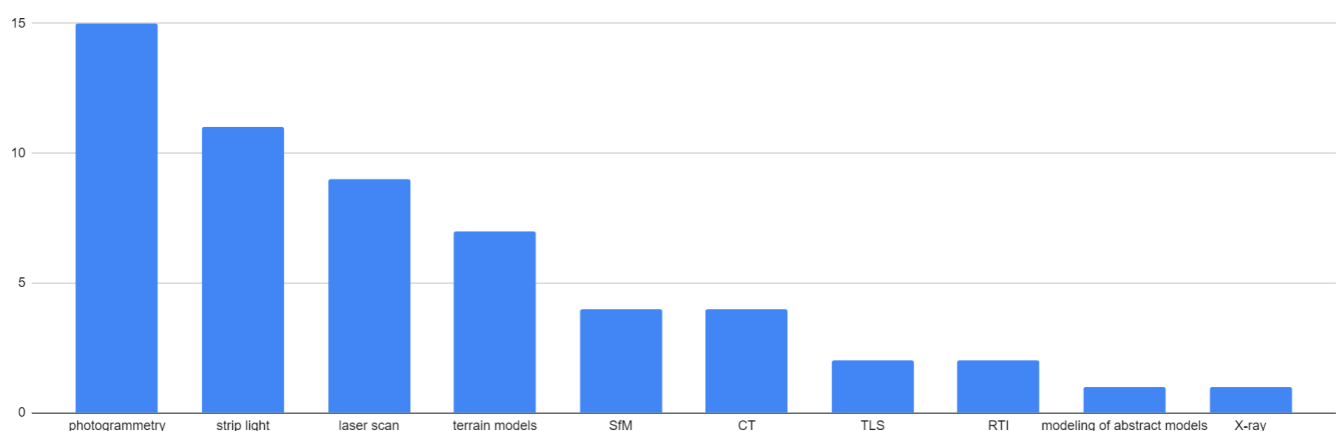


Figure 2 Which methods do you use to digitise 3D objects?

When asked which digitisation methods our participants used, photogrammetry was used fifteen times, followed by strip light methods eleven times, laser scanning methods nine times, as well as terrain models seven times, structure-from-motion (SfM) and computer tomography (CT) four times each, as well as terrestrial laser scanning (TLS) and Reflectance Transformation Imaging (RTI) twice each. Finally, the modeling of abstract models such as molecules but also the use of X-ray were mentioned once each.

3. Digitised objects

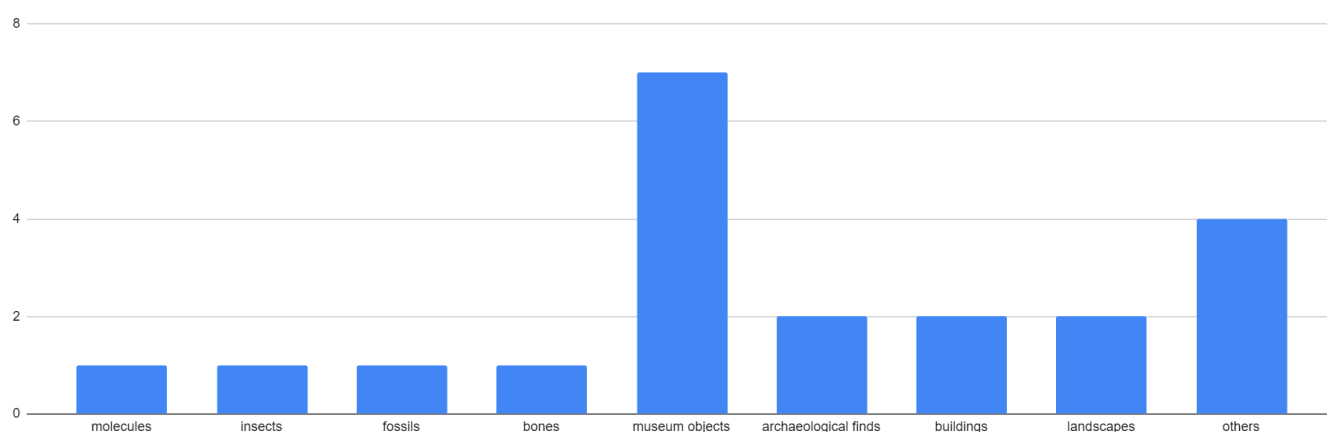


Figure 3 What are you digitising? (Ordered by object size)

With regard to the objects to be digitised, the variety in terms of size, material and surface structure was significant. Museum objects were frequently listed, from pottery to sculptures made from various materials such as ceramics or metal. Very large objects such as buildings and landscapes followed, which were mostly mentioned in connection with capture using stationary laser scanners or drones using photogrammetry methods. In addition to the previously mentioned modeled molecules, very small objects such as insects, fossils or even bones were also listed, for which CT or X-ray methods were mainly used.

4. Metadata schemas

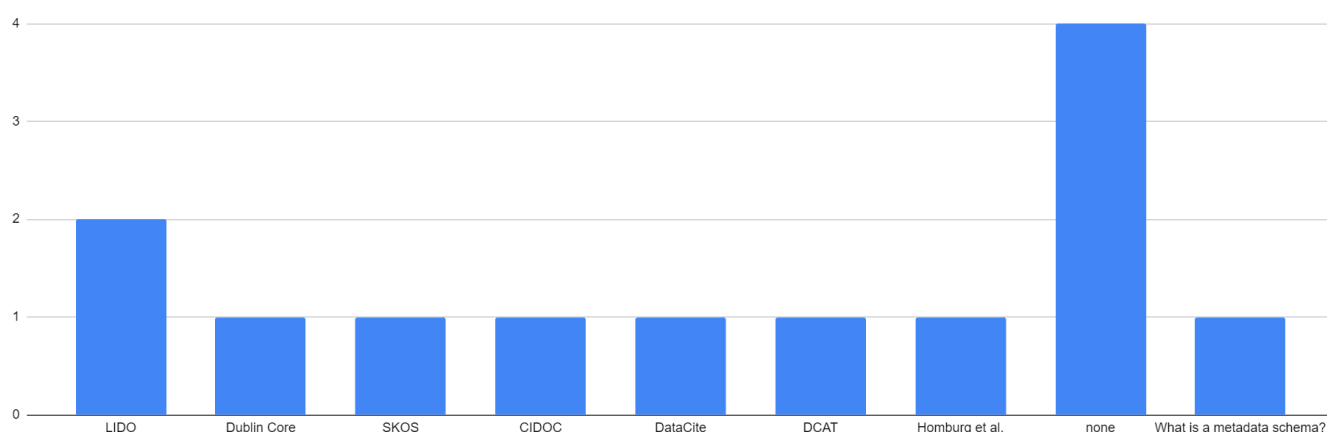


Figure 4 Which metadata schemas do you use?

When asked about the metadata schemas used to describe the physical and digital objects, the XML schema Lightweight Information Describing Objects (LIDO) was mentioned twice and Dublin Core, Simple Knowledge

Organization System **SKOS**, Conceptual Reference Model of the International Council **CIDOC**, DataCite, Data Catalog Vocabulary **DCAT** and a metadata schema by Homburg et al. once each. Data representation forms or concepts for structuring bibliographic information were also occasionally listed, which will not be discussed further here.

Finally, four respondents reported not using metadata schemas. One participant also stated that they were unfamiliar with the term metadata schema. Furthermore, seven participants did not provide any information on this question.

5. Digitisation software

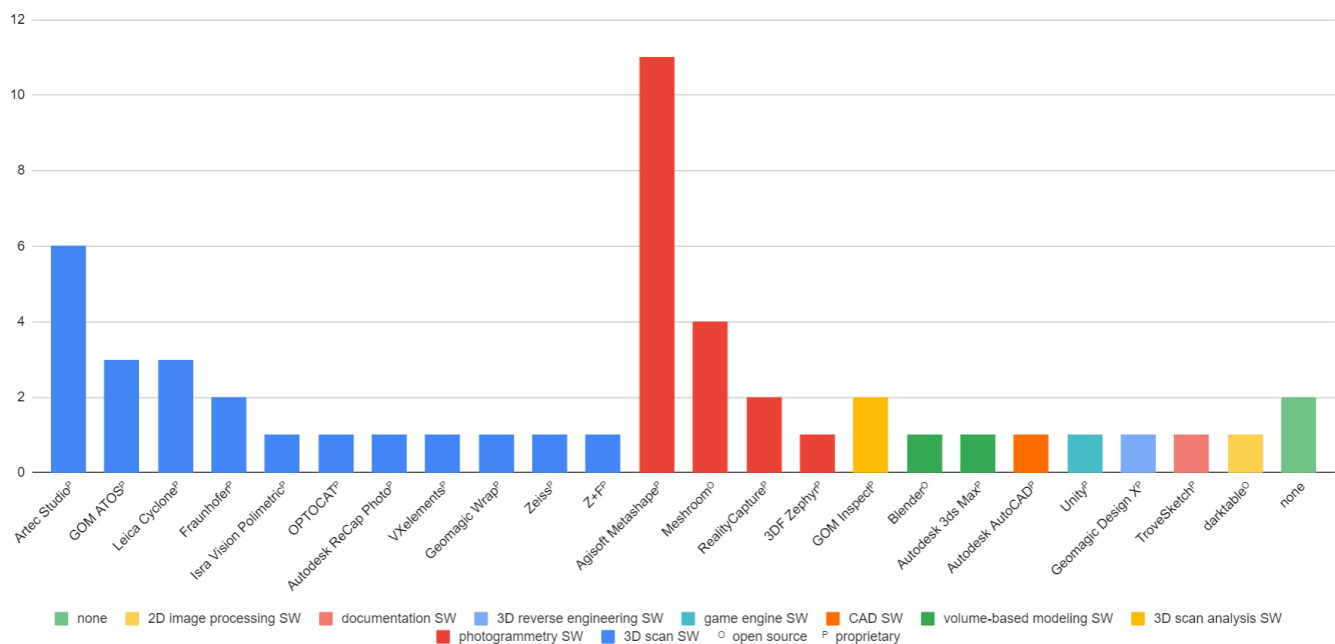


Figure 5 What software do you use to digitise 3D objects?

Regarding 3D digitisation software, 3D scanning and photogrammetry, 3D scan analysis, modeling software and a game engine, as well as 3D reverse engineering, documentation and 2D image processing software were mentioned. Artec Studio was listed as 3D scanning software six times, followed by GOM ATOS and Leica Cyclone three times each, and 3D scanning software from the Fraunhofer Institute for Graphical Data Processing twice. Isra Vision Polymetric, OPTOCAT, Autodesk ReCap Photo, VXelements, Geomagic Wrap and 3D scanning software from Zeiss and Z+F were also listed once each. In terms of photogrammetry software, Agisoft Metashape was mentioned eleven times, Meshroom four times, RealityCapture twice and 3DF Zephyr once. In addition, the 3D analysis tool GOM Inspect was listed twice. Finally, the 3D modeling software Blender and Autodesk 3ds Max, the CAD modeling software Autodesk AutoCAD, the game engine Unity, the 3D reverse engineering

software Geomagic Design X, the documentation software TroveSketch and the 2D image processing software darktable followed once each.

6. Capture hardware

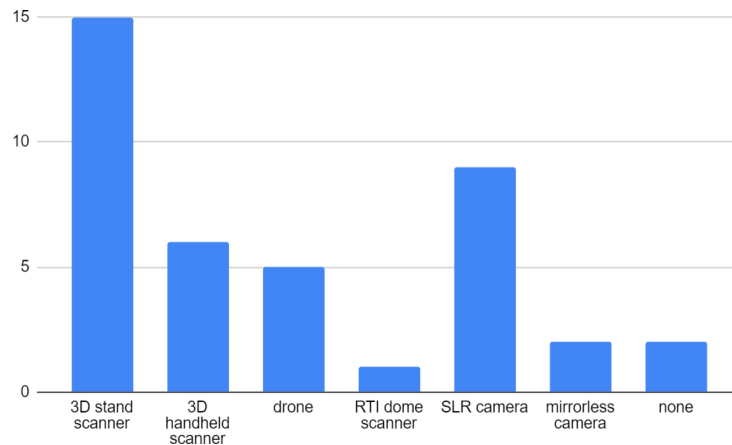


Figure 6 What hardware do you use to capture 3D objects?

In terms of hardware for object detection, stationary scanners were mentioned fifteen times, which primarily work with laser light and are used for medium to large objects. In addition, hand-held scanners were listed six times, particularly for detecting small to medium-sized objects, which use striped light or laser scanning methods. The use of drones to capture buildings and landscapes was mentioned five times. A Reflectance Transformation Imaging (RTI) dome scanner was also used to capture small to medium-sized objects. Furthermore, nine single-lens reflex (SLR) and two mirrorless cameras were listed, with which photos were taken either for documentation purposes or to calculate 3D models using photogrammetric methods.

7. Display devices for 3D data

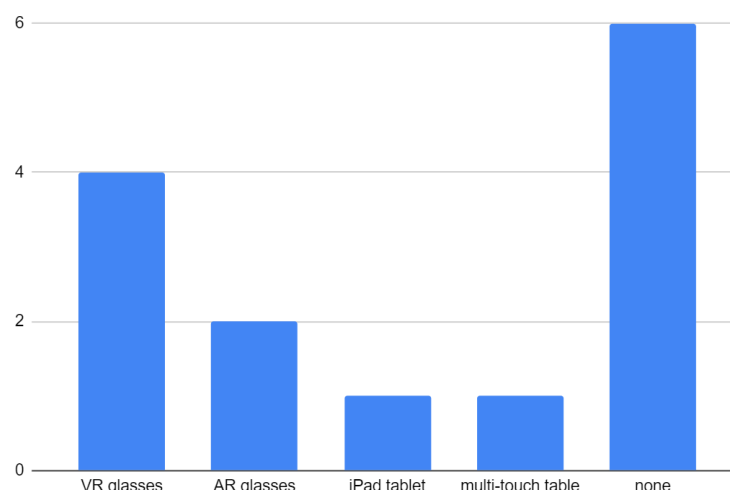


Figure 7 Do you use special hardware to display your 3D data?

Six participants stated that they did not use any special devices to display 3D data. Apart from that, virtual reality VR glasses were mentioned four times, including the Oculus Rift glasses twice and the Oculus Quest and HTC Vive once each. The augmented reality AR glasses Microsoft Hololens were also listed twice. An Apple iPad tablet and a Multi-touch table were also mentioned once each.

8. 3D formats

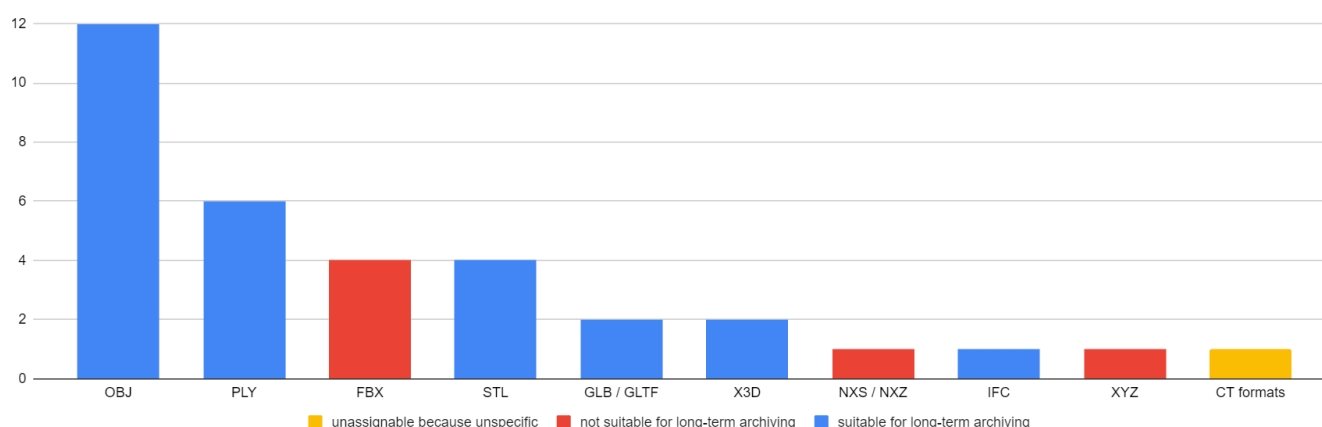


Figure 8 Which 3D formats do you work with for digitisation, presentation or digital preservation?

The most important question for us in this survey was which 3D formats the community uses for various application scenarios, such as digitisation, presentation and digital long-term archiving. The participants listed Wavefront OBJ twelve times, Stanford Polygon PLY six times, Filmbox

FBX and Stereolithography STL four times each, Graphics Language Transmission Format GLTF or the binary variant GLB and Extensible 3D (X3D) twice each. In addition to the 3D format Nexus NXS or the compressed variant NXZ, the Industry Foundation Classes IFC format, which is common in construction, as well as the point cloud format XYZ and general CT formats were each mentioned once.

9. 3D data viewing software

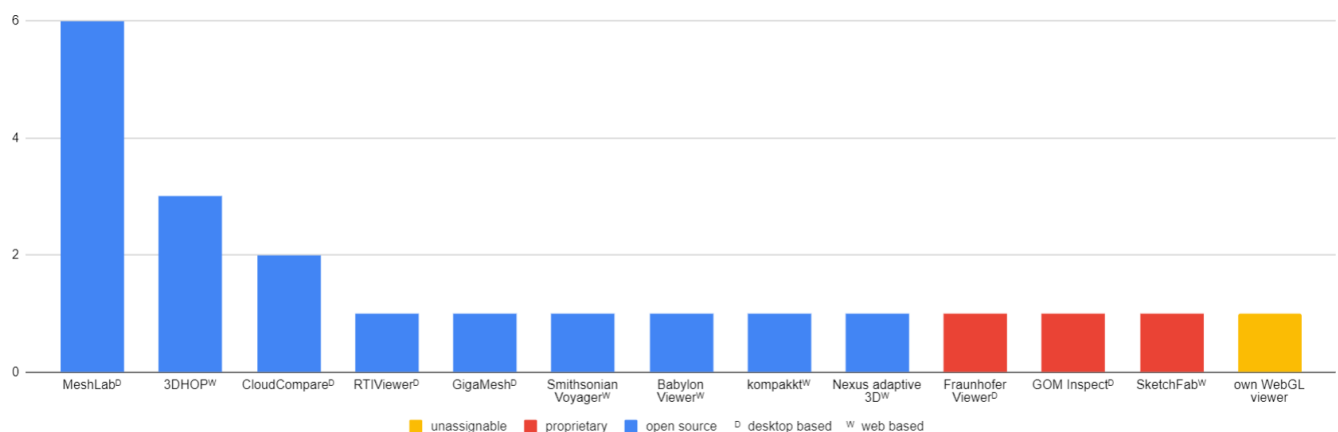


Figure 9 Do you use special software to display your 3D data?

Finally, we asked what software is used to display 3D data. MeshLab was listed six times, the 3DHOP Viewer, which is based on the Nexus adaptive 3D viewer, three times, and CloudCompare twice. The RTIViewer, GigaMesh, Smithsonian Voyager, the Babylon Viewer and the kompakkt viewer, which is based on the Babylon.js framework, were also mentioned once. A viewer from the Fraunhofer Institute for Graphical Computing as well as GOM Inspect and SketchFab were also listed. One participant reported using an own WebGL-based viewer.

Conclusion and outlook

This survey gave us valuable insight into the formats, devices and software used by our community for 3D digitisation and long-term archiving. This lays the foundation for further investigations and developments.

A first interesting aspect of this survey was the question about the use of metadata schemas. It remains unclear why seven out of twenty participants (35%) did not provide any information here. In addition, current metadata schemas, e.g. CARARE or Extensible Metadata Platform XMP, were not mentioned. The latter raises the question of the extent to which existing schemas for 3D data are accepted by the community and meet their needs. To better understand this, a deeper study of this topic would be required.

A second interesting aspect was the question of 3D formats used. Formats such as Extensible 3D **X3D** and Graphics Language Transmission Format **GLTF** or Digital Asset Exchange **DAE** are particularly suitable for long-term archiving due to their high degree of disclosure, self-documentation and timeliness. However, these were only mentioned in the minority or the latter was not mentioned. Formats that are not suitable for long-term archiving were also listed: Filmbox **FBX** is a proprietary format, which means that there is no openly available format specification that describes the exact structure of an FBX file. If this format ceases to be widely used in the future, long-term archives may not be able to develop their own tools for rendering or converting FBX files. In the case of Nexus **NXS** or **NXZ**, the distribution is too low to ensure that there will still be tools that can display this format in 20 years. In the text-based point cloud format XYZ, the data structure is not clear, which is why this format is unsuitable for software-independent exchange. The standardised ASTM E57 format would be an alternative to the long-term archiving of 3D point clouds. However, we have not yet examined CT formats for their archivability. Finally, the information provided by the participants also showed that, on the one hand, when querying formats in further surveys, we must differentiate between the purposes and, on the other hand, continue to point out the advantages of 3D formats that can be archived for a long time.

It is also noticeable that the 3D scanning software used is largely proprietary and device-specific. Likewise, most of the photogrammetry software used is not freely accessible. However, the large number of open source viewers mentioned for displaying 3D data is positive.

Future surveys will expand on these findings and focus on specific 3D use cases in addition to long-term preservation purposes. To draw general conclusions, it is crucial to reach a larger number of participants and better understand their specific backgrounds and needs.

Finally, it is planned to create a guide for the digital long-term archiving of 3D cultural heritage objects based on collected findings and further investigations. This guide will help establish good practices in the NFDI4Culture communities and promote collaboration between different actors in the field of cultural heritage.

Appendix: Survey data

Anonymised survey source data:

<https://github.com/NFDI4Culture/3d-digitalisierung-und-langzeitarchivierung/tree/main/survey-data>

Acknowledgements

The authors would like to thank all participants for their valuable contributions and commitment to this survey. Your insights will help

advance the future of 3D digitisation and long-term preservation of cultural heritage objects. We would also like to thank the following colleagues for their helpful comments: Eva Bodenschatz and Sabrina Herzog.

Imprint



NFDI4Culture – Consortium for Research Data on
Material and Immaterial Cultural Heritage,
<https://nfdi4culture.de/>

Authors:

Heseler, Jörg; ORCID iD [0000-0002-1497-627X](https://orcid.org/0000-0002-1497-627X); Saxon State and
University Library Dresden (SLUB Dresden); ROR ID <https://ror.org/03wf51b65>

Arnold, Matthias; ORCID iD [0000-0003-0876-6177](https://orcid.org/0000-0003-0876-6177); Ruprecht-Karls-
Universität Heidelberg, Universitätsbibliothek; ROR ID <https://ror.org/038t36y30>

Büttner, Alexandra; ORCID iD [0000-0002-4950-0941](https://orcid.org/0000-0002-4950-0941); Ruprecht-Karls-
Universität Heidelberg, Universitätsbibliothek; ROR ID <https://ror.org/038t36y30>

Worthington, Simon; ORCID iD [0000-0002-8579-9717](https://orcid.org/0000-0002-8579-9717); German National
Library of Science and Technology (TIB); ROR ID <https://ror.org/04aj4c181>

Citation:

Heseler, Jörg; Arnold, Matthias; Büttner, Alexandra; & Worthington, Simon.
2023. *3D Digitisation and Preservation – Survey Results from the
NFDI4Culture Community*. Digital Publications and Data Working Group
(NFDI4Culture). <https://doi.org/10.5281/zenodo.8033413>.

German version: Heseler, Jörg; Arnold, Matthias; Büttner, Alexandra; &
Worthington, Simon. 2023. *3D-Digitalisierung und Langzeitarchivierung –
Umfrageergebnisse aus der NFDI4Culture Community*. Arbeitsgruppe
Digitale Publikationen und Daten (NFDI4Culture).
<https://doi.org/10.5281/zenodo.8033536>.

HTML version:

<https://nfdi4culture.github.io/3d-digitisation-and-preservation/>

German version: <https://nfdi4culture.github.io/3d-digitalisierung-und-langzeitarchivierung/>

Keywords:

NFDI4Culture; survey; 3d model; 3d object; digital preservation; digitisation; research data; metadata scheme; file format

Funding:

NFDI4Culture and this research are funded by the German Research Foundation **DFG** under project number 441958017.

DFG <https://www.dfg.de/>; ROR ID <https://ror.org/018mejw64>

Licenses:



This work is licensed under a Creative Commons Attribution 4.0 International License.

© Copyright of the texts: The Authors, 2023.

Fonts:

Inter. These fonts are licensed under the SIL Open Font License, Version 1.1. The Inter project is led by Rasmus Andersson. <https://github.com/rsms/inter/>

Playfair Display. These fonts are licensed under the SIL Open Font License, Version 1.1. Designer is Claus Eggert Sørensen. <https://github.com/clausteggert/Playfair>

Roboto Slab. These fonts are licensed under the Apache License, Version 2.0. Designer is Christian Robertson. <https://github.com/googlefonts/robotoslab>

Technical information:

ADA Semantic Publishing Pipeline: <https://github.com/NFDI4Culture/ADA-semantic-publishing>

Multi-format CSS theming: [Interpunct.dev](https://interpunct.dev)