Rewind, Pause, Playback: Addressing a Media Conservation Backlog at the Denver Art Museum
By Eddy Colloton and Kate Moomaw

INTRODUCTION

Due to the exhibition and acquisition of electronic media art and time-based media art often out pacing the development of best practices for the preservation of such artworks, the Denver Art Museum, like many museums, has developed a "backlog" of untreated and minimally cataloged objects necessary for the realization of the museum's media artwork collection. To combat this backlog, the museum has adopted an iterative, project-based approach to identification and treatment of such objects.
The Denver Art Museum's ongoing IMLS funded conservation project, centered on electronic media, will result in the migration of all media works in the collection to a central server, and updated catalog records for each object, either physical or digital, associated with those works.

REWIND: BACKGROUND

The museum has been collecting electronic media artworks for several decades, both from local artist like Gary Emrich, and canonical video artist like Nam June Paik. Before established media preservation practices were instituted at the museum, media objects were often stored and described as traditional art objects. Their dimensions carefully noted, minor scratches on cases detailed in condition fields, the objects were then placed in acid free boxes in art storage. These objects were treated as carefully as they could be by those charged with their stewardship, but with the benefit of hindsight, a lack of familiarity with the technologies involved in these works may have hampered those efforts.

A seminal point in the development of the electronic media preservation practices of the Denver Art Museum (DAM) occured in 2011, spurred by an exhibition that focused on media works. "Blink! Light Sound and the Moving Image" featured dozens of media works and resulted in the acquisition of 15 artworks, mostly single-channel video, including: *Of the North* (2008) by Steina, *Telephones* (1995) by Christian Marclay, *Winchester* (2002) by Jeremy Blake, *We Can Make It Rain but No One Came to Ask* (2005) by Walid Ra'ad and many others. This exhibition, and subsequent acquisitions, functioned as a bit of a "wake up call" for the museum.

The growing awareness of the complexities of electronic media at the museum lead to the formation of the Variable Media Working group, an organization of stakeholders from a variety of different departments. This group was responsible for several policy changes, including actively pursuing grant funding and other projects to develop their media conservation program. This research resulted in a summer internship, which was eventually performed by Eddy Colloton, at the time a student in NYU's Moving Image Archiving and Preservation program.

The 2015 internship focused on items from the AIGA archive, part of the DAM's Architecture, Design, and Graphics collection. While the emphasis of this paper is the ongoing IMLS grant funded media preservation project, this internship functioned as a pseudo pilot project for many of the practices and policies that the DAM has since implemented. To that end, a brief summary of the museum's goals and decision making process for that project will better contextualize the conservation department's current activities.

The AIGA collection contains a very wide variety of objects including posters, food packaging, bar soap, and websites. The electronic media from the AIGA archive arrived at the DAM stored on 148 VHS tapes, 193 optical discs, and 38 floppy discs. The collection includes music videos, websites, computer games,

and powerpoint presentations.

While the DAM was very aware of a large backlog of VHS tapes, without playback equipment available, and without funding to transfer the tapes through a vendor, it was determined that the only work that would be feasible with these tapes would be to transcribe their labels and take pictures of them - not a particularly exciting internship project, or a particularly large step forward for the accessibility of the collection. The removable media objects, on the other hand, would need to be mounted to be able to be described in any meaningful way, which also posed the opportunity to migrate media off of those objects.

With the help of the museum's Technology department, Archivematica was installed on a BitCurator virtual machine to eventually process and catalog 70 unique digital media works. This would not have been possible without the collaboration between the conservation department, the Technology department (running Archivematica on a BitCurator VM is no small feat), collections, and curatorial. Notably, the Development department was kept aware of internship activities and Colloton's desk became a stop on VIP tours, where he would amuse patrons with anecdotes about floppy discs.

PLAYBACK: IMLS E-MEDIA PROJECT

At the conclusion of the 2015 internship in August, the Conservation department was already working on a grant proposal for a broader preservation project, one that would touch every media object in the museum's collection. The grant's narrative highlighted the progress made through the internship, citing this as an "initial framework" for preserving digital media at the organization. The grant application was successful, and the project began in December of 2016. Colloton was invited back in an 18-month position as an Assistant Conservator beginning in March of 2017.

The goals set out in the project proposal are somewhat analogous to a traditional condition survey or risk assessment, in that the museum is hoping to develop a better understanding of the contents and condition of the collection. It differs, however, in that, as opposed to simply noting condition and prioritizing treatment for a later date, preservation actions are performed as objects are processed and assessed, in attempt to change their condition. Certain preservation actions are essentially required in order to have performed a thorough condition assessment of electronic media. Just as with the 2015 pilot project, the DAM's approach has been to marry identification and assessment of the media works in the collection, with migration, normalization, and other digital preservation actions.

While the success of the 2015 pilot project had bolstered confidence, the participants began this project knowing that both throughput and output would be much greater, and that the demands on the museum would be much higher. In order to facilitate this, investments in the museum's media conservation infrastructure were built into the grant. With over 200 works in the collection still stored on videotape, migration of tape-based video formats to file-based formats was clearly going to be a large part of the project, and so playback decks, signal generators, and waveform monitor/vectorscopes were purchased for quality assurance, along with hardware upgrades to the media lab's Mac Pro computer to ease wait times on processing large video files. Grant funds were also allocated to secure a support contract with Artefactual Systems (the developer of Archivematica), funding for cloud storage through DuraCloud, and contracts with 2 video migration specialists - one local, Post-Modern Media Services, and one in New York, Maurice Shechter and Bill Seery.

Due to the number of objects included in the scope of this project, a uniformity to treatment and description were necessary. This presented somewhat of a challenge, as electronic media in museum collections often resists uniformity. The necessity to treat works on a "case by case basis" is often a refrain in the field of electronic media conservation for a reason. Not every videotape, optical disc, or hard drive in the collection should be treated equally, so each object needs to be evaluated and understood in context. The balance struck between these two opposing needs, one of assembly-line productivity and the other bespoke contextualization, was to develop a workflow that captured and preserved a large amount of information quickly, while leaving options open for future treatment and study. Determining this scope, is one of the more complicated aspects of undergoing this type of project. What information to capture, how thorough to be, and when to know to move on, are always challenging aspects of a survey, and the preservation actions conservators are performing as a part of this project did not simplify this condonrum.

At a very high level, the DAM's workflow is as follows: All removable media in the collection (hard drives, optical discs, and floppy discs) are disk imaged, most commonly using the Guymager application, part of the BitCurator suite. Metadata describing the disk image is generated at that time, and then packaged with the disk image and transferred to Archivematica, where, through an ingest process, data is transferred to a CentOS server, and then backed up twice, both locally, and at an offsite location through DuraCloud.

Simultaneously, existing catalog records are modified, and additional records are created to describe the media found on those objects. Descriptions of these objects would be primarily technical, such as "this file, on this server, came from this physical object, and it is a video file with this MediaInfo output." Much of what one needed to know about a particular object in order to process and catalog that object, was ascertained during the first step of the process, disk imaging.

Disk imaging allows for sector-by-sector verification of a volume, ensuring the creation of a bit-for-bit copy. Therefore, simply for the assurance of securely copying information, it is a very valuable tool. On top of that the disk image will preserve the original filesystem of a volume in a write-protected environment, and allow the user to mount the volume and explore the directory structure much as a user of the original volume would have done. Following evaluation, the disk image is transferred from the BitCurator Virtual Machine to the staging directory for Archivematica as one file, instead of many files, averting risk of file corruption or modification upon copy.

The next step of the DAM's workflow is to move files through the open-source digital preservation software Archivematica. Archivematica is essentially a suite of open source tools stacked on top of each other. The software uses SleuthKit, MediaInfo, ExifTool, FFmpeg, BagIt, and many other tools to describe and package files for long term storage, as well as access. DAM's conservators submit disk images to Archivematica, and extract the files in those disk images through automation, which is then documented as a PREMIS event in the METS wrapped XML file that the software produces.

The museum has chosen to keep both the original disk image and the extracted files. Extracting the files as part of the ingest process, through the tsk_recover command, part of SleuthKit, allows one to take advantage of the automated identification and characterization tools that are built into Archivematica. Keeping the original disk image ensures a bit-for-bit copy of the original volume (assuming that no bad sectors were encountered upon imaging).

As a part of the ingest process, Archivematica allows for normalization. Normalization is commonly defined in the archive and libraries community as reducing the number of file types a steward of digital media is required to monitor by converting similar file types to an "archival standard." In general, but certainly in media conservation, this practice is less and less common. However, there is a benefit in creating a more sustainable derivative of a file, if that file is at risk at the time of ingest. While no digital video format that is truly obsolete, that is to say, completely unrenderable under any circumstances, has been encountered during this project, the possibility does not seem unimaginable. The International Association of Sound and Audiovisual Archives (IASA) will slowly be releasing their "Guidelines for the Preservation of Video Recordings" this year, but the introduction has already referenced the recommendation of transcoding Windows Media recordings to other formats due to their lack of interoperability.[1] Even in this instance keeping the original file type is the best course of action, but having a sustainable derivative, a copy of the original media in a known file type commonly used by the museum, serves a dual purpose of guarding against the obsolescence of a file format while providing an easily accessible version of the file. The only downside to this is transcoding time and storage space.

An aspect of this project which has evolved over time is the project staff's emphasis on, and understanding of, server performance. Given the workflow's reliance on Archivematica, our ability to monitor, understand and improve the performance of the servers is very important. The museum's original cloud backup strategy serves as a good example. Backups were initially scheduled to run as soon as new media was added to the directory. However, the Amazon Web Service S3 servers, that the museum rents as part of its contract with DuraCloud, have a file size limit of 10 gb per file. DuraCloud has of course thought of this, and the service comes with a java-based application that automatically "chunks" larger files into 10 gb pieces, and then re-compiles them upon retrieval. This chunking and uploading, though, demands a significant amount of memory. These problems were revealed through a series of server failures during ingest of files through Archivematica, while the server was attempting to upload files to DuraCloud.

The demands of the two software were competing for too few resources and as a result the server would crash and work would have to be put on hold while the issue was troubleshot. Moving the backups to the evening solved this issue, but several other solutions, such as adding processing power, additional memory, and additional storage space were tried first.

PAUSE: LESSONS LEARNED

This realization of the challenges of working at scale essentially summarizes most of the lessons learned from this project. Complex systems are complex to maintain, and the amount and diversity of media that is in the DAM's collection (not atypical of an art museum), effectively amounts to a stress test for those systems. There have been an exhausting list of issues with Archivematica, many of them due to the museum's instance of Archivematica being updated to an unreleased version of the software, shortly before the developers hit a major delay causing the next stable release to be approximately 7 months late. The need to troubleshoot the errors presented by Archivematica and exhaustive testing of each subsequent release, until finally hitting a stable release in May, has caused significant delays.

The need for almost constant contact with Archivematica support, and the need for frequent communication with our other vendors, underlines the labor involved with partnerships. When this project

[1] Guidelines for the Preservation of Video Recordings, IASA-TC 06, A-7, https://www.iasa-web.org/sites/default/files/publications/IASA-TC_06-A_20180518.pdf

started the project organizers had viewed the previously mentioned "infrastructure" of contracts with vendors as an investment in efficiency, something that the DAM could "contract out," and therefore not have to worry about. This is not the case. Museum staff have been highly involved with the work performed by all of its contractors, which takes time and energy away from the pressing task of moving media through the workflow.

Pace has been hard to measure in that context. How costly is a delay, if solving the delay will improve the workflow? The grant proposal stated that the museum intended to process 10 objects a week, a goal that conservators sometimes exceeded but more often fell short of. In the early days of struggling with Archivematica, delays were seen as "growing pains" that could eventually be overcome, and eventually benefited from. Essentially, the goal was to make up for early delays through gained knowledge. Unfortunately, gained knowledge often led to further complexities, and growing pains of the project's early stages, once overcome, did not yield exponentially faster progress.

Realistic quotas for a project of this nature in theory could be a very helpful tool, but from the perspective of this project could be very difficult to set. The early stages of a project are going to need to account for unexpected delays, including configuration, and likely re-configuration of systems essential to the process. Moreover, not all objects are going to demand the same amount of attention, and so "weighting" certain materials, works, or processes, to ensure due amount of resources have been allocated for specific tasks, is of significant importance.

In spite of these challenges, much has been accomplished in a relatively short amount of time, and we know much more about our collection than ever before. The DAM currently has 365 AIPs in the museum's Archivematica storage service. Those AIPs are comprised of 43,456 files, or around 2.5 terabytes of data. Additionally, 1,117 catalog records have been created or modified: 325 of those records describe Quicktime video files, 37 html files, 117 ffv1 files, 33 WAV files, and 193 records that describe raw disk images.

The DAM's collection will be in much better condition by the end of this project, but there will still be a good deal of work yet to be done. As works have been processed, project conservators have taken note of works, and types of works, that still need further research and treatment. Chief among these are software based works that rely on obsolete operating systems. Another example are works that were not collected fully, or that there are better preservation elements for, such as websites from the AIGA collection that were delivered to the museum as screenshots. For the more complex media in the collection, more thorough documentation and display specifications would make the works more accessible and less intimidating for exhibition. There's also the matter of justifying continuing costs related to digital preservation. The museum plans to forego further Archivematica support, but will still have costs related to digital storage, such as the offsite storage provided through DuraCloud. Communicating how important those costs are will be an ongoing challenge.

At the Denver Art Museum the ability to complete two preservation projects that have not been driven by programming or acquisition schedules is admittedly a luxury. While the projects have not been without time constraints, these projects have given the museum's conservators the ability to plan and work

exclusively on preservation, without interference. This model has been beneficial to the museum for developing procedures and policies, and now has resulted in collection wide preservation actions. The DAM's iterative, project-based approach to media conservation has allowed the museum to make large strides in a relatively short amount of time. Hopefully, this approach can serve as a model for addressing similar "backlogs" at other museums.