IJDC | *General Article*

Volumetric Video: Preservation and Curation Challenges of an Emerging Medium

Zack Lischer-Katz
University of Arizona

Bryan Carter
University of Arizona

Rashida Braggs
Williams College

Abstract

Volumetric video is an emerging media format that uses multiple cameras to record live-action subjects and produce three-dimensional, time-based digital media. The resulting digital objects encode visual and spatial information, colour, textures, and sound in a format that allows for users to view the subject from any angle and use the assets in video games, virtual reality, augmented reality, or films. The technology has been pioneered by Hollywood production companies but is now being experimented with by digital humanities scholars. As it becomes more popular, information institutions, particularly academic libraries and others that support researchers, will likely need to support this new format throughout its lifecycle, which may draw on research data management, digital preservation, and repository services. This article introduces volumetric video capture, discusses some of its current applications outside of the commercial film industry, and outlines the curation and preservation challenges that this new media format presents. The paper compares two different production workflows that result in different output qualities: professional and prosumer studio-based workflows. The analysis explores the digital curation challenges that volumetric video raises within these workflows, with considerations for selection and appraisal criteria, file format sustainability, metadata requirements, legal/ethical considerations, and directions forward for future research in digital curation.

Submitted 17 May 2024 ~ *Revision received* 11 March 2025 ~ *Accepted* 12 March 2025

Correspondence should be addressed to Zack Lischer-Katz. Email: zlkatz@arizona.edu

The *International Journal of Digital Curation* is an international journal committed to scholarly excellence and dedicated to the advancement of digital curation across a wide range of sectors. The IJDC is published by the University of Edinburgh on behalf of the Digital Curation Centre. ISSN: 1746-8256. URL: http://www.ijdc.net/

Copyright rests with the authors. This work is released under a Creative Commons Attribution License, version 4.0. For details please see https://creativecommons.org/licenses/by/4.0/

Introduction

Volumetric video capture is an emerging media technology (and proto-format[1]) that records live-action subjects with an array of cameras (from 2 to 40 and beyond) to produce three-dimensional, time-based digital media. The resulting digital objects encode spatial information, textures, and sound in a format that allows for users to view the subject from any angle, and use these assets in video games, virtual reality (VR), augmented reality (AR), or in films. Volumetric video is essentially a series of video frames in which each frame is a 3D model. This can be encoded as a series of 3D mesh files (e.g., a stream of OBJ files and associated textures) or as 2D video files (e.g., MP4) with an associated file used to reassemble its geometry. The application of this technology to narrative storytelling has been pioneered by Hollywood studios and used in some cases to record the stories of important historical figures (e.g., the testimonies of Holocaust survivors; see Schreer et al., 2022). Academics working in the digital humanities (DH) are now experimenting with ways to use volumetric video as a capture and streaming tool for presenting new pedagogical experiences to students and for research purposes. As volumetric video becomes more popular, information institutions, particularly the academic libraries that support DH researchers, will likely need to support this new format throughout its lifecycle, requiring research data management, digital preservation, and repository services.

Volumetric video is still highly experimental and is typically limited to use in specialised labs at large research universities. To "democratize" the technology and ensure that the potential benefits of this new technology can be realized by DH scholars more broadly (Carter et al., 2021), a group of researchers from University of Arizona (Tucson, Arizona) and Williams College (Williamstown, Massachusetts), in collaboration with technical innovators from the world-renowned volumetric capture studio VoluCap, embarked on a project to explore the challenges and potential benefits of volumetric video capture for BIPOC (Black, Indigenous, People of Colour) storytelling. The project was funded in part by a Digital Humanities Advancement Grant from the National Endowment for the Humanities, titled "Preserving BIPOC Expatriates' Memories During Wartime and Beyond: Building a Volumetric Archiving Platform for Immersive Storytelling and Historical Preservation." The team travelled to Berlin/Potsdam, Germany, in June 2023 to visit VoluCap Studios and record several volumetric capture videos. Dr. Bryan Carter is the director of the Center for Digital Humanities (CDH), which houses a prosumer-level volumetric capture studio. Comparing the workflows at CDH with observations from the visit to the VoluCap Studios allowed the project team to better understand the challenges and benefits of volumetric capture at different scales and levels of quality. Future digital libraries may be populated with volumetric videos of historical (or everyday) people, or reenactments of famous events for scholarship and pedagogy. This paper presents on the creation challenges and pedagogical benefits of volumetric video, as well as the preservation and curation challenges of volumetric video assets.

What is Volumetric Video?

Volumetric video (VV) consists of a fixed stream of 3D models, displayed at a sufficiently rapid frame rate or refresh rate to produce the impression of motion to a human observer, similar to the effect of film or video recordings. Typically, sound is recorded using synchronised sound

[1] Virtual reality, augmented reality, volumetric video, and other new media technologies rely on a range of digital file formats, but they can also be studied as media formats in themselves, as these technologies become stabilised and standardised over time. They are currently inchoate, and perhaps the idea of "format" is a bit premature to apply. Media theorist Jonathan Sterne (2012) suggests that format "denotes a whole range of decisions that affect the look, feel, experience, and workings of a medium. It also names a set of rules according to which a technology can operate" (p. 7). Format theory, as articulated in current media studies research, posits that formats "frame and configure media in fundamental ways while also linking different domains of media production, distribution, and reception" (Volmar, 2017, p. 9).

recording equipment, edited, and synchronised with the VV elements in post-production. Unlike film or video, however, wherein each frame is a flat, 2D image, in VV, in each "frame" is a 3D model. This enables viewers to perceive the depth of the objects in the scene or rotate the recorded scene in space and take up different viewing positions to view the scene from different perspectives. Some of the more popular presentation methods include displaying VV via augmented reality on smartphones or other devices, producing hologram-like effects, as well as incorporating VV into virtual environments, in order to incorporate realistic representations of real-world objects in virtual reality.

Volumetric video is being used in DH research and pedagogy, often for communicating historical topics through re-enactments and restaging of famous figures, or through oral history storytelling by contemporary people. Volumetric video has emerged at a time when virtual reality (VR), augmented reality (AR), and other technologies falling under the umbrella of "XR", or extended reality, are becoming increasingly affordable for scholarly use. At the same time, 3D capture techniques, such as photogrammetry, laser scanning, and structured light scanning, have become more widely available for producing high-resolution digital 3D models of artifacts, sites, and people. Investigating VV from a DH perspective entails both exploring the possibilities of VV for humanities representation and pedagogy, as well as interpreting how VV, integrated as it is into other media, such as feature-length Hollywood films or the AR apps on our smartphones, shapes our experiences as humans in the early twenty-first century.

Benefits of 3D and Immersive Media for Research and Pedagogy

3D imaging technologies have shown great promise for supporting research and pedagogy in a range of fields. Fields where the object of study has an important spatial component can particularly benefit from 3D digital imaging tools (Donalek et al., 2014). Archaeologists have been using 3D imaging technologies to capture data from excavations and other types of research sites (e.g., Fragkos et al., 2018; Hostettler et al., 2024, Karasik & Smilansky, 2008). Architects and historians are also using 3D tools to capture rich spatial data about cultural heritage sites for analytic and preservation purposes (e.g., Bozorgi & Lischer-Katz, 2020). Humanities scholars are also finding benefits to 3D capture for understanding written texts in new ways. For instance, Bill Endres (2019) has shown how 3D tools can provide new scholarly insights into "flat" medieval manuscripts, and Roberts Thompson and Williams (2016) used reflectance transformation imaging (RTI) to create 3D scans of a historic copper plate. Since the early 2000s, a range of projects in Europe have been exploring the potential uses of 3D imaging technologies for preserving cultural heritage artifacts and sites (e.g., the 4CH Project: Competence Centre for the Conservation of Cultural Heritage)[2] and developing tools for presenting 3D-based scholarship (e.g., Pure3D: an Infrastructure for the Publication and Preservation).[3] 3D data can be presented as 2D images on computer monitors, but increasingly affordable XR (extended reality) technologies such as virtual reality (VR) and augmented reality (AR) are making it possible to view 3D data with its spatial dimension fully represented, with important implications for visual understanding, such as pattern recognition (Donalek et al., 2014), and for immersive and engaging teaching purposes. Recent research suggests the benefits of immersive technologies for student learning in higher education, often supported by university libraries, including the use of VR within an introductory anthropology class (Lischer-Katz, Cook, & Boulden, 2018) and in undergraduate chemistry lab instruction (Qin, Cook, & Courtney, 2020). Spatial cognition has been identified as an essential, and often underdeveloped, dimension of student learning in engineering, mathematics, and other fields (Sorby et al., 2013; Buckley, Seery, & Canty, 2018), and the potential of VR and other XR technologies to support improvements to spatial skills for students have been suggested (e.g., Rafi

[2] https://www.4ch-project.eu/about-the-project/
[3] https://editions.pure3d.eu/

et al., 2005; Rafi, Samsudin, & Said, 2008). Research on the effects of XR's affordances of presence and immersion suggests enhancements to historical and cultural understanding and for increased student engagement in the classroom (Aguayo & Eames, 2023; Chang et al., 2019; Dodds, 2021; Meccawy, 2022; Sadler & Thrasher, 2023). Carter et al. (2021) identify the transformative power of volumetric video (VV) to improve student learning as volumetric technologies "connect the perception of comfortable proximity and 'presence' with increased attention and engagement" (p. 4845). As with other emerging spatial media, the high cost and complexity of the VV creation process have limited its application to big budget Hollywood film productions and advertising campaigns, or to a handful of innovative digital humanities labs.

Identifying Curation and Preservation Challenges of Volumetric Video Capture

Volumetric video can take several different forms (3D sequences, video files, etc.), but in many workflows, volumetric video is encoded in the form of a sequence of 3D models (textured, 3D meshes). As such, we can look to research on preserving 3D models for some guidance for how to think about the challenges of preserving volumetric video.

Interest in the use of 3D capture technologies in libraries emerged in the mid-2010s, in part due to the integration of maker spaces and other innovation labs into library spaces, as well as a decrease in the cost of the hardware and software needed to carry out 3D scanning and processing. Parallel developments in the availability and affordability of virtual reality (VR) technologies encouraged some researchers to explore 3D data holistically with a consideration of VR-related hardware and software. Over the last decade, several library projects have attempted to explore the preservation and curation challenges of 3D data formats, such as the CS3DP (Community Standards for 3D Data Preservation) Project,[4] and the LIB3DVR Project (Developing Library Strategy for 3D and Virtual Reality Collection Development and Reuse),[5] Building for Tomorrow,[6] and the Smithsonian 3D Digitization Project.[7]

These projects have found that 3D data formats have particular characteristics that distinguish them from other types of research data, namely the variety of different creation techniques used, the mixtures of different formats, and the scale of data collection, processing, and storage needs. One of the outputs of the CS3DP project was the book *3D Data Creation to Curation: Community Standards for 3D Data Preservation*, which notes "the features that necessitate treating preservation of 3D data differently emerge from the co-occurrence of multiple data types [...] and multiple intents for end use (e.g., use to create fixed 2D visualizations, use to produce physical objects, use to allow interactive inspection, or use to produce a virtual experience)" (Moore, Rountrey, & Scates Kettler, 2022, p. 298). Similarly, the LIB3DVR project found that while 3D data formats share many of the same challenges as other complex data types (such as digital images, or video recordings), including large files sizes, proprietary formats, interdependent file structures, and critical post-processing decision making, they also have unique characteristics that require rethinking existing digital curation approaches. The distinctive characteristics of 3D data require digital curators to look more closely at the creation and reuse phases of the 3D data lifecycle and more fully consider the needs of creators and users of 3D data who may want to access or distribute these contents using virtual environments with immersive interfaces. Cook and Lischer-Katz (2019) have defined three important preservation areas related to 3D/VR where libraries should take the lead:

- Managing VR hardware and software obsolescence;
- Establishing file formats for archiving 3D content;

[4] https://cs3dp.org
[5] https://web.archive.org/web/20240811094135/https://lib.vt.edu/research-teaching/lib3dvr.html
[6] https://www.gsd.harvard.edu/frances-loeb-library/building-for-tomorrow/
[7] https://3d.si.edu

- Developing metadata standards.

Case Studies: Two Levels of Volumetric Video Capture

The following section describes two levels of volumetric video capture, professional and prosumer, in order to illustrate the common challenges and differences across different types of workflows that digital curators may encounter and need to integrate into the digital curation lifecycle.

Professional Level: VoluCap Studios

The first case study presents the highest level of production quality. This "professional level" is illustrated by the workflows of the media production company VoluCap, based in Potsdam, Germany (just outside Berlin). VoluCap works with Hollywood production studios on big budget films, such as *The Matrix: Resurrections*.[8] The VoluCap system consists of a cylindrical room embedded with 40 (or more) custom-built cameras that encircle the subject, who is lit from all sides by custom Arri lighting rigs. The video stream from the cameras has an approximate resolution of 3000 megapixels (although 800MP is often enough for many projects). The walls of the cylinder are primarily white screens, with some green areas, which makes it easy to remove any background from the captured volumetric image. See Figures 1–3 for images of the volumetric capture space. As of February 2024, VoluCap's studio rates were as follows (VoluCap, personal communication, February 11, 2024):

- Studio rental and crew: "The daily rate for our studio, including the crew and equipment, is €10,000" (~$10,800USD)
- Export: "For a minute of export at 800 megapixels, we charge €5,000" (~$5,400USD)
- Setup fee: "The setup fee is €3,000" (~$3,200USD)

While these rates are likely to go down over time, it shows how costly volumetric video production is when implemented at the professional level. Given the high cost and the large file sizes involved, developing digital curation guidelines is necessary to support reuse of costly data.

In June 2023, the project team visited the VoluCap studio within the historic Studio Babelsberg. As part of the NEH grant, the team had enough budget to carry out a short volumetric video capture session. Ordinarily, the cost to use the studio is over $10,000 per hour, but we were given a significant discount by VoluCap. VoluCap is collaborating with the project team and is interested in contributing their expertise to the development of research and teaching applications for volumetric video technology. The project is focused on capturing BIPOC stories from the first half of the twentieth century, particularly the World War II era, and to that end we were able, through co-author, Dr. Rashida Braggs's connections, to locate someone in Berlin who had a compelling story to tell. Mike Russell, an American expatriate living in Berlin, offered a range of stories about his late father, James D. Russell, an American military service member stationed in Italy during WWII. His story was captured in several takes using the VoluCap system, and the best take was selected and processed to produce a volumetric video clip that was 3 minutes and 50 seconds long (see Figure 4 below).

[8] https://www.imdb.com/title/tt10838180/

Figure 1. VoluCap studio (exterior).

Figure 2. VoluCap studio enclosure (interior).

Figure 3. VoluCap's capture studio.

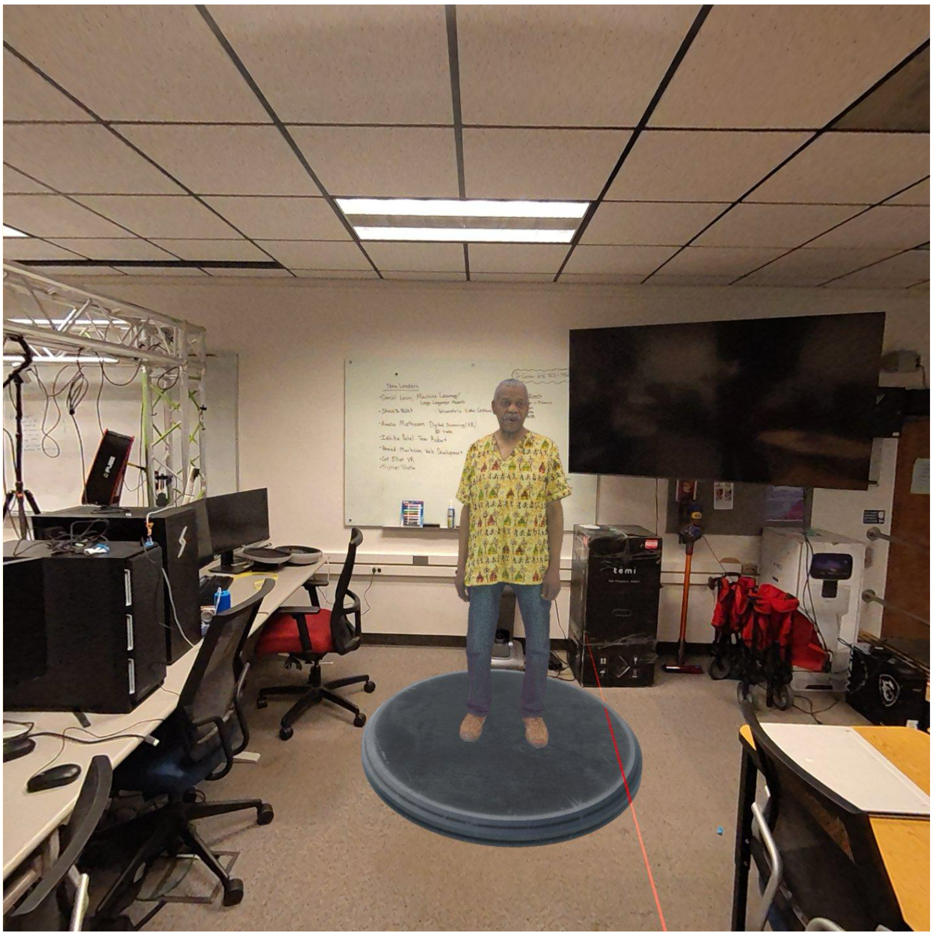

Figure 4. Still image from final volumetric video capture of Mike Russell.

The overall experience was like a film shoot, with makeup, rehearsal, multiple takes, and direction given by the VoluCap team in collaboration with the project team. Understanding that volumetric video sessions need to be rehearsed and involve makeup, wardrobe, and lighting considerations is important for interpreting the quality of the volumetric video produced. Pre-capture steps are just as important as the capture steps and equipment selection in the quality of the resulting outputs. The workflow that we observed at VoluCap Studios consisted of the following steps (some steps were hidden from the project team because of proprietary software and techniques):

- Pre-capture
    - Calibrate cameras and system
    - Rehearsal of actors outside of the studio space with their scripts
    - Select wardrobe and apply makeup
    - Rehearsal of actors in the studio under lighting
- Capture
    - Start video recording
    - Performer performs
    - Stop video recording
    - Setup for next take
    - Repeat capture process until all takes are completed
- Post-capture
    - Preview video clips as 2D lower res image to select suitable takes
    - Select clips for processing by VoluCap's custom software that uses machine learning to increase the speed and accuracy of its 3D processing
    - Produce derivative files:
        - 4K/2K: MP4s for use in VR (using video game engines, such as Unity or Unreal authoring software)
        - 1/1.5K: files for use in augmented reality using 8th Wall (cloud-based AR)

The visit to VoluCap studios produced several volumetric video clips of Mike Russell as well as project team members (and co-authors) Drs. Bryan Carter and Rashida Braggs, each one 3–4 minutes in length. Only Mr. Russell's clip was fully processed and converted into a volumetric video file suitable for viewing on a standard personal computer or VR headset. This prototype helped the project team better understand how volumetric video capture of historical figures and about historical events could be used in the classroom. Please see Table 1 below for data rates of the files generated by VoluCap for the volumetric prototype. By comparison, 4K HD video (at 30 frames per second) requires approximately 350MB of storage per minute of video.[9]

[9] https://help.encoding.com/knowledge-base/article/understanding-bitrates-in-video-files/

Table 1. Volumetric capture levels and related data rates (GB/minute) and file formats (note: excluding size of any associated audio files).

Production level	Raw capture		Processing		Output (GB/min)		Derivatives (AR) (GB/min)	
	File format	Data rates	File format	Data rates	File format	Data rate	File format	Data rates
Professional	Unknown proprietary format (7K)	1666	Compress to PNG or JPEG sequence	833.33	MP4 (H.264, 2K)	0.986	MP4 (H.264, 1K) + .ts files	0.701
					MP4 (H.264, 4K)	1.862	MP4 (H.264, 1.5K) + .ts files + M3U8 files	0.954

In this project, we identified several limitations of the professional-level volumetric capture process: its three-minute time limit; having to stand may be prohibitive for elderly and less mobile interviewees; difficulty transporting participants to an international or nationally distant location of the studio; the time, knowledge, cost, and person power it took to share, disseminate, and activate Mike Russell's VV; and VR's immersive and individualised quality can also limit how many can experience it together.

Key benefits of the volumetric prototype were also identified:

- High-resolution, professional scan detail that contributes to more "authentic"-feeling representations of people.
- Volumetric captures are most effective when they are presented alongside existing biographical materials of the participants, augmenting them with greater emotional, embodied and sensorial knowledge.
- The VR medium creates opportunities for deep immersion and direct, intimate connection with these stories.

Going through the process of planning and creating the prototype volumetric capture also illustrated the possibilities/opportunities for future uses. One of the best applications of volumetric capture is using it to bolster other biographical resources in order to more authentically illustrate emotion, portray the details of body movement, and capture the visual and sensorial details of the environment.

Prosumer Level: Center for Digital Humanities (University of Arizona)

The previous section described the professional-level volumetric capture process and impact of the VoluCap approach on the project's first case study. Clearly, this approach is cost-prohibitive for many academic faculty, but the workflow process and impact of volumetric video at the professional level illustrates the fundamental techniques behind it and the potential for enhancing student engagement and historical understanding. As a second case study, we present volumetric video capture at the "prosumer" production level at the University of Arizona's Center for Digital Humanities (CDH).

CDH[10] uses volumetric video capture for humanistic inquiry through immersive storytelling and immersive pedagogy to enhance student engagement and understanding. In their volumetric projects, the use of presence afforded by volumetric video helps to transform users'

[10] https://digitalhumanities.arizona.edu/

experiences with the past, and they are currently working with museums and other cultural heritage institutions to offer new ways of engaging the public with the stories, places, and artifacts of the past.

CDH started experimenting with volumetric video in 2020 after receiving a grant from the Knight Foundation. The award invited the project investigators, The Colored Girls Museum, and Dr. Carter to the Microsoft headquarters, where volumetric capture had been introduced as one of the emerging technologies on which Microsoft was focusing. This grant funded the development of a prototype volumetric capture stage at CDH. Much has changed since then; as technology has continually evolved, prosumer-level volumetric capture has also evolved and gotten cheaper and easier to manage.

After using other platforms, such as DepthKit, CDH currently uses a platform called Forma Vision[11] to capture and process volumetric video. Using Forma Vision, lab staff can record locally, upload to the CDH AWS server, and stream directly to its web AR platform, 8th Wall.[12]

The following subsection describes the current workflow at CDH using Forma Vision software. This workflow supports live streaming as well as recording functionality.

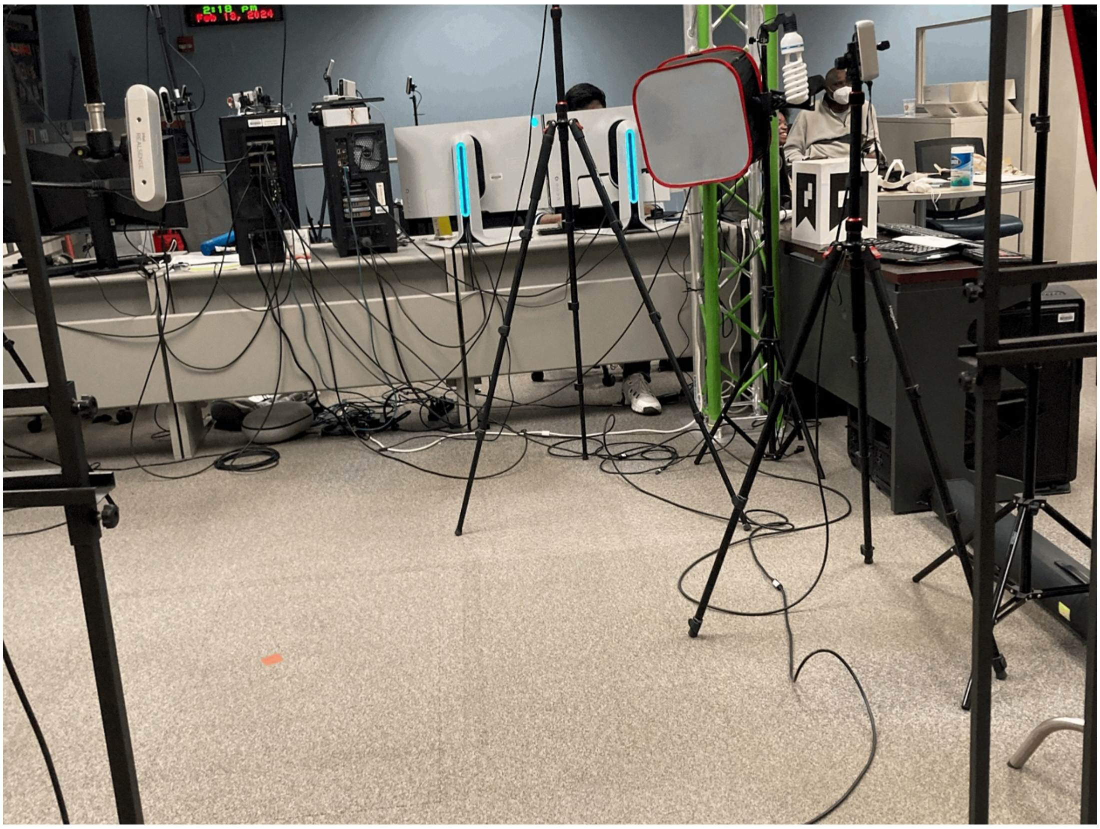

Figure 5. Center for Digital Humanities (CDH) volumetric capture studio.

Volumetric capture workflow at Center for Digital Humanities

The following workflow uses the Forma Vision volumetric capture system with 8 Intel RealSense D455 Cameras to capture and process volumetric video:

- Pre-capture
    - Calibrate cameras and system (takes approximately 15 minutes)

[11] https://www.formavision.io/
[12] https://www.8thwall.com/

    - Set up scene with “talent” (i.e., actors or other subjects)
    - Set up lighting
    - Set up teleprompter with script
    - Set up sound: attach wireless mic to talent, test recording for ten seconds, check levels.
    - Choose volumetric method: Select either "Record" or "Broadcast Live"

- Capture
    - Start capture: record or broadcast subject
    - Performer performs
    - Stop capture when content is complete
    - Repeat until all takes or scenes are recorded (for recording sessions only)

- Post-capture (for recording sessions only):
    - Process video files with Forma Vision software (time depends on length of recording time)
    - Publish/disseminate content for augmented reality (AR) applications
        - Push processed MP4 file and associated txt file to folder on Center for Digital Humanities Amazon Web Services (AWS) cloud storage
        - Copy file path for use with AR provider 8th Wall
        - Open 8th Wall Forma Vision AR Template and paste file path copied earlier
        - Process on 8th Wall and generate QR code for AR experience
        - Share QR code to disseminate volumetric record or use in class

This new process creates the following file formats and file sizes (approximate):

Table 2. Prosumer level: Volumetric capture file formats and related data rates.

Production level	Raw capture/processed file	
	File format	Data rates
Prosumer	MP4 + BYTES file (proprietary Forma Vision binary file)	~98MB (~0.098GB) / minute (MP4 File) ~2.9MB (~0.0029GB)/ minute (BYTES file)

By observing the workflows at CDH for Case Study 2, we identified a set of basic digital curation challenges related to prosumer-level volumetric video capture. At present, there are no archival workflows in place at CDH, but the scholarly, educational, and cultural value of the volumetric videos that are created there would likely benefit from being archived by an information institution, such as an academic library, and issued persistent, unique identifiers so that they can be cited in journal articles or accessed through open educational resources (OERs).

There is also some concern about the long-term sustainability of the files produced using the latest capture workflow that employs the Forma Vision system. While this workflow is considerably simplified, the master files with the highest resolution consist of an MP4 file and a required BYTES file. While MP4 files are well documented and low-risk file formats for preservation, the BYTES file is a proprietary format developed by Forma Vision. Documentation about this file format does not appear to be publicly available. The BYTES file is necessary for properly assembling the volumetric video file from the video frames of the MP4 file. Without this file, the MP4 looks like Figure 6, below, which shows how a frame of MP4 video looks without the BYTES file.

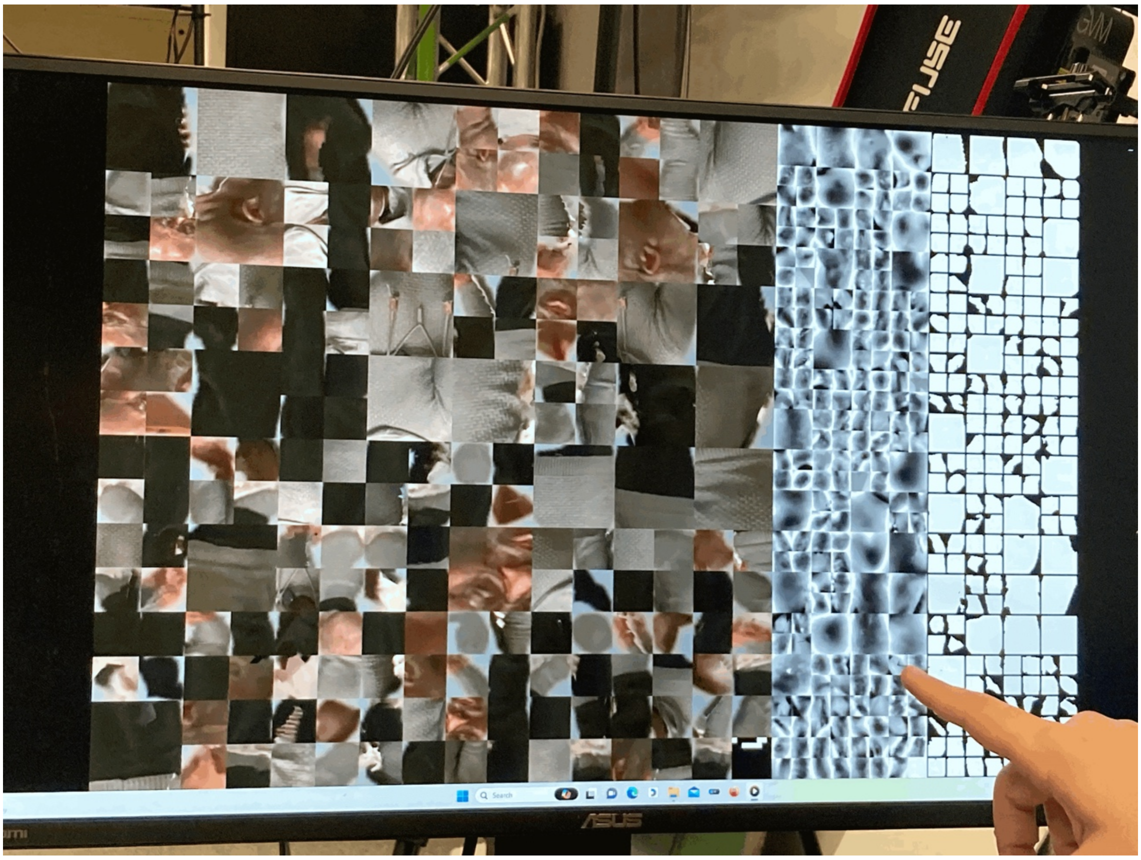

Figure 6. Video frame from MP4 file produced via Forma Vision workflow without rendering with BYTES file.

Discussion: Challenges of Preserving and Curating Volumetric Video Capture

From these two case studies, it is possible to identify several critical curation and preservation challenges for volumetric video:

1. Storing very large files

2. Establishing appraisal/selection criteria

3. Choosing sustainable file formats for preservation and future reuse

4. Metadata for description, administration, documentation, and paradata

5. Need for digital repositories

1. Storing Very Large Files

In planning any volumetric capture project that involves recording, the large file sizes produced at different stages (capture, processing, and distribution) need to be considered and planned for. In the case of VoluCap, they use local servers and mirror the data to off-site backups. For CDH, raw capture files are stored and processed locally, and then processed files are uploaded to their cloud storage, hosted by AWS. Project teams need to plan for and allocate the necessary storage at each point along the production pipeline, either using local hosting, cloud-based hosting, or a mixture. At the same time, storing all of these files for the long term may not be possible given the cost required to store and transfer very large files, so appraisal and selection criteria need to be established to determine which files are selected for long-term preservation.

2. Establishing Appraisal and Selection Criteria

Even with seemingly ever-expanding data storage resources, the large size of volumetric video files requires decision making about which files to keep and which to dispose of. There are both economic and environmental costs to data storage, and archival researchers are grappling with environmentally sustainable models for long-term digital preservation (see Lischer-Katz, 2017). Jinfang Niu (2014) has outlined a set of criteria for appraising digital records: mission alignment ("whether the resource supports the mission and falls within the scope of the collection policy of a preserver" [p. 71]); value of the digital resources (primary and secondary value); cost; and feasibility ("often determined by the technical capacity of the preserver" as well as the cost [p. 73]). Mission alignment means that the collecting and preserving of particular materials aligns with the mission of the preserving organisation. In the case of volumetric video projects at CDH, project outputs could be seen to align with the collecting imperatives of the University of Arizona Libraries, which systematically collects key faculty and student scholarly outputs. In determining the value of resources, Niu (2014) points to several clusters of factors to consider, including factors related to quality, including "authenticity, reliability, integrity and accuracy of resources" (p. 72); factors related to the use of resources, including "usefulness, usability and accessibility" (p. 72); and factors related to how the resources fit into the universe of records, including "uniqueness, diversity, and representativeness" (p. 72). Records have primary value to their producers, including "administrative value, fiscal value, and legal value", (Tibbo, 2003) but they often have secondary value to other users, including "evidential values and informational values" (Niu, 2014, p. 72). In the case of volumetric video, we need to consider these factors in relation to the record creators (the project team at CDH) and other users (other faculty and students across campus and at other institutions). The value of volumetric video, thus, depends on the content and who will find that content useful into the future, as well as the technical characteristics of the volumetric video, including file sizes, file formats, and metadata schemas. The third criterion, cost, "may become the decisive factor in selecting multiple versions of the same content" (Niu, 2014, p. 72).

3. Evaluating the Sustainability of File Formats

Carrying out appraisal and selection also requires considering which file formats are most likely to be supported in the future and choosing to keep those formats. The longevity of support for a file format is not always predictable, but the Library of Congress has developed a set of criteria to use when evaluating the sustainability of digital file formats. These "sustainability factors"[13] can be applied to evaluating file formats used in volumetric creation.

The file formats that appear in the volumetric video capture workflows discussed in the case studies in this report include MP4, PLY, OBJ, PNG, JPEG, and BYTES files. Other production pipelines may include other file formats. If we evaluate these file formats using the Library of Congress Sustainability Guidelines, we see that most of these files are suitable for preservation, at least notwithstanding the different software that can play them back properly.

Current video player codecs can playback MPEG4 files, although in a flat, unwrapped form optimised for 3D processing software (see figure 7 below); and OBJ files can be opened in a variety of software packages, but OBJ sequences require more specialised software. The only concern we have is with the BYTES file used in the Forma Vision capture workflow. This format is essential to properly rendering the 3D volumetric capture from the associated MP4 file. It appears to be a proprietary format. Additional research is needed to determine how widely supported this format is.

Summary of the Library of Congress recommendations for these formats:

- MP4 (.h264): raw video files
  - Library of Congress does not have a recommendation yet[14]
- PNG: still image files
  - "a preferred format for photographs in digital form, other graphic images in digital form and 2D and 3D Computer Aided Design raster images"[15]
- PLY: 3D mesh files
  - "acceptable format for Scanned 3D Objects"[16]
- OBJ: 3D mesh files
  - "acceptable format for Scanned 3D Objects (output from photogrammetry scanning)"[17]
- BYTES files: proprietary file format developed by Forma Vision.[18]
  - Required for assembling MP4 files created via the Forma Vision workflow.
  - Unknown to the Library of Congress. Additional research is necessary to find the documentation for this format. May be protected by Forma Vision intellectual property.

As we can see from this list, besides some of the proprietary capture formats, most of the file formats used in volumetric video pipelines are well documented and have strong sustainability factors based on the Library of Congress framework. However, it is important to note that just because a particular file format is renderable, does not mean that it will be rendered properly. In the case of the volumetric MP4 files, they can be rendered in a standard video player but require additional processing or software to be rendered properly as a 3D volumetric video file.

---

[13] https://www.loc.gov/preservation/digital/formats/sustain/sustain.shtml
[14] https://www.loc.gov/preservation/digital/formats/fdd/fdd000081.shtml
[15] https://www.loc.gov/preservation/digital/formats/fdd/fdd000153.shtml
[16] https://www.loc.gov/preservation/digital/formats/fdd/fdd000501.shtml
[17] https://www.loc.gov/preservation/resources/rfs/design3D.html
[18] https://www.formavision.io

Figure 7. Frame of volumetric video improperly rendered as 2D video.

This demonstrates the importance of creating documentation that specifies how the volumetric video was created and how it should be rendered (e.g., describing the software used to create it and the types of software that should be used to display it). It may also be important to plan for archiving any essential viewing software, too (e.g., Unity plug-ins used to author VR experiences). Volumetric video can be integrated into VR and AR experiences, but there may be compatibility issues in the future, so preservationists should consider archiving some of the VR authoring applications to ensure that volumetric video renders properly in the future. One format that may be appropriate to consider as a preservation master is an OBJ or PLY stream, which is a series of 3D models, one for each frame of video. However, audio and metadata need to be archived alongside the 3D model stream, and later re-integrated with it.

4. Metadata for Volumetric Video

In the case studies discussed in this report, detailed metadata is not integrated into the workflows. Beyond the use of basic file names, groups of files are described at the level of individual takes and projects. To support the discoverability and the reusability and interpretation of the volumetric videos, additional metadata considerations should be made. The FAIR data principles (Jacobsen et al., 2020; Wilkinson et al., 2016) stress the importance of ensuring that data are Findable, Accessible, Interoperable, and Reusable. Metadata plays an important role in supporting FAIR principles. For instance, to be "reusable", data (or metadata) are "richly described with a plurality of accurate and relevant attributes [...,] are associated with detailed provenance [... and] meet domain-relevant community standards" (Wilkinson et al., 2016, p. 4).

Given the high cost in time and resources required to produce one volumetric video, it seems logical to follow the FAIR principles to ensure that volumetric video is usable and reusable for a variety of users and uses outside of its original context. Therefore, this report recommends that metadata should be used in conjunction with digital repositories and search engines to make volumetric video findable, accessible, interoperable, and reusable.

Technical metadata

Creators and curators of volumetric video should consider the technical characteristics that they will need to document to ensure the long-term use and interoperability of their 3D content. Recording technical metadata is essential for being able to render the files in the future, interpret their quality and creation process, as well as use them on other platforms or convert them to other formats for use and reuse. Creators of volumetric video content might consider documenting the following characteristics of volumetric video files:

- File information:
  - File format: specify the format of the volumetric video file (e.g., MPEG-4 for video files, OBJ stream for VR/AR files).
  - File format codecs (if video)
  - File size: record the size of the file in bytes.
  - Checksums and fixity information: include information about checksums or fixity checks to ensure data integrity.
- Capture and acquisition details:
  - Camera information: describe the cameras and camera configuration used for capturing the volumetric video, including make, model, and specifications.
  - Calibration data: include details about the calibration process used to align multiple cameras in the volumetric capture setup. Provide a link to any calibration files.
  - Depth sensing technology: specify the technology or sensors used for capturing depth information.
- Spatial information:
  - Resolution: indicate the spatial resolution of the volumetric video, including dimensions in pixels or other relevant units.
  - Spatial coordinates: capture information about the spatial coordinates and dimensions of the captured scene.
  - Coordinate reference system: specify the coordinate reference system used for spatial data.
- Temporal information:
  - Frame rate: record the frame rate of the volumetric video.
  - Temporal range: indicate the start and end times of the volumetric video sequence.

- Software and processing information:
    - Capture software: specify the software used during the capture process.
    - Post-processing tools: document any software or tools used for post-processing and rendering volumetric video.
    - Processing parameters: include information about parameters used during processing steps, such as compression settings.
- Spatial audio information (if applicable):
    - If the volumetric video includes spatial audio, include metadata about the audio capture, processing, and spatialisation.

Preservation metadata

Digital preservation best practices require that any preservation actions carried out with a digital resource are well-documented. At a minimum, digital preservationists should record the following types of information for preserving any digital objects:

- Migration history: document any migrations or format changes that have occurred during the preservation process.
- Fixity check results: include information about the results of fixity checks performed on the volumetric video over time.
- Versioning information: if applicable, include version information for the volumetric video file and associated metadata.
- External dependencies: document any external dependencies, such as specific software versions or libraries required for playback or rendering.

PREMIS (PREservation Metadata Implementation Strategies)[19] offers additional guidance for digital preservation metadata, but these are general, minimum requirements for preserving digital objects.

Descriptive metadata

Volumetric video has properties of both 2D video (frame-based audiovisual media with a linear form) and 3D models (spatial information stored as 3D mesh files with associated texture maps), thus, this section will explore relevant video and 3D metadata recommendations and see which aspects together might be sufficient for describing the content of volumetric video files.

These are some of the essential properties to consider recording as descriptive metadata for volumetric video projects (most are relevant to all types of video projects):

- Title and project information
    - Title: include a title for the volumetric video piece.
    - Project: indicate the project that the digital object is a part of.
- Creators and contributors:
    - Creator(s): identify the individuals or entities responsible for creating the volumetric video.

[19] https://www.loc.gov/standards/premis/v3/premis-3-0-final.pdf

  - Contributors: record additional contributors, such as technicians, researchers, or artists involved in the creation process.
  - Actors/subjects: record the names of any actors, faculty, students, historical figures, etc. captured in the volumetric video.
- Rights and usage information:
  - Copyright information: include details about copyright holders and licensing information.
  - Usage restrictions: specify any restrictions on the use or distribution of the volumetric video.
- Subject information (ideally using controlled vocabularies and authorities, e.g., those maintained by the Library of Congress[20]).
  - Main topics of the video, based on what the video is about and any dialogue
  - Any material objects included in the video
  - Historical figures included in the video
  - Temporal coverage (time)
  - Geographic coverage (place)
- Shot-level description
  - Actions/events (linked to specific timecodes)
  - Dialogue

Metadata schemas for describing audiovisual media already exist. While they are not designed to capture 3D characteristics, metadata schemas designed for describing audiovisual media, such as PBCore[21] and MPEG-7[22], should also be considered. And generic schemas for describing digital objects, such as Dublin Core Metadata Initiative (DCMI) terms,[23] may also be worth considering when designing basic, easily implemented descriptive schemas for volumetric video. An important element to consider is ways of forming a link between the volumetric video and information about the subject matter (the real-world people and objects captured in the recording).

Documentation and paradata

Because of the complex creation and processing pipelines used to create 3D digital objects and volumetric video, there has been a call in various fields for recording "paradata" for these workflows. Paradata in this sense is information recorded about creation processes, techniques, and technologies used to create 3D outputs. Paradata has also been used to describe interpretive practices and other uses of datasets, as well as the interpretive processes by which digital, intellectual objects are created. Current projects being pursued through the Pure3D initiative,[24] such as *Paradata in 3D - Scholarship Intellectual Transparency and Scholarly Argumentation in Digital Heritage* (January–December 2025), seek to "enhance the intellectual transparency and scholarly argumentation within digital heritage by focusing on the development and implementation of

[20] https://id.loc.gov/authorities/subjects.html
[21] https://pbcore.org/what-is-pbcore
[22] https://www.mpeg.org/standards/MPEG-7/
[23] https://www.dublincore.org/specifications/dublin-core/dcmi-terms/
[24] https://pure3d.eu/

paradata standards" (Pure3D, 2025). The particular methods of creation and the decisions made during processing have a significant impact on the final 3D model output and how accurate measurements can be on the resulting models. Since the publication of the *London Charter (for the Use of 3-Dimensional Visualisation in the Research and Communication of Cultural Heritage)* in 2006,[25] which codified the importance of paradata, paradata in its various forms, is increasingly seen as important for archaeological 3D data production and cultural heritage preservation. Documentation helps future users to interpret the visual and aural content that they encounter, as well as the historical and intellectual content, helping to support the research values of transparency, reproducibility, and reusability of findings. This documentation may be structured (as in the case of Cultural Heritage Imaging's Digital Lab Notebook software)[26] or unstructured (more narrative-based or lab notebooks).

5. Need for Digital Repositories

A big challenge identified in this project was the need for suitable digital repositories for storing and making accessible collections of volumetric video files. Creators, users, and other stakeholders of volumetric video need to ask these types of questions: Where should one store these resources for the long term? How do we ensure that users can find them and access them? Who takes on the costs of storing them long term? What preservation actions might need to be taken in the future?

Some form of repository to host and provide access to volumetric video is necessary. There are scholarly repositories that accept 3D data, such as Morphosource[27] and tDAR,[28] but no repositories exist that can properly ingest and make accessible volumetric video. Generalist repositories, such as Figshare[29] and Zenodo,[30] limit the size of files that can be uploaded (20GB for Figshare and 50GB for Zenodo), and they cannot render volumetric video through an online viewer. Currently, volumetric video can be made accessible to users by uploading it to a third-party host such as 8th Wall,[31] which uses various WebAR tools to make volumetric video accessible via augmented reality apps, or by incorporating the volumetric video into a VR application, using Unity3D or Unreal authoring software. While 8th Wall maintains a library of projects, it is not searchable or accessible in the same way that a digital collection hosted by a repository would be. A basic repository should provide access to a lower-resolution, streamable format, a higher-resolution VR- or AR-ready version, and a high-resolution archival master file that can be used for creating new derivative files. At the very least, a WebVR viewer should be embeddable and streamable for web and mobile viewing, while the higher-resolution versions could be stored and accessed as download-only, since current browsers and mobile devices may have difficulty rendering higher-resolution volumetric video.

Implications for Digital Curation Models

The Archaeology Data Service, in their preservation manual for archiving 3D data, suggest an approach that considers what they call "preservation intervention points" in which digital curators can become involved and assist, making it more likely that the integrity of 3D data will be maintained throughout the research lifecycle: from planning to collection and creation to processing to curation and access (see p. 19 in CS3DP book, Moore et al., 2022). Because volumetric video shares common properties of 3D, VR, and audiovisual formats, looking at past research on the curation and preservation challenges of these formats can help guide inquiry

[25] https://londoncharter.org/fileadmin/templates/main/docs/london_charter_1_1_en.pdf
[26] https://culturalheritageimaging.org/What_We_Offer/Downloads/DLN/index.html
[27] https://www.morphosource.org/
[28] https://core.tdar.org/
[29] https://figshare.com/
[30] https://zenodo.org/
[31] https://www.8thwall.com/products-web

into volumetric video challenges. Lischer-Katz et al. (2019) found that while 3D data formats share many of the same challenges as other complex data types (such as digital images or video), including large files sizes, proprietary formats, interdependent file structures, and critical post-processing decision making—they also have a unique mix of characteristics that will require ingenious and creative uses of existing approaches and further technical refinement. The distinctive characteristics of 3D data require digital curators to look more closely at the creation and reuse phases of the 3D data lifecycle and more fully consider the needs of creators and users of 3D data who may want to access or distribute these contents using virtual platforms (e.g., VR and AR).

Conclusion

Volumetric video capture is a new technique for creating complex digital objects that may be useful in a variety of research, teaching, and commercial contexts. Multiple scales of production are available, meaning that practitioners of different means may adopt these tools to create various levels of quality in their volumetric captures. Cost and storage of these materials are still significant barriers of entry to using these tools, but as processing power continues to increase in personal computers and mobile devices, and new compression algorithms and AI tools can produce smaller files sizes, it is likely that volumetric video capture will become increasingly common in a number of contexts, and volumetric will be truly “democratized” (Carter et al., 2021).

Additional work needs to be done to provide guidance on preservation planning from the beginning of volumetric capture projects. The high cost of high-quality volumetric capture (professional and prosumer levels) makes it essential that efforts are made to ensure that these digital objects are discoverable and usable by others in the future. While it is not entirely certain how future users might want to use or interact with volumetric video materials, it is clear that flexibility and interoperability are essential. Users will likely want to access the volumetric material on any number of digital devices, using augmented reality, virtual reality, and new technologies that have yet to be developed. Making sure that digital objects are preserved with this idea in mind is important to maximise usability and return on investment. The selection of file types and the selection of which stages in the production process should be preserved are essential considerations. In the professional and prosumer cases we looked at, the 3D mesh sequence files are the highest-quality file format in the workflow that is non-proprietary or locked into the specific cameras or capture software used to produce it. In addition, the high-resolution MP4 files that are used to create virtual reality or augmented reality experiences are also important to preserve because they make it easy for a variety of users to reuse the volumetric video for their own uses.

One final area in need of future research is the ethical and legal issues associated with volumetric video. This is an area that could benefit from bringing both philosophers and humanists across campus into conversation with legal scholars to consider these challenges. Like other creative works, there are copyright and licensing issues associated with volumetric video. What types of licences should be used by academic institutions for the volumetric videos they create? Additional research needs to be done in this area to ensure that those whose bodies, movements, and voices are captured are respected and their rights maintained, which is especially pressing as volumetric video becomes more and more common and the potential for use with AI systems becomes ever more apparent. For instance, volumetric video could one day be used to train AI in order to produce very convincing, three-dimensional deep fakes or other content that the volumetric video subject did not agree to or expect.

In closing, from these two case studies and a review of related literature, we can see the potential benefits from volumetric capture for enhancing research and pedagogy in the humanities, and at the same time, we can see the challenges. Developing portable volumetric capture systems that can be more easily calibrated and produce professional results is necessary, as is additional research on the challenges of digital curation and preservation, in particular, the development of volumetric video access and preservation platforms. Developing consortia and other strategies for reducing costs and sharing infrastructure across the DH community would help make it easier for a broader range of scholars to engage with this technology. The future of volumetric video for widespread use in the humanities and related fields will depend on research that addresses these challenges.

Acknowledgements

This project was made possible in part by the National Endowment for the Humanities (Digital Humanities Advancement Grant, Level I). The views, findings, conclusions, and recommendations expressed in this publication do not necessarily represent those of the National Endowment for the Humanities.

An earlier version of this paper is available as the final report for the project: https://repository.arizona.edu/handle/10150/674673.

References

Aguayo, C., & Eames, C. (2023). Using mixed reality (XR) immersive learning to enhance environmental education. *The Journal of Environmental Education* 54(1), 58–71. https://doi.org/10.1080/00958964.2022.2152410

Bozorgi, K., & Lischer-Katz, Z. (2020). Using 3D/VR for research and cultural heritage preservation: Project update on the virtual Ganjali Khan project. *Preservation, Digital Technology and Culture* 49(2), 45-57. https://doi.org/10.1515/pdtc-2020-0017

Buckley, J., Seery, N., & Canty, D. (2018). A heuristic framework of spatial ability: A review and synthesis of spatial factor literature to support its translation into STEM education. *Educational Psychology Review* 30(3), 947-972.

Carter, B., Robinson, L., Acuna, E., Uralov, D., Ketema, T., & Gordon, M. (2021). The Robinhood of 3D: Democratizing volumetric capture for more accessible and innovative learning environments. In *ICERI2021 Proceedings* (pp. 4845-4850).

Chang, C., Debra Chena, C., & Chang, W. (2019). Research on immersion for learning using virtual reality, augmented reality and mixed reality. *Enfance* 3, 413-426. https://doi.org/10.3917/enf2.193.0413

Cook, M. N., & Lischer-Katz, Z. (2019). Integrating 3D and VR into research and pedagogy in higher education. In K. Varnum (ed.), *Beyond reality: Augmented, virtual, and mixed reality in the library* (pp. 69-85). Chicago: ALA Editions.

Dodds, H. (2021). Immersive learning environments: Designing XR into higher education. In S. Conklin, B. Oyarzun, R. M. Reese, & J. E. Stefaniak (eds.), *A practitioner's guide to instructional design in higher education.* EdTech Books. https://dx.doi.org/10.59668/164.4220

Donalek, C., Djorgovski, S. G., Cioc, A., Wang, A., Zhang, J., Lawler, E., ... & Longo, G. (2014). Immersive and collaborative data visualization using virtual reality platforms. In *Proceedings of 2014 IEEE International Conference on Big Data* (pp. 609-614).

Endres, B. (2019). *Digitizing medieval manuscripts: The St. Chad Gospels, materiality, recoveries, and representation in 2D & 3D.* Amsterdam: ARC Humanities Press. https://doi.org/10.1515/9781942401803

Fragkos, S., Tzimtzimis, E., Tzetzis, D., Dodun, O., & Kyratsis, P. (2018). 3D laser scanning and digital restoration of an archaeological find. In *MATEC Web Conference, 178*. https://doi.org/10.1051/matecconf/201817803013

Hostettler, M., Buhlke, A., Drummer, C., Emmenegger, L., Reich, J., Stäheli, C. (2024). Digital archaeology between hype and reality: The results of a survey on the use of 3D technologies in archaeology. In M. Hostettler, A. Buhlke, C. Drummer, L. Emmenegger, J. Reich, and C. Stäheli (eds.), *The 3 Dimensions of digitalised archaeology* (pp. 181–203). Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-031-53032-6_11

Jacobsen, A., de Miranda Azevedo, R., Juty, N., Batista, D., Coles, S., Cornet, R., ... & Schultes, E. (2020). FAIR principles: Interpretations and implementation considerations. *Data intelligence* 2(1–2), 10–29.

Karasik, A., & Smilansky, U. (2008). 3D scanning technology as a standard archaeological tool for pottery analysis: Practice and theory. *Journal of Archaeological Science* 35(5), 1148–1168. https://doi.org/10.1016/j.jas.2007.08.008

Lischer-Katz, Z. (2017). Studying the materiality of media archives in the age of digitization: Forensics, infrastructures and ecologies. *First Monday* 22(1–2). https://doi.org/10.5210/fm.v22i1.7263

Lischer-Katz, Z., Cook, M., & Boulden, K. (2018). Evaluating the impact of a virtual reality workstation in an academic library: Methodology and preliminary findings. *Proceedings of the Association for Information Science and Technology* 55(1), 300–308.

Lischer-Katz, Z., Cook, M. N., Hall, N., Hardesty, J., Wittenberg, J., Johnson, J., … Carlisle, T. (2019). Supporting virtual reality and 3D in academic libraries: Defining preservation and curation challenges. In *Proceedings of the 16th International Conference on Digital Preservation iPRES* 2019 (pp. 114–125), Amsterdam, The Netherlands (Sept. 16–20).

Meccawy, M. (2022). Creating an immersive XR learning experience: A roadmap for educators. *Electronics* 11(21), 3547. https://doi.org/10.3390/electronics11213547

Moore, J., Rountrey, A., & Scates Kettler, H. (eds.). (2022). *3D data creation to curation: Community standards for 3D data preservation*. Chicago, Illinois: Association of College and Research Libraries.

Niu, J. (2014). Appraisal and selection for digital curation. *International Journal of Digital Curation* 9(2), 65.

Pure3D. (2025). *Paradata in 3D scholarship - Intellectual transparency and scholarly argumentation in digital heritage*. https://pure3d.eu/paradata-in-3d-scholarship/

Qin, T., Cook, M., & Courtney, M. (2020). Exploring chemistry with wireless, PC-less portable virtual reality laboratories. *Journal of Chemical Education* 98(2), 521–529.

Rafi, A., Anuar, K., Samad, A., Hayati, M., & Mahadzir, M. (2005). Improving spatial ability using a Web-based Virtual Environment (WbVE). *Automation in Construction* 14(6), 707–715. https://doi.org/10.1016/j.autcon.2004.12.003

Rafi, A., Samsudin, K. A., & Said, C. S. (2008). Training in spatial visualization: The effects of training method and gender. *Journal of Educational Technology & Society* 11(3), 127–140.

Roberts Thompson, A.D., & Williams, M. (2016). Reflectance transformation imaging (RTI) of a little known Southeastern copper plate. *Southeastern Archaeology* 35(1), 70–80. https://doi.org/10.1179/2168472315Y.0000000008

Sadler, R., & Thrasher, T. (2023). XR: Crossing reality to enhance language learning. *CALICO Journal* 40(1), i-xi.

Schreer, O., Worchel, M., Diaz, R., Renault, S., Morgenstern, W., Feldmann, I., … Eisert, P. (2022). Preserving memories of contemporary witnesses using volumetric video. *i-com 2022* 21(1), 71-82. https://doi.org/10.1515/icom-2022-0015

Sorby, S., Casey, B., Veurink, N., & Dulaney, A. (2013). The role of spatial training in improving spatial and calculus performance in engineering students. *Learning and Individual Differences* 26, 20-29.

Sterne, J. (2012). *MP3: The meaning of a format.* Durham, NC: Duke University Press.

Tibbo, H. R. (2003). On the nature and importance of archiving in the digital age. In M. Zelkowitz (ed.), *Advances in computers, Vol. 57: Information repositories* (pp. 1-67). https://doi.org/10.1016/S0065-2458(03)57001-2

Volmar, A. (2017). Formats as media of cooperation. *Media in Action: Interdisciplinary Journal on Cooperative Media* 2, 9-28. https://doi.org/10.25969/mediarep/16226

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., ... Mons, B. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific Data* 3(1), 1-9.