A Catch 22 of 3D Data Sustainability: Lessons in 3D Archaeological Data Management & Accessibility

Heather Richards-Rissetto  
*University of Nebraska-Lincoln*, richards-rissetto@unl.edu

Jennifer von Schwerin  
*German Archaeological Institute*, schwerin@dainst.de

Follow this and additional works at: [http://digitalcommons.unl.edu/anthropologyfacpub](http://digitalcommons.unl.edu/anthropologyfacpub)

Part of the Archaeological Anthropology Commons, Digital Humanities Commons, Geographic Information Sciences Commons, Remote Sensing Commons, and the Social and Cultural Anthropology Commons
Introduction

3D technologies are revolutionizing archaeological practice. Airborne laser scanning now allows archaeologists to map inaccessible places covered in lush forest canopy or off the beaten path (Brewer et al., 2017; Chase et al., 2012, 2014; Fernandez-Diaz et al., 2014; Opitz and Cowley, 2013; Prufer et al., 2015; von Schwerin et al., 2016a, 2016b). Archaeologists use terrestrial laser scanning to acquire 3D point clouds and close-range photogrammetry to convert overlapping 2D photographs into 3D point clouds—x, y, z data that we transform into 3D digital models for research, education, and outreach purposes (e.g., Remondino et al., 2009; Pattee et al., 2015; von Schwerin et al., 2016a, 2016b). Archaeologists use terrestrial laser scanning to acquire 3D point clouds and close-range photogrammetry to convert overlapping 2D photographs into 3D point clouds—x, y, z data that we transform into 3D digital models for research, education, and outreach purposes (e.g., Remondino et al., 2009; Pattee et al., 2015). While we can now collect an inordinate amount of high resolution 3D data from the scale of objects to vast landscapes, we face two critical and interrelated challenges for sustaining these data: management and accessibility.

While data sustainability falls not only under the purview of the digital, the seemingly infinite amount of digital data generation, rapid technology changes, non-standard data formats, application mashups, storage costs, and more make digital data sustainability a critical challenge (e.g., Clarke, 2015). 3D content in archaeology and cultural heritage is certainly not immune to this challenge. Extremely large datasets, evolving data standards or best practices, rapid developments in 3D technologies, and high costs make 3D data sustainability difficult (e.g., Koller et al., 2009; Simbulan, 2013).

The scope of 3D data sustainability is large; thus we narrow our discussion to focus primarily on two interrelated aspects of data sustainability that prove most challenging for 3D data in particular: data management and accessibility. Increasingly, archaeologists are publishing on the technical challenges of making 3D data accessible. Instead, we emphasize that current limitations to 3D data accessibility are not only technical, but rather the technical challenges are embedded within institutional practices that require best practices/standards in archaeological data management, in addition to “standardized” file formats. While this topic is multi-faceted, in this paper, we focus on four issues that were the most challenging (and therefore interesting) for the MayaArch3D project and that we believe will prove crucial for institutions to address in the coming years: data sensitivity/security, web-based dissemination, conveying uncertainty, and data storage/reuse/peer review. These are significant current challenges to making 3D archaeological data sustainable.

Abstract

Archaeologists can now collect an inordinate amount of 3D data. But are these 3D data sustainable? Are they being managed to make them accessible? The MayaArch3D Project researched and addressed these questions by applying best practices to build four prototype tools to store, manage, visualize, and analyze multi-resolution, geo-referenced 3D models in a web-based environment. While the technical aspects of these tools have been published, this position paper addresses a catch 22 that we, as archaeologists, encounter in the field of 3D archaeology—one that formed the initial impetus for the MayaArch3D Project: that is, while the quantity of 3D archaeological data is increasing, these data are not usually accessible. By researching and addressing 3D data integration and accessibility, we learned many lessons that group around four main issues: sensitivity/security, web-based dissemination, conveying uncertainty, and data storage/reuse/peer review. These are significant current challenges to making 3D archaeological data sustainable.

A catch 22 of 3D data sustainability: Lessons in 3D archaeological data management & accessibility

Heather Richards-Rissetto1 and Jennifer von Schwerin2

1 University of Nebraska-Lincoln, United States; richards-rissetto@unl.edu
2 German Archaeological Institute, Germany; schwerin@dainst.de

1 University of Nebraska-Lincoln, United States; richards-rissetto@unl.edu
2 German Archaeological Institute, Germany; schwerin@dainst.de
While efforts for 3D archaeological data sustainability are growing in the U.S. (e.g., [www.cyark.org/]), European institutions and organizations particularly have been working in the area of 3D data management toward establishing best practices and developing infrastructure for 3D data management in archaeology and cultural heritage for the past decade, (Remondino and Campana, 2014; Cignoni and Scopigno, 2008; Fresa et al., 2015; Lyons, 2016; Stylianidis and Remondino, 2016; Taylor and Gibson, 2016; Vecchio et al., 2015; von Schwerin et al., 2009, 2011; [www.dc-net.org/getFile.php?id=467]; [www.dc-net.org/getFile.php?id=450]). Efforts include CARARE ([www.carare.eu/]) and 3D-COFORM ([www.3d-coform.eu/]), which built tools, infrastructure, and workflows, and were essential in establishing a community of experts. Europeana ([www.europeana.eu/portal/]) integrates workflows and software developed through CARARE to promote re-use of cultural heritage data. The 3D-Icons Project ([http://3dicons-project.eu/]) built on these earlier efforts, focusing on the documentation and distribution in 3D data for UNESCO World Heritage Sites (Coms et al., 2015).

Current projects including VCC-3D ([www.vcc-3d.com/]) and ARIADNE ([www.ariadne-infrastructure.eu/]) concentrate on providing centralized expertise and centralized infrastructure to the community, whereas the INCEPTION project ([www.inception-project.eu/]) continues to innovate in the development and implementation of new technologies, and focuses on 3D semantic modeling. The Archaeological Data Service (ADS), located in the UK, provides a guide to good practice for archaeological data that includes close-range photogrammetry, laser scanning, and virtual reality with information on archiving these data ([http://guides.archaeologydataservice.ac.uk/g2gp/Main]). Applying these good practices, ADS recently developed a web-based 3D Viewer for the management and analysis of archaeological 3D data. The ADS 3D Viewer integrates the 3D Heritage Online Presenter (3DHOP), a software package for the web-based visualization of 3D geometries, with the infrastructure of the Archaeology Data Service (ADS) repository (Galeazzi et al., 2016; Potenziani et al., 2015) to make individual 3D models with associated archaeological data more accessible.

While the development of standards and best practices for 3D content is an essential component of data management, in this paper we address data sustainability with an end goal of reuse. Our key point is that to achieve 3D data sustainability, we not only need to guard against obsolete file types and disappearing data, but more work is required to make the data reusable. We understand that best practices and standards underlie data reuse, but continual management of the data is also required—this will allow for the adjustment of data to be used for multiple purposes, and as standards and best practices change. If we facilitate 3D data reuse for multiple purposes, this will lead to greater data sustainability because institutions have an impetus to sustain data that are being used and reused.

To make data reusable, it must be accessible. The common definition of data accessibility refers to a user’s ability to access or retrieve data stored within a database or other repository. But what does it actually mean to be able to access data? Does this mean users can view data online, or must they be able to retrieve, move, and manipulate data? We contend that a key component to making data truly accessible is providing the ability to do something with these data, that is, users must be able to reuse them in some way to further research. In this regard, the concept of linked data, i.e., interrelated datasets on the Web, which necessitates “standard [data] formats and reachable and manageable Semantic Web tools” provides a framework for “defining” data accessibility ([www.w3.org/standards/semanticweb/data]). Open Context—a web-based open access publishing platform—uses open linked data to provide data management as well as publishing for archaeological data that facilitates not only data access but data reuse ([https://opencontext.org/about/]; Arbbcule et al., 2014; Buccellati and Kansa, 2016). However, currently they do not offer a platform for visualizing and querying 3D data.

The task of making 3D archaeological data reusable is difficult not only because of limited data standards, heavy datasets, and rapidly changing technologies, but also because as archaeologists we must seriously consider to whom we can “ethically” make these data accessible (Vitelli and Colwell-Chanthaphonh, 2006). Furthermore, the difference between 3D surveying and 3D modeling, which underscores the difference between data acquisition vs. data creation, adds a layer of complexity, fuzziness, and interminability to the task.

1. The problem of “Infinite” 3D data

3D surveying is data collection. It is the raw x, y, z data we collect, for example with airborne LIDAR, terrestrial laser scanning, photogrammetry as well as theodolites. Data resolution ranges from microns to meters depending on technology and acquisition choices; the higher the resolution, the more 3D points, and thus the greater the storage requirements. To sustain these data, they must be stored, and data storage is expensive. Additionally, questions of data format arise—many data are collected in proprietary formats. What happens if we store only the proprietary format and the software to visualize and manipulate them disappears? Are inaccessible data that are stored on hard drives, servers, or the cloud truly sustained? No. This is why data management—not only data storage—is crucial to data sustainability.

To further complicate this problem, much 3D archaeological data comes from 3D modeling. 3D modeling is what we do with these data—in other words it is another layer of data creation; for example, we convert point clouds into polygonal meshes requiring decisions, interpretations, and transformations of the raw data. Along these lines, we reference two model types: (1) base models derived from laser scanning, photogrammetry, etc., that is, 3D technologies that collect data on extant archaeological features such as standing architecture, and (2) 3D reconstructions that hypothesize how a building or object might have looked in the past based on integrating sources ranging from architectural plans to laser scans to excavation data.

With both types of modeling, interpretive decisions are made and these decisions (paradata) are critical to the 3D models because they explain the choices modelers have made, for example, as they decimate meshes of base models for online use or add a third story to a building (Bentkowska-Kafel et al., 2016; Lyons, 2016; von Schwerin et al., 2011). The take-away points are three-fold. First, a 3D model is an interpretation and several reconstructions might exist for an individual building—which reconstruction do we keep? Do we keep all of them? Second, modeling involves choices (paradata) based on the input of data sources. How do we sustain the original data sources and paradata as well as the 3D data? Third, software is continuously updated and unfortunately newer versions are not always compatible with previous versions; thus it is always necessary to record the software version used to create and visualize 3D data. Which version of a 3D model do we preserve? Do we preserve all of them?

To provide an illustration, a quick search for 3D file formats on Wikipedia ([https://en.wikipedia.org/wiki/List_of_file_formats]) returns over seventy-four file formats for 3D models (graphic), and with the growing popularity of virtual and augmented reality this number is increasing. Among these many file formats, a few serve as “standards” for 3D archaeological data; however, these file types have pros and cons based on storage, visualization, and analytical needs. For example, the STL (STereoLithography) format is often used for 3D printing because it stores closed, solid (“watertight”) models that can be printed in slices, but a major drawback is that STL files have large sizes and do not store textures or other material properties. Moreover, in archaeology we often do not have “watertight” models—particularly for 3D landscape data. And, what about textures or other material properties? Is it enough to only store geometry when materials are critical for archaeological analysis? Thus, while STL might be an ideal format for 3D printing, it only...
serves a partial solution to the broad needs of 3D archaeological data sustainability and accessibility. Typically used for terrestrial laser scanning data is PLY—an ASCII polygon or binary format with no compression that stores both geometry and textures. However, for archaeology additional material properties are often critical, and in this case the Wavefront OBJ format, a simple data format that stores 3D geometry with texture coordinates along with a Material File(s) (.MTL) to define specific material properties such as shininess, is preferable. The common denominator among these three files types, STL, PLY, and OBJ, is that they can all be stored as ASCII and this is critical for sustainability as it is simply text that can be read without special software.

In the end, the number of ways we can transform our data is almost infinitesimal—so what do we preserve? How do we sustain the infinitesimal? To discuss some of these issues in more detail, we use as an example our experiences as members of the MayaArch3D Project—a project that developed a 3DWebGIS to bring together GIS and 3D for studies of ancient Maya architecture and landscapes (www.mayaarch3d.org).

2. The MayaArch3D project

The MayaArch3D project developed an open source, 3D WebGIS for the documentation, visualization, and analysis of complex archaeological sites (Reindel, 2014, von Schwerin, 2013, von Schwerin et al., 2016a, 2016b; Auer et al., 2014). This system has four prototype tools that can store, manage, query, visualize, and analyze 3D data of different formats and resolutions. Project partners were the Commission for the Archaeology of Non-European Cultures (KAAK) of the German Archaeological Institute (DAI), Department of Geoinformatics at the University of Heidelberg, Department of Anthropology and Center for Research in the Digital Humanities (CDRH) at the University of Nebraska-Lincoln, the Honduran Institute of Anthropology and History (IHAH), and the 3D Optical Metrology Unit of the Bruno Kessler Foundation (FBK). The technical aspects of this system have already been published (Auer et al., 2014; Billen et al., 2013; Loos et al., 2013, Lyons, 2016, Reindel et al., 2016, von Schwerin et al., 2015, von Schwerin et al., 2016a, 2016b). In this section we give a brief overview of the system tools; however, because technology so rapidly changes, we focus on data management issues underlying 3D archaeological data accessibility that ultimately impact sustainability.

The case study is the archaeological site of Copan, which today is a UNESCO World Heritage site located in Honduras (Fig. 1). The site’s occupation dates back to at least 1800 BCE (Hall and Viel, 2004) and from the fifth to ninth centuries CE. Copan was a prominent Maya kingdom (Fash, 2001). Copan is an ideal site for development and testing of 3DWebGIS because it has a long history of excavation and research providing large and diverse datasets from many researchers. In addition to archival and published data, the MayaArch3D Project has also collected close-range photogrammetric data, terrestrial laser scans, airborne lidar, and created CAD and Geographic Information Systems (GIS) data including, for example, 3ds, .obj, collada, shapefiles, and Digital Elevation Models (Remondino et al., 2009; Reindel et al., 2016; Richards-Rissetto, 2010; Richards-Rissetto, 2013; Richards-Rissetto and Landau, 2014; Richards-Rissetto, 2017; von Schwerin et al., 2013, 2016a, 2016b). We have used these data for the development of four prototype web-based tools: 2D Geobrowser, 3D Single Object Viewer, 3D Scene Viewer, and Virtual Panoramic Tour. We emphasize that these tools are prototypes, and that during tool development, we learned several lessons in managing 3D archaeological data with the intent to make it accessible. We believe these lessons learned can assist others in more effective planning for sustainable data management as they design archaeological projects involving 3D data (see also von Schwerin et al., 2016a, 2016b).

3. MayaArch3D prototype tools

3.1. 2D Geobrowser

The 2D Geobrowser of the MayaArch3D 3D WebGIS serves as a portal for georeferenced 2D and 3D geometry stored in a PostgreSQL database with a PostGIS extension. These spatial data are linked to archaeological and metadata in a FileMaker Pro database housed at the DAI and photographs from ARACHNE—a free object database of the DAI and the Institute of Classical Archaeology in Cologne (www.ariadne-infrastructure.eu/Portal/Related-services/DAI-online-services). For more information see the User Guide to the FM Pro Database: http://www.mayaarch3d.org/wordpress/wp-content/uploads/2015/06/User-Handbook-for-the-MayaArch3D-DAI-field.pdf. With the 2D Geobrowser users can navigate an interactive map (based on geomajas (www.geomajas.org) an open source collection of free and open source GIS libraries, tools and APIs for web mapping) to access coordinates and basic descriptive information such as site name and project director for 6554 archaeological sites in Mesoamerica (Fig. 2) (Loos, 2013; Reindel et al., 2016). We received 6000 of these GPS points courtesy of the Electronic Atlas of Ancient Maya Sites (http://mayaisms.gov.org) and then edited these data to eliminate doubles, fix misspellings, or improve coordinate accuracy as well as added 500 more sites. Additional descriptive data are available for sites in Honduras (Fecher, 2015). For Copan, data derived from 3D airborne lidar are integrated in the 2D viewer (von Schwerin, 2016a, 2016b). For more information see the User Guide to the geobrowser: http://www.mayaarch3d.org/wordpress/wp-content/uploads/2015/07/User_Guide_Geobrowser_eng.pdf.

3.2. 3D single object viewer

The 3D single object viewer allows users to interact with a collection of individual 3D models derived from multiple sources, e.g. photogrammetry, laser scanning, CAD (http://www.mayaarch3d.org/language/en/research/tools-in-development/3d-object-viewer/?lang=en). 3D models are stored and delivered as Three.js-JSON-Scene format. The viewer uses the open source GIScene.js library, which builds on top of three.js and adds geometric concepts (e.g. layers, projections, coordinates). Internally it uses WebGL—a JavaScript API for rendering interactive 3D computer graphics in browsers without the use of plugins (Auer et al., 2014). WebGL is an open web standard for online visualization. 3D viewer tools include rotation, zoom in/out, measurement, lighting, screen capture, and the ability to query information about 3D models or their segmented parts from a linked database (Fig. 3) (Lyons, 2016; von Schwerin, 2016a, 2016b).

3.3. 3D scene viewer

The 3D scene viewer contains geo-referenced 3D models created from 2D footprints (shapefiles) that are situated within a terrain model derived from airborne lidar data (Richards-Rissetto, 2013; von Schwerin, 2016a). The simulation is ca. 800 CE during the reign of Copan’s final dynastic ruler, and comprises over 3000 models that serve as visual storage “containers” (Fig. 4). These 3D models are served via a PostgreSQL database and are linked to the FileMaker Pro database allowing users to click on them to access archaeological data and metadata (von Schwerin, 2016a, 2016b). To reuse the developed code base, the 3D Scene Viewer also uses GIScene.js.

3.4. Virtual panoramic tour

To promote public outreach and support the Honduran Institute of Anthropology and History (IHAH) in encouraging tourism to Copan, the project produced a Virtual Tour of the Copan Archaeological Park (Fig. 5). This tour was created by linking twenty-four panoramic images captured with a 360° spherical camera system and GNSS GPS to georeference the images. The tour includes text, images, sound clips, and interactive 3D models. Users can tour Copan by clicking on hot spots on a map or following arrows through the virtual environment. The tour is available at http://3dmap.fbk.eu/repository/files/vt_copan/index-EN.html.
All four tools draw upon archaeological and 3D attribute data stored in the project archaeological database that can be searched via query tools. For reasons of inherited project data, and differences between data formats needed to connect to 3D segmented models and those needed for archaeological data analysis, it was decided to keep the FM Pro database as the foundational archaeological data management tool and link it via a parallel system to query builders in the 3DWebGIS tools (see von Schwerin, 2016a, 2016b for a more detailed discussion of the reasons for this decision). Because the FM Pro database is in the process of being mapped to the CIDOC-Conceptual Reference Model (common and extensible semantic framework for cultural heritage information), open linked data will be possible in the future (http://www.cidoc-crm.org/) (e.g. Bruseker and Carboni, 2014).

Based on our experiences in the design and development of these four prototype tools, four key issues have come to the forefront in regard to 3D archaeological data sustainability and accessibility. They are:

- Data Security/Sensitivity
- Web-based Dissemination
- Conveying Uncertainty
- Data Storage, Reuse, and Peer Review
4. Data security/sensitivity

Archaeological data are often sensitive. Data sustainability and accessibility require some degree of "open" access to data; however, many archaeological data must remain confidential for ethical, security or intellectual property reasons. The backbone of 3D WebGIS is georeferenced data; that is, data that have real-world coordinate information, but providing precise locational information of archaeological sites to the public is a concern for looting and vandalism. Moreover, some archaeological data are also culturally sensitive; for example, indigenous knowledge or sacred locations may not be appropriate for public consumption (e.g., Frank et al., 2015; Vella et al., 2015). Finally, it is not always possible to...
publish data that has been collected in a research database automatically to an online format. This may be because that data has not yet been published, or the web tool to visualize the data does not always have a direct link to the source information or to detailed enough source information. So often some data has to remain “project internal” while other data can be made available to other archaeological colleagues, or the public.

This conundrum creates a catch 22—we want/need to make data accessible, but to whom can we actually make sensitive data accessible? How do we decide who should have data access? And, after these decisions are made, how do we deal with data security? While each project has specific data sensitivity and security concerns that might require tailored solutions, the MayaArch3D Project has done four things to secure sensitive data. First, we have designed the 2D Geobrowser and 3D Scene Viewer with five levels of user access. The five user levels are: (1) public—access to view archaeological parks including linked data, no edit permissions; (2) external researchers—access to view all data (e.g. unguarded archaeological sites in addition to archaeological parks), no edit permissions; (3) project researchers- access to view all data, permission to add and edit attribute, GIS (geometry), and 3D data, others cannot view or edit project data until project director gives permission; (4) institutional- full access to all archaeological sites and linked data with permission to edit data via the web-based interface, and; (5) administration- full permissions for front-end and back-end established via the database. Currently, we have these five user-levels implemented—public, external researcher, internal/project researcher, MayaArch3D project member/IHAH, and administration.

Second, making use of these user-levels, the GIS data in the 2D Geobrowser are divided into two categories—(1) protected archaeological parks or sites with guards and (2) unprotected sites. Only protected archaeological parks are visible to the public and unprotected sites are available with password log-in. In this way, the data are protected and yet still accessible in a web-based environment to those with permission.

A third option that we discussed with the Honduran Institute of Anthropology and History (IHAH) is to implement zoom limits for archaeological sites in Honduras. In this way, the public could view relative site locations in relation to terrain and satellite data; however, they could only zoom into, for example, a twenty-five kilometer resolution. This approach would allow the public to access data and yet still protect precise coordinate information.

Fourth, to ensure security, data are stored on a secure server at the DAI headquarters and backed up on servers at the Computing Center of Cologne University (RRZK). Furthermore, the data are stored on the DAI–Cloud, which encompasses the constant and long-term expansion of a high end NAS storage area.

5. Web-based dissemination

3D data are heavy. Raw 3D data points acquired from laser scanning and photogrammetry afford high resolution data capture, but to visualize these raw data in real time in a web-based environment is difficult even with the best of broadband conditions. 3D models that capture finite details and curves contain more data but are often too big to visualize on the web, never mind perform scholarly analysis, and models must be decimated (number of polygons in mesh reduced) for online visualization. While we can use methods such as normal mapping to texture 3D models and make them look high-resolution, behind the scenes, the geometry is typically reduced to a level inappropriate for quantitative analysis or highlighting fine details for archaeological analysis. Additionally, while technologies are now allowing us to visualize single 3D objects in greater detail online (Verhoeven, 2016), it continues to be a challenge to visualize high-resolution 3D data across landscapes (e.g., Fanini et al., 2011; Forte, 2011; Frischer and Dakouri-Hild, 2008; Loos et al., 2013; von Schwerin et al., 2016a, 2016b).
This challenge creates a second Catch 22—researchers require high-resolution data for certain analyses and yet to make high resolution data accessible for real time interactivity and analysis online is difficult. The MayaArch3D Project researched solutions to this challenge in the development of the 3D Single Object Viewer and the 3D Scene Viewer. In regard to 3D Single Object Viewer, before development our colleagues in the Geoinformatics Group at the University of Heidelberg carried out testing on the pros and cons of different 3D libraries for delivering high resolution on the web. The open source library Three.js was selected because it is lightweight 3D library that runs directly in the browser, it handles shadows, rendering, vertex processing, shaders, etc., and importantly, OBJ and Collada—common 3D mesh formats for archaeology—can be converted into Three.js (Auer et al., 2014). While the Three.js format as an ASCII format is ideal for sustainability, it is not an efficient for usage. Therefore, for the 3D Single Object Viewer and 3D Scene Viewer, we selected the Three.js- JSON Format—an ASCII format promoting sustainability and yet unlike simple ASCII files, Three.js JSON also allows for efficient web usage because JSON is a lightweight data-interchange format.

Geoinformatics specialist, Michel Auer tested performance—loading, parsing, and rendering and system requirements—and memory consumption. Results indicate that large models cause long loading and parsing times, i.e., browser does not respond and large meshes require higher memory consumption; thus, depending on user hardware, the browser could crash. From these data, it was determined that we could comfortably visualize up to 1,000,000 triangles in the 3D Single Object Viewer.

But what about multiple 3D models within a landscape—how do we visualize and interact with these data on the web? The 3D Scene Viewer also uses Three.js and WebGL; however, the terrain data generated from airborne LiDAR needed to be converted from 3D point data to a mesh and decimated. However, to minimize decimation and maximize geometry accuracy, there is a Web 3D Service to deliver the terrain in tiles that load based on user viewpoint and a Geometry Service is used for hierarchically segmented single objects (Auer et al., 2014; Billen et al., 2013; Loos et al., 2013). It is critical to note that while on-screen view is triggered by user distance and field of view, for analytical purposes, such as visibility analysis, the entire terrain is taken into account. This solution allows higher resolution and large landscapes to not only be visualized and navigated in a web-based environment, but importantly data integrity is maintained to permit quantitative analysis, also online. In a similar vein, 3D models of buildings can be visualized as Levels of Detail (LOD), which involves decreasing or increasing the complexity of a 3D model representations (e.g., LOD 1=un-textured schematic blocks, LOD 2=refined building models with doorways, stairs, etc. and texture, etc.). These LODs allow users to experience vast landscapes and relationships among archaeological features, for example, at LOD 1 the 3D scene viewer visualizes over 3000 3D models in Copan’s terrain (Agugiaro et al., 2011).

6. Conveying uncertainty

Do we use hypothetical reconstructions to present a vision of the past that gives a sense of ancient life (realism) or do we visualize strictly what we can reconstruct from archaeological data (reality)? The answer often depends on how these data are to be used; thus, most archaeologists have come to realize that realism vs. reality sets up an unnecessary dichotomy, and instead we require approaches to convey the uncertainty of our 3D data and models.

While both 3D visualizations for public consumption and 3D digital scholarship require solutions to convey uncertainty about 3D data as well as resultant models and visualizations, 3D scholarly publications seeking peer-review have more stringent requirements (Kanter, 2000; Lyons, 2016; Richards-Rissetto, 2013). Currently, various visualization methods exist to convey modeling certainty such as combining reality-based with transparent reconstructions or using color coding, color coding + opacity, wireframes, and other combinations of these techniques (http://gabi-server.adsroot.itsc.umich.edu/gabiqoesdigital/gabii_unity_sandbox.html; www.aquaepatavinae.it/portal/?page_id=2174&lang=en; Apollonio, 2016; Danielova et al., 2016; Kensek et al., 2004); however, these focus on visual solutions and many scholars question how effective these visuals actually are in conveying uncertainty. But what can be done to increase clarity about uncertainty in 3D archaeological data? We contend that if the associated archaeological data (attributes) and other data sources (e.g., site plans) along with para data (modeling choices) are made accessible then uncertainty can be much more clearly conveyed.

When data about 3D models is made available to users, specifically via linked data, then the uncertainty of 3D reconstructions becomes more evident and peers are more easily able to review the quality of the data. Information about how 3D data are collected or generated creates greater clarity for peer-reviewers moving us towards an increased valuing of 3D scholarship. If para data is made accessible, users can “see” and explore the relationships between end-products, modeling choices, and source data. However, for these para data to be accessible requires a range of solutions from hyperlinked text documents or annotations to sophisticated database solutions. One example is the Gabii Goes Digital Project (http://gabi-server.adsroot.itsc.umich.edu/gabiqoesdigital/gabii_unity_sandbox), which employs text and reality-based modeling overlaid with a transparent hypothetical reconstruction to convey model certainty and also provides links to data in a database (Nebbia, 2014; Opitz et al., 2016).

In a similar vein, the 3D Single Object Viewer and 3D Scene Viewer of the MayaArch3D Project link georeferenced 3D models to a database. These models range from reality-based models (what we can “see” today) to 3D models generated from excavation data to hypothetical models based on survey footprints and comparative data. The system links to a database that users can query to obtain both associated archaeological information with specific features but also the metadata about the models themselves—how, when, by whom they were created, etc.

While uncertainty is not immediately visually apparent in the 3D Scene Viewer, users can click on the 3D models in the single object viewer or 3D scene viewer to access information and photos in the linked database. Given our research and development of these prototype tools, we advocate a Visual + Textual approach with visual illustrations of uncertainty via for example transparency, color-coding, etc. that are linked to textual descriptions stating this is what we “know” vs. “reconstructed” as well as active links to a database to enable greater data accessibility to promote ongoing research with the 3D data.

For example, in a follow-up project, MayaCityBuilder, we have brought some of the hybrid models into Unity—a 3D gaming engine—to begin to visualize the simulations that will be possible. You’ll notice in video 6 we have integrated a pop-up system triggered by distance that gives information on the data sources used in this reconstruction as well as short blurb about the process including the decisions made and data used in the modeling process (Day and Richards-Rissetto, 2016) (Fig. 6). (This reconstruction of Temple 18 at Copan was done by Mike Lyons as part of his MA thesis at the University of Bonn. See Lyons (2016)).

The Copan Virtual Panoramic Tour offers another option to convey uncertainty through a tour of the present-day landscape that offers reality-based photogrammetric or laser scanned 3D models in conjunction with pop-ups that provide additional information (http://3dgeom.fbk.eu/repository/files/vt_copan/index-EN.html). While we have not done so yet, we plan to juxtapose 360° views from the tour with 3D reconstructions of specific buildings and architectural complexes that can be found in the 3D Object or 3D Scene Viewers. We note that the virtual tour uses the KRPano viewer, which while a flexible high-performance viewer, our current version is based on flash technology (not supported on all platforms such as mobile devices), and therefore, a lesson learned is to use the HTML5
dissemination version for future deployment and accessibility. However, still the tour has made these 3D data accessible to a wider public, who can easily navigate the UNESCO site in a web-based environment.

7. Data storage, reuse, and peer review

Archaeological data repositories such as TDAR serve the long-term preservation of data (https://www.tdar.org/about/), others such as Open Context (https://opencontext.org/) focuses on fostering data reuse with a web-publishing platform (Kansa et al., 2014), and the Digital Index of North American Archaeology (DINAA) (http://jux.opencontext.org/archaeology-site-data/) is working to create interoperability models for archaeological site databases in the eastern U.S. (Wells et al., 2014) (see Clarke, 2015 for comprehensive discussion). While these projects and others are making great strides forward in archaeological data preservation and reuse, their main focus is not 3D data. In contrast, a key goal of the MayaArch3D Project was to research and use 3D data standards and best practices to promote 3D data reuse. However, given that Open Context is now mirroring its data with the German Archaeological Institute, we may be able to increase data reuse through establishing interoperability with Open Context.

To make the MayaArch3D System sustainable, we have begun to further standardize our datasets and place them in the e-Repositories of the German Archaeological Institute (DAI) (Reindel et al., 2016). The DAI is a recognized government repository for archaeological data in Germany and is the largest archaeological research institution in Germany with about 250 projects around the world. Many of these are dealing also with the problem of managing large amounts of data—including 3D data. The DAI maintains an important national repository of archaeological data and provides quality-assured research data and analysis tools. The data infrastructure is led by Prof. Dr. Reinhard Forstsch, Scientific Director of Information Technologies. The DAI-Cloud is a storage, virtual-server- and tape-archive-based data infrastructure for storing, processing, and long-term-archiving DAI’s data.

This exemplary solution to the problem of data management and access is a research data infrastructure called iDAI.world (http://resources.dfg.de/).1 iDAI.world is a modular research-data-environment with a data-layer, a standards-layer and an analysis-layer, described in greater depth on the DAI homepage (http://www.dainst.org/forschung/forschung-digital/daiwelt). The data layer contains repositories of data from its numerous archaeological projects, reports, publications, photo archives, 3D models, etc. The archaeological field data repository of the DAI (iDAI.field) was used to collect test data from Copan. Translation tables (in English, German and Spanish) were created for the iDAI.field repository in order to begin to optimize the accessibility of the data for an international scientific community.

Beyond the MayaArch3D Project, Reinhard Forstsch has overseen the design of the standards layer of iDAI.world to allow for the standardization of the data in its repositories — this contains a gazetteer of place names, vocabulary lists, definitions of temporal periods, etc. The place names and GPS points of Maya archaeological sites that the MayaArch3D project has prepared are now in the DAI gazetteer geoserver, and the project photos are in the object database, ARACHNE. Here these data receive digital object identifiers (DOIs) and our 3DWebGIS — which is an analytical system rather than only a storage or archiving system — can then access these data for visualization and analysis.

To ensure interoperability and usability of the data, we collaborated with the Foundation for Research and Technology– Hellas (FORTH) to map our database structure to CiDOC-CRM, an ISO standard for a semantic ontology to record cultural heritage data (http://www.cidoccrm.org/) (Bruseker and Carboni, 2015). CRM has been implemented at the British Museum and German Archaeological Institute, as well as various European Union projects (e.g., Guillen et al., 2015; Makela et al., 2016), and also the Text Database and Dictionary of Classic Mayan at University of Bonn (https://www.sub.uni-goettingen.de/en/projects-research/project-details/projekt/text-database-and-dictionary-of-classic-mayan/). Because the CRM ontology makes the data semantically intelligent, it expands the interoperability of the mapped data with other databases and so it will ensure accessibility of the data long into the future. Moreover, source code (developed by Michael Auer) is available on GitHub to promote on-going research and development. The GSbrowser library (a 3D WebGIS framework based on Three.js) is open source allowing other researchers to build on the existing code (https://github.com/GIScience/GSbrowser.js).

During the research and development of the four prototype web-based archaeological tools of the MayaArch3D project, issues of data sensitivity and security, web-based dissemination, conveying uncertainty, and data storage, reuse, and peer-review came to the forefront. While various options exist for dealing with these issues, 3D archaeological data brings particular challenges such as expensive storage requirements, rapidly changing software, and heavy datasets. While major goals in the field of 3D archaeology are to make 3D archaeological data sustainable and accessible, these terms are complex and thus researchers must explicitly define what they mean by them as the field quickly moves forward.

8. Conclusions and future directions

The rapidly increasing amount of 3D archaeological data from technologies such as laser scanning and photogrammetry creates a catch 22—more 3D data creates more challenges in how to make this data accessible. Issues of accessibility are multifaceted. While much focus is on the technical challenges of making 3D data accessible via the internet, at a foundational level, all data requires long-term archival (storage) strategies and the funds to maintain these storage facilities otherwise the data are not available. However, simply saving 3D data is not sufficient. We must continually manage the data to keep it in accessible formats for the public, teachers, and researchers for reuse. The challenges we have highlighted in this paper should be considered and addressed before a 3D data project begins, in order ensure that the 3D data collected is published in a peer-reviewed format, and is sustainable and accessible for future generations, particularly for data reuse.

Archaeologists must work toward the reuse of 3D data because it affords:

• data downloads for research, artistic, or educational applications (each of which has different needs with regard to formats and associated metadata)

1. http://resources.dfg.de/detail/RI_00402_de.html; http://resources.dfg.de/detail/RI_00401_de.html; http://resources.dfg.de/detail/RI_00401_de.html; http://resources.dfg.de/detail/RI_00403_de.html; http://resources.dfg.de/detail/RI_00405_de.html; http://resources.dfg.de/detail/RI_00403_de.html; http://resources.dfg.de/detail/RI_00405_de.html; http://resources.dfg.de/detail/RI_00405_de.html; http://resources.dfg.de/detail/RI_00405_de.html; http://resources.dfg.de/detail/RI_00405_de.html
However, often reuse requires that archived data be reformatted or standardized, and it requires meta- and paradata that allow users to understand the data, particularly with regard to levels of uncertainty. Thus data management plays a central role in the sustainability of 3D data, particularly as new possibilities for reuse of 3D data continue to be invented.

The MayaArch3D Project developed four prototype tools to foster collaborative research and data reuse. While the archaeological and 3D data are stored and hosted by the repositories of the DAI, much of the data also are accessible in the online tools of the MayaArch3D project for a wide range of uses. In the future, the open source infrastructure of the MayaArch3D system could be used to present and analyze 3D data from other Maya sites or for comparative studies of 3D models between sites.

The 3DWebGIS developed by the MayaArch3D project has the potential to be an online publishing venue for 3D models connected to geo-referenced archaeological data throughout the Maya region, and a tool for collaborative international research on, and management of, complex archaeological sites. Given the already established institutional connections between the DAI and Open Context, the MayaArch3D Project could serve as a prototype platform for the development and testing of an online publication platform for 3D archaeological data linked to meta- and paradata. As a web resource, it could help archaeologists structure and upload their data, run analyses against other data sets, all the while keeping their data, when necessary, invisible and proprietary. Then when they are ready, researchers could submit their data to peer review for publication in the system within a peer-reviewed, virtual research environment, for others to cite, use, and download if allowed. But to arrive at this point, this prototype system needs about 1.5 years of additional development on the user interface and security functions and of course, interest and support from the Maya archaeology community. In any case, for the 3D data that is linked, visible, and analyzable in the 3DWebGIS to be truly sustainable and accessible, ideally, the 3DWebGIS must become an online, peer-reviewed digital publication venue with an institution or journal that continually manages it. This brings us again to the point about security. While as archaeologists we have ethical responsibilities to manage our data and make it available to future generations, we cannot make sensitive data widely available, and we must ensure that any sensitive data are secure. As mentioned above, in the MayaArch3D project we have designed tailored solutions to deal with sensitive 3D georeferenced data (as well as other types of data). However, while additional 2D and attribute data can be added at any time to the existing 2D and 3D geometries, to allow for the upload and integration of NEW 3D data sets, as well as for the security of these and links to their sources, the system requires additional development.

As for future directions, with regard to management and reuse of the MayaArch3D data, we are moving forward with a new project that will use the prototypes tools and data from the MayaArch3D Project to carry out archaeological analyses on Copan. To start, we are beginning formal collaborations with other Copan archaeologists to integrate additional datasets (architectural, iconographic, textual, environmental and bioarchaeological) into the system in order to run new kinds of analyses with these combined datasets to better understand the social organization at the city. We are also considering to organize a summer school to train Honduran students to use the system.

In addition, we have begun the MayaCityBuilder Project (www.mayacitybuilder.org) to further promote data reuse. The long-term goal of this project is to allow users to create alternative architectural reconstructions of ancient Maya citiescapes within a web-based environment to foster collaborative analysis and discourse. To do this, the MayaCityBuilder has been reusing data from the MayaArch3D Project as well as employing procedural modeling (Richards-Rissetto and Plessing, 2015). We have developed workflows and user modules to quickly generate 3D buildings and vegetation from georeferenced data in the free version of Unity. Programming in Unity using C#, we have created four modules in Unity that allow users to bring in georeferenced and digitized building footprints (from any archaeological site) and automatically generate 3D buildings stored in an assets folder in Unity. 3D models can be brought into the assets folder offline (desktop) or online, for example, from 3D models stored on SketchFab—a free web platform for publishing 3D models (http://sketchfab.com). These 3D buildings can be generated in two ways: (1) using an already assigned name in the imported footprint file or (2) linking to a PostgreSQL database. Moreover, we have created modules to easily scale and rotate 3D models in relation to footprints (Fig. 7).

Finally, to promote further data reuse by providing examples of what is possible with 3D data, the MayaCityBuilder Project is using these same 3D data to test immersive technologies such as the Oculus Rift and gesture-based technologies such as Kinect and LEAP Motion for public consumption (http://mayacitybuilder.org/?page_id=295).

To close, we re-emphasize here that as archaeologists who want to achieve data sustainability for 3D content in cultural heritage, we must not only archive the data, but continually manage it with regards to standards, storage requirements, uncertainty, and security issues. And most importantly—we must actively promote data accessibility by preparing our data for—and perhaps even becoming involved with projects that engage in—data reuse.
Acknowledgments — We thank Jared Katz for organizing the Chacmool session and this special issue as well as the three anonymous reviewers for their invaluable comments. Special thanks go to Markus Reindel of the Commission for the Archaeology of Non-European Cultures at the German Archaeological Institute who was Coordinator of the MayaArch3D Collaborative Project, as well as to Professor Alexander Zipf and his GIScience Group at the University of Heidelberg, particularly Michael Auer, Lukas Loos, and Nicolas Billen who were responsible for the technical design and implementation of the MayaArch3D tools. Mike Lyons, Franziska Fecher, and Laura Stelson were staff members on the archaeological sub-project at the DAI. Thanks also to Reinhard Förtsch, of the IT Department of the German Archaeological Institute, for collaboration and support. We thank Belen Jimenez and Fabio Remondino of the 3D Optical Metrology Unit of the Bruno Kessler Foundation for the terrestrial laser scanned and photogrammetric models and the implementation of the Virtual Tour. The MayaArch3D Project was funded by the E-Humanities program of the German Ministry for Education and Research. The MayaCityBuilder Project thanks University of Nebraska-Lincoln students—Zachary Day, Rachel Plessing, Stephanie Sterling, and Shane Bolan. MayaCityBuilder received pilot funding from the Layman Award, University of Nebraska-Lincoln and undergraduate research funding from UCARE—Undergraduate Creative Activity and Research to support student participation on the project. Final and most important thanks go to the Honduran Institute of Anthropology and History for permission to collect data in Copan.

Appendix A. Supplementary material — Supplementary material associated with this article is as follows (all are mp4 files):

1. Video of 3D Object Viewer illustrating functionality, MayaArch3D (120 MB)
2. Video of 3D Scene Viewer, MayaArch3D (62 MB)
3. Video of Virtual Tour of Copan’s Main civic-ceremonial group, MayaArch3D (144 MB)
4. Video of working virtual reconstruction of Copan, MayaCityBuilder (102 MB)

These are archived with this document, and can be accessed from the html webpage addressed on the first (cover) page of this version.

References


