Final Report

LIS 545 - Data Curation I

Erik Mendoza

Data and Metadata Profile

([Link](https://figshare.com/articles/dataset/3D_data_for_recreating_a_bronze_portrait_figure_from_Cleveland_Museum_of_Art/5234890))

The data to be analyzed are a collection of various image files which are cumulatively intended for a 3d modeling software. When this data is input into the proper software, the output would be a recreation of *Bronze portrait of a man*, a 1st century bronze head. The original head was found in Northern Italy and is believed to be a likeness of Marcus Vipsanius Agrippa, but is possibly some other high-ranking Roman official. It is an excellent example of the realistic portraiture of the late Republican Period of the Roman Empire. The statue’s body, however, is lost to time. Although the data states that the original figure is housed at the Cleveland Museum of Art, the piece now exists at the Metropolitan Museum of Art.

The 3D modeling data come from a technologically-minded archaeologist who creates 3D representations of priceless pieces of art in an attempt to preserve them in a digital format. The key stakeholders are the creator, Daniel Pett, as well as the organizations that fund and utilize Daniel’s work, The Fitzwilliam Museum and The University of Cambridge to name a couple. The Cleveland Museum of Art is also a stakeholder as the original user of the model and, presumably, the entity which called for its creation in the first place. Interestingly, the MoMA is also a key stakeholder through its decision to not use the model in its digital collection. The data itself include 81 photographs of the statue’s head (JPG), monochrome masks of the same photos (PNG), the final 3D model file in four file formats (JPG, MTL, OBJ, and STL), a README.md file and a LICENSE.md file. Although a specific program would be required to render the images into a 3D model, the specific program used for this model is not named. After some research, it was discovered that the MTL file is opened in Alias/Wavefront, the OBJ file can be opened by a number of 3D CAD programs such as Rhino or Blender, and STL files are native to CAD programs created by the company 3D Systems. The data itself is licensed under creative commons and no other usage restrictions are mentioned.

The metadata exists within the text files included with the data and is entirely unstructured. The README.md file includes metadata pertaining to the object being analyzed such as the title, approximate time and location of its creation, and a brief description. The LICENSE.md file contains metadata specifically about the creative commons license associated with the data. The metadata does not include a lot of information about the model itself, such as what the intention of its use is or what software(s) were used to make it. None of the metadata is structured to a standard, but it is organized to be human readable.

If there were perhaps a file specifically intended for metadata, and on top of that, the metadata were structured to a specific standard, then this data would be much more easily found on a searchable repository environment. The folder and file names are also very weak in their descriptions. For example, the folder is titled ‘Clevelandportraitofaman.zip’ when the Cleveland Museum of Art has severalpieces entitled *Portrait of a man*, thus potentially confusing this data with the data for those other pieces. The images all also were named via the generic ‘IMG\_23…’ file naming convention used by the camera that took the pictures. Also, if more information were included about how to stitch the images together to create the final model or about how the final model(s) were created, it may help a user unfamiliar with the process understand the purpose of every image clearer.

Although there is no evidence that the data lead to an official publication, the model itself is viewable as a fully rendered 3D model on sketchfab. The implication is that this model was available on the Cleveland Museum of Art’s website when the foundation housed the actual piece, however now that the piece is now on view at the MoMA, the page no longer exists. This theory is given more traction as the Cleveland Museum of Art website does include several 3D models of their current pieces, while the MoMA does not. The model itself can no longer be found online besides sketchfab, but its original intention was to be used to represent the actual piece for the Cleveland Museum of Art’s digital collection.

Repository Profile - iDAI.world / ARACHNE

([Link](https://arachne.dainst.org/))

Of those available from re3data.org, iDAI.world is a digital repository whose collection objectives most closely match the data analyzed in the previous assignment. The repository would be used to store a collection of images and 3D modeling files of ‘Portrait Bust of a Man’, a bust from the Augustan period of the Roman Empire which is currently stored at the Cleveland Museum of Art.

ARACHNE is the German Archaeological Institute’s (DAI) central object database and is part of the iDAI.world project. In the project’s own words, “Arachne is an internet research tool for the Archaeologies and Classical Studies. It offers a means to access objects and their states and to search efficiently in hundreds of thousands of datasets. Arachne feeds from analogue and digital sources.” iDAI.objects / ARACHNE was chosen because the repository goes beyond collecting images of cultural heritage artifacts, but builds stories and projects around the originating geographical location. On top of this, each object is given a thorough treatment beyond a single image, but includes a plethora of images from various angles and lighting situations, just as with the featured data.

Though the images would be stored in iDAI.objects/ARACHNE, they would then be compounded with the modeling data and all of the data in total would then be displayed in a different subset of the iDAI.world project. An example of this would be the iDAI.field project, which displays photos and other data in a geographic context, allowing users to browse a globe to find artifacts based on where they were found. In the case of our data, it would be placed somewhere in Italy.

Due to the nature of the refined scope of the database, only datasets provided by established institutions are considered for acceptance. The concepts and goals behind iDAI.world are of open source availability of archaeology and art history while ensuring that each object model has a common set of attributes and can be thoroughly linked. The repository has been around since 1995 and has adapted and grown considerably since then, with plans to remain fluid and ready to adopt technologies as they premier in the future. The database was first created as a solution made in FileMaker but has since been updated continually since then, with the most recent version, ARACHNE 4, premiering in 2016. In 2001, the University of Cologne created a chair of Digital Humanities, the students of which used ARACHNE as a test environment for their own collections. Since that time, the database has been growing with only a few additions listed here: 40,000 sculpture negatives from Rome in 2003, the complete catalog of sculptures in the antiquities collection of Berlin State Museums including 3,500 plaster casts in 2013, and 2,300 printed engravings which were ran through a TEI viewer were added in 2013 as well. In total, the database now includes 2.9 million images, 600 3D models, and 23,226 books. In addition to the massive collection, context for each piece is given through the metadata of approximately 975,000 location references, 160,000 pieces of dating information, and 275,000 bibliographic references. The collection has subdivided its collection into the following categories: buildings, building components, individual objects, scenes, images, types, collections, topographies, receptions, reproductions, individual motifs, multipart monuments, inscriptions, books, and book pages.

Creating a login gives the user the ability to generate their own unique collections and displays the images at a higher resolution. All of the data contained within the ARACHNE database are available as open source within iDAI.world, the larger software architecture of the DAI. The data are available through several different modules which each provide different methods of displaying or connecting the data making up the collection. For example, the module *iDAI.field* allows the user to utilize a geographic UI to search for objects based on where in the world they were excavated or otherwise found by modern society. The database also supports Semantic Web, with every object assigned URIs. Because the photos are available via the creative commons license, the database provides the user with the contact information for the actual owners of the images for the purposes of ordering photos as publication templates. This includes the Research Archive for Ancient Sculpture and the various offices of the German Archaeological Institute itself. Any photos owned by anyone else are not available for order from ARACHNE. Image metadata are available, however, they are not formatted to any particular standard, but are simply listed under the headings ‘Technical Metadata’ and ‘Other Metadata’.

Files are not actually available for download through iDAI.objects, however, when viewing images, a dissemination information package is provided in the form of a linked HTML page listing metadata regarding the image. The metadata is not to any one standard, but includes information such as *photographer, image carrier, film number, negative number, scan number, file name, project name, file format, file size,* as well as *created, and modified* dates.

Additional Information

In MLA format, this data can be cited the following way:

Pett, D. 3D Data for Recreating a Bronze Portrait Figure from Cleveland Museum

of Art. 1, figshare, 23 July 2017, doi:10.6084/m9.figshare.5234890.v1.

In addition, the data can be found using a unique DOI:

<https://doi.org/10.6084/m9.figshare.5234890.v1>

In regards to long-term data preservation, the images (both JPEG and PNG) should have no problem remaining relevant and available to access as long as we have a method for accessing image files. The only maintenance required involves upkeep on the files themselves to avoid corruption. The modeling files are a different story. The .OBJ file is openable by any modeling software, so as long as modeling programs continue to use that filetype, it will remain usable. The other two are only openable using proprietary software. This means that if that company were to go under, stop using that file type, or stop producing new software that can read that file type, the file becomes unusable. This is why the data includes multiple ways to access the model.

All of the data is licensed under creative commons and is available for free use by anyone.

There is only one identifiable person in the data and that is the Roman soldier depicted in the bust. The identity of the man is a complete mystery to history, and even if he were well known, he has been dead for approximately 2000 years. Also the creator of the data is easily found as well, so it does not seem any attempts were made to anonymize or adjust the data based on privacy concerns.

REFERENCES

*Arachne*. arachne.dainst.org.

*Bronze portrait of a man | Roman | Early Imperial, Julio-Claudian | The Metropolitan Museum of Art*. (n.d.). The Metropolitan Museum of Art. <https://www.metmuseum.org/art/collection/search/248892>

*Daniel Pett, museologi.st*. (n.d.). Museologi.st. <https://museologi.st/>

figshare. (2023, May 31). *3D data for recreating a bronze portrait figure from Cleveland Museum of Art*. Figshare. <https://figshare.com/articles/dataset/3D_data_for_recreating_a_bronze_portrait_figure_from_Cleveland_Museum_of_Art/5234890>

*iDAI*. idai.world.

*Search the collection | Cleveland Museum of Art*. (n.d.). <https://www.clevelandart.org/art/collection/search?is_3d=1>

Sketchfab. (n.d.). *Portrait of a man - Download Free 3D model by danielpett*. <https://sketchfab.com/3d-models/portrait-of-a-man-9c148f09ec5e4b95835ef34e15f2acc0>

Xometry, Team. “All About OBJ File Format for 3D Printing.” Xometry, 5 Feb. 2024, www.xometry.com/resources/3d-printing/obj-file-format/#:~:text=The%20OBJ%20file%20format%20is,tools%20like%20SolidWorks%20and%20FreeCAD.