Electronic Media Review

The postprint publication of AIC's Electronic Media Group

Making Practice Practical: Developing Digital Preservation for Storage of Artworks at the Museum of Modern Art

Amy Brost

Electronic Media Review, Volume Six: 2019-2020

Full text PDF

ABSTRACT

In "A Case for Digital Conservation Repositories," Barbra Mack and Glenn Wharton outlined an approach to documentation and information management for time-based media artworks. Their model aimed to comprehensively organize

the characteristics and dependencies of these works in a component-based, non-hierarchical repository ecosystem. Wharton, then time-based media conservator at the Museum of Modern Art in New York, was leading research and development of what would become the museum's Digital Repository for Museum Collections. This article provides an overview of the decade that followed, including the workflows and technologies employed, and the growing pains during implementation. For collections of digital art objects, this project demonstrates the value of taking the approach that Mack and Wharton outlined, which is technology agnostic but involves understanding and embedding the status of components according to artist-defined values, implementing cataloging policies, arranging and describing artworks at the component level, exposing technical metadata, and defining relationships between components. What comes to the fore in this history is not the story of technical challenges alone but also the more holistic story of how the David Booth Conservation Department at the Museum of Modern Art ultimately spearheaded the development of a cross-departmental team to collaborate on the ongoing development of the Digital Repository for Museum Collections. In this way, dialogues with the museum's leaders and specialists in information technology infrastructure, applications, database administration, collection management systems, and digital asset management led to practical, sustainable solutions that involve and benefit stakeholders in the care of digital material across the museum.

Introduction

Few art museums have such a large collection of digital art objects as the Museum of Modern Art (MoMA) in New York. The collection comprises many thousands of digital files from more than 3,000 artworks across curatorial departments. The largest amount of material in terms of total size falls within the Film Department, which collects born-digital films and also stores digital files created from

scanning and digitally restoring chemical film elements. Out of necessity, MoMA has been a leader in developing digital preservation storage and management practices specifically designed to meet the needs of an art museum.

More than 10 years ago, when the collection was a small fraction of the size it is today, MoMA began taking steps to assess the unique needs of these high-value digital materials. Taking a holistic approach, the museum started to develop policies and procedures for acquiring, handling, cataloging, documenting, and displaying time-based media artworks. As part of this comprehensive effort, the museum also began to research safe storage practices. This article will focus on the development of those storage practices, and specifically, the development of MoMA's Digital Repository for Museum Collections (DRMC). It has been a decade since MoMA provided its initial report conceptualizing a "DRMC" to the *Matters in Media Art* project team. Passing this milestone provides an opportunity to reflect on the development of the DRMC and its future.

Background

MoMA began collecting digital media in 1991 and subsequently amassed diverse digital storage media from acquisitions. In 2005, Jim Coddington, then chief conservator at MoMA, tapped objects conservator Glenn Wharton to conduct a survey of the time-based media artworks in the collection. There were 62 works with digital media, with a footprint of 500 GB. In 2007, Wharton became the first time-based media conservator at MoMA. This was a long time ago in terms of digital preservation tools and practices. Technologies have evolved rapidly since then: in 2007, the first iPhone was released, and Netflix introduced streaming video services. There were advances in the digital preservation field around that time as well. Artefactual Systems first started development on its Archivematica software with funding from the UNESCO Memory of the World Subcommittee on

Technology in 2007. This new software was to be based on the Open Archival Information System (OAIS) functional model that had been released as ISO 14721 in 2003. The subcommittee had just written a report that defined the open-source software requirements to implement a digital archival and preservation system that would align with OAIS (Van Garderen et al. 2012). As this article will show, throughout the life of the DRMC, contemporaneous developments within and beyond the museum would continually inform the rapid evolution of the museum's digital storage ecosystem.

In 2008, MoMA launched an interdepartmental research project to investigate models for safe server storage of its digital art collections, spearheaded by Coddington and Wharton in the Conservation Department. It concluded that these collections were at high risk of loss. Artworks that were wholly or partially digital were being collected across many departments, including Media & Performance Art, Architecture and Design, Paintings and Sculpture, Photography, and Prints and Illustrated Books. In the meantime, the collection was growing; there were more than 100 works with digital elements at the time. They ranged from TIFF files for still images to complex installations with artist-written code that depended on specific software and hardware to run. All were still on the original hard drives, computers, discs, and other storage media on which they were delivered to the museum at the point of acquisition. Representatives from conservation, curatorial, collections management, and the Information Technology (IT) Department held a series of meetings to discuss the problem of digital media in the collections. This group correctly identified computer-based art as one of the most fragile types of digital art in the collection. There were no digital repositories for artworks in existence, and only several certifiable digital repositories in libraries and archives. One was located at New York University (NYU) Libraries and Information Technology Services, which had developed a digital repository for library special collections and was working on a project to develop a repository for digital public television (Rubin 2010, 25–26). MoMA

reached out to the team at NYU to learn more about what these projects entailed.

In 2009, MoMA began working on a proposal to obtain grant funding for the Digital Collections Conservation Repository (DCCR). It would be a phased project that would allow the museum to perform further research on its needs and then build a digital repository specifically designed for storage of digital elements of artworks. Following a series of information-gathering meetings with the NYU team, which included members from NYU Libraries, IT, Computer Science, and NYU's Moving Image Archiving and Preservation (MIAP) graduate program, MoMA sketched out a short proposal for the DCCR that included high-level functionality, staffing needs, and approximate budget (Wharton 2009). The museum hoped that this repository would also be a model that could be followed by other museums, creating a community of practice around responsibly storing and preserving technology-based art. Additionally at this time, NYU partnered with MoMA to form a working group for the Conservation of Computer-based Art. Its leaders were Glenn Wharton for MoMA and NYU professors Deena Engel, Howard Besser, and Mona Jimenez, with funding provided by the university's Visual Arts Initiative (Wharton and Engel 2015, 113). The case studies performed by the working group helped elucidate the documentation needs and preservation risks associated with some of the collection's most complex and fragile works (Engel and Wharton 2014). That acquired understanding helped shape the unique functional requirements for the museum's digital repository.

Additionally in 2009, Glenn Wharton and Barbra Mack, portfolio manager and systems analyst for NYU Information Technology Services, presented their paper "A Case for Digital Conservation Repositories" at the AIC Annual Meeting in Los Angeles (Wharton and Mack 2012). In this important paper, the authors articulated the value of *component-based*, rather than object-focused, data models for digital repositories for conservation. They explained that object-based approaches structure metadata hierarchically, placing the artwork at the "parent"

level and adding technical and descriptive metadata below. However, they showed how the relationships and dependencies within software-based artworks make this approach problematic because "multi-component digital works of art employ an abstract IT system that cannot be understood hierarchically" (Wharton and Mack 2012, 33). Put another way, "The components have distinct roles in maintaining the entire system. Like a human body in which organs such as the heart and lungs are cared for in unique ways, each component is a member of a whole functioning system. Therefore, one cannot readily understand . . . relationships using the hierarchical concepts applied to still or moving image art (e.g., video) such as *siblings* and *derivatives*" (Wharton and Mack 2012, 26). Component-based models allow for granular description and management of diverse digital objects comprising a single artwork.

Component-based approaches were being advanced in other contexts as well. Tate had previously adapted its component-based model for analog media to digital elements, and in subsequent years, it became clear that this approach was well suited for packaging, describing, and storing digital material (Roeck 2016, 24–26). Another contemporaneous model, the Documentation Model created by the DOCAM project (Documentation and Conservation of Media Arts Heritage), also emphasized the importance of having a "component" level. In fact, the model is based on the Functional Requirements for Bibliographic Records model with "Component" added below "Item." As it states on the project website, "components are at the very heart of the changes affecting most media artworks. The addition of this level promotes the identification and collection of documents that make reference to a specific component in the item, which in turn facilitates the tracking of changes made to the work throughout its lifecycle" (DOCAM Research Alliance 2010a). In the DOCAM online Glossaurus, a "component" is defined as "any physical or logical part of an artwork from which particular characteristics, behaviours or functions can be identified" (DOCAM Research Alliance 2010b).

The superiority of component-based data models for time-based media artworks is a foundational concept in the conservation field. Wharton and Mack emphasized two additional foundational concepts in their paper: recording the *artist-assigned value* of each component, as well as the *relationships* between components. When recording the status of each component in the work—from artist-created elements to replaceable ones—it is particularly important to capture artist-assigned values, which in turn inform conservation decision making.[1] Wharton and Mack wrote, "Clearly it is not just the IT system relationships, nor the technology variations that challenge an object-based focus for conservation. Individual components may be assigned different values by the artists. This requires a shift towards applying different conservation strategies to individual components, instead of a single strategy that is holistically applied to the work of art" (Wharton and Mack 2012, 27–28). Also key are the relationships between components, especially non-hierarchical relationships such as diverse dependencies. Tate has long documented the genealogy of media works using color-coded "media production diagrams" that show the parent/child relationships and status of related media, usually masters and various derivatives (Roeck 2016, 24–26). This concept was gradually expanded to include the entire web of relationships between components of an artwork.[2] In their paper, Wharton and Mack used a work in MoMA's collection as a case study to demonstrate that it is necessary to document this complex web of relationships for conservators to truly understand how to care for the work.

Finally, Wharton and Mack emphasized the importance of robust component-level metadata because individual digital components of the same work may each have different preservation risks. They suggested an expanded PREMIS schema and showed how it could be applied to computer-based art. However, they noted that the specific schema chosen was less important than adopting a component-focused model and generating robust technical, descriptive, and preservation metadata. Practices at MoMA today conceptually align with all of these early

ideas.

In 2010, digital components of the approximately 230 works in the collection were being minimally catalogued in MoMA's collection management system, TMS (The Museum System), to signal the existence of digital elements to TMS users. They were also backed up to a server designated by IT for artwork storage. Working with AudioVisual Preservation Solutions (AVPS), MoMA generated MD5 and SHA-1 checksums for these 250,000 files and extracted technical metadata for them using a Technical Metadata Processor Script by archivist/technologist Dave Rice. This data was stored in spreadsheets intended for further analysis. Additionally during this year, the influential paper "Digital Forensics and Born-Digital Content in Cultural Heritage Collections" demonstrated how digital forensics concepts and practices could be adapted to cultural heritage contexts to establish and maintain chain of custody for digital acquisitions (Kirschenbaum, Ovenden, and Redwine 2010). This would ultimately lead to the adoption of write-blockers, disk imaging, and other forensic acquisition techniques by time-based media conservators at MoMA and other institutions.

The First 10 Years of the DRMC

By 2010, MoMA had been a part of the *Matters in Media Art* project for about six years. Funded by the New Art Trust, the project gathered colleagues from MoMA, San Francisco Museum of Modern Art (SFMOMA), and Tate to collaboratively develop guidelines for the care of time-based media artworks. The first phase focused on loans and launched in 2005. The second phase, on acquiring time-based media art, launched in 2007. After that, the partners shifted their focus to digital artworks, and each institution worked on a separate project. In 2010, the partners met to share their work. The report on repositories for digital artworks was the MoMA project that would become the DRMC (Smith 2020, 45). The

authors of the report, titled "Design for a Digital Repository for Museum Collections" (Wharton et al. 2010), were primarily Glenn Wharton and Kara Van Malssen, then senior research scholar at NYU working on the Preserving Digital Public Television project (Rubin 2010, 25), but there were many voices and contributors. The risk assessment reports were written by Engel and Wharton. The Media Working Group at MoMA served to organize and implement the project, and funding was made possible through the New Art Trust as part of *Matters in Media Art*.

In 2011–2012, the museum prepared for an ambitious exhibition called *MoMA Media Lounge*, developed by Sabine Breitwieser, then chief curator of Media and Performance Art. The exhibition ran from 2012 to 2013 and was a modular display with multiple viewing stations that allowed visitors to select and view videos from the Media and Performance collection of video art. The digitization of hundreds of videotapes to prepare for this exhibition was performed by contract conservator Peter Oleksik and the vendor DuArt (Oleksik 2015). This project added 30 terabytes and hundreds more works to the museum's existing digital collection of 280 works occupying 12 TB. During the project, Oleksik permanently joined the staff as assistant media conservator. The digitization project added to the urgency to develop a storage system for the museum's valuable digital collection materials.

By 2012, the museum had nearly completed its plan to develop this unique storage system. Recognizing that it was the first art museum to venture into these waters, MoMA convened a small group of external advisors to meet with key staff, including representatives from the conservation, IT, collections management, registration, and TMS and digital asset management system (DAM) teams. Among the participants were consultants Chris Lacinak and Kara Van Malssen from AVPS, as well as Stephen Abrams from California Digital Library, Howard Besser from the Moving Image Archiving and Preservation graduate program at NYU, Ben Fino-Radin from Rhizome, Hannah Frost from Stanford University Digital Library

Systems & Services, Jerome MacDonough from the University of Illinois "Preserving Virtual Worlds" project, and David Millman from NYU Digital Library Technology Services. This was the first of several small-group meetings funded by MoMA to ensure that the museum leveraged a wide range of expertise during the planning process. The external participants varied. Numbering more than two dozen in total, they included time-based media conservators and leaders in the digital preservation field who had experience in one way or another with the kinds of challenges MoMA was facing. At several key stages, these advisors provided invaluable feedback on proposed system requirements and software development.

Soon after the first advisors meeting in 2012, AVPS and MoMA, led by Wharton, completed the DRMC system requirements document. It described a system that would perform a number of functional requirements in accordance with the Reference Model for an OAIS, ISO 14721:2003. These included definition of Submission Information Packages (SIPs) and Archival Information Packages (AIPs), validation and ingest of packages, storage and integrity management, indexing and data management, long-term preservation planning, and providing access to file-based artworks for authenticated users and external environments. The DRMC would perform these services on an ongoing, managed basis to ensure the longevity of digital artworks. The document stated, "Primary features of the DRMC will include a *storage system* that reliably maintains digital files with maximum integrity, a *database* that houses detailed technical and structural metadata required for the conservation and presentation of digitally-based artworks and their operating platforms, and a *workflow management system* that facilitates valid ingest, conservation, and access procedures. The DRMC will be a highly secure and specialized infrastructure that integrates with broader museum collection management policies, procedures, and system" (emphasis added). The stated aim of the DRMC was "to enable effective installation of all digital artworks over time" (MoMA and AVPS 2012, 4). At the same time, the authors acknowledged that "the problem of software-based art preservation will not be

solved immediately; rather, the aim is to take adequate preparatory measures so that as tools and strategies to address this challenge evolve, MoMA will be ready to implement those" (MoMA and AVPS 2012, 5). As of 2020, MoMA is still in the early stages in its care of software-based artworks.[3] The DRMC functional requirements as devised in 2012 and their status in 2020 are summarized in the appendix.

The museum began to build the DRMC under the leadership of Glenn Wharton, with the support of then chief conservator Jim Coddington, Chief Technology Officer Juan Montes, IT Director James Heck, and Database Administrator Steven Moore, as well as Chris Lacinak and Kara van Malssen from AVPS. In this period, Ben Fino-Radin joined MoMA as a DRMC manager. Up to this point, although the museum had started storing copies of digital files on the server to achieve redundancy, checksums and metadata were not packaged with the files in any way. Fino-Radin spearheaded the implementation of digital forensics workflows, including the use of disk imaging and write-blockers, and the use of BagIt, the file packaging tool from the Library of Congress, enabling the museum to package files with their checksums. In this way, digital files could be safely "bagged" from artist-provided carriers, transferred to the dedicated art storage server set up by the museum's IT department, and validated to ensure their complete and correct transfer.

In 2013, Wharton left MoMA to accept a professorship at NYU, and Kate Lewis, a time-based media conservator from Tate, joined the museum. At this time, it was apparent that MoMA had a large and rapidly growing collection that necessitated automating the creation of AIPs for long-term digital storage. Fino-Radin worked with IT and Artefactual Systems on the implementation of a customized version of Archivematica that would run on MoMA servers in a virtual machine, and leverage automation tools to ingest both bagged and unbagged material. With the addition of a server at a second location in the city, MoMA then had two copies of

the collection in two different geographic locations (two EMC VNX5500 arrays for primary and replicated collections storage). Then, a specification was developed for a management application layer to support search, discovery, access, and conservation planning. This new set of requirements was issued as a part of a request for proposals for vendors that could develop the DRMC Management Application. Artefactual Systems, the developer of Archivematica, was selected to build this custom software, which would be based on the company's AtoM (Access to Memory) product.

This management application was later called *Binder*, and it launched in 2014. The conceptual work done on Binder was essential for developing an understanding of what management tasks were required to be performed in a repository ecosystem for art objects. There had been several meetings of internal and external experts at which the functions of the management application were refined and tested. In 2014, those tasks were bundled and performed by the custom software application. The full functional requirements are listed in the appendix. Broadly speaking, the functions supported the following collection care objectives (Van Malssen and Fino-Radin 2013, 12–13):

1. *Description*: Provide access to granular artwork descriptions and an interface for users to input new metadata about digital collections.
2. *Search & Retrieval*: Facilitate discovery through robust search and browse features.
3. *Versioning*: Facilitate user management of AIP versions when new components or artwork versions are added to the DRMC.
4. *Fixity*: Audit and report on the integrity of files stored in the DRMC.
5. *Access*: Enable limited user access to content files stored in AIPs.
6. *Administration*: Provide tools for configuration of the management application (Binder), including user and vocabulary management.
7. *Usability*: Support ease of use and utilize contemporary user experience design.

8. *Conservation*: Considered a future phase, the management application (Binder) would ultimately utilize new modules and web services to support conservation planning.

In addition to the launch of Binder, the other notable change in 2014 was the anticipated receipt of half a petabyte (500 TB) of new film scan files from the Film Department's preservation work, to add to the overall collection size of 75 TB. The Film Department also anticipated growing their digital footprint at a rate of about 40 TB per year. The collection size was ballooning, and acquisitions of digital elements were rising across the board, with about 100 new works coming in per year. In light of this, storage discussions came to the fore. Although there were two copies of the data, the second copy was created by mirroring the Manhattan art server, which was not ideal for preservation purposes. In addition, the projected costs of keeping this collection on fast server storage were prohibitive. The decision was made to change from a server-based to an LTO-6 tape-based storage system, and to have that storage located on-site like the rest of the collection.

In 2015, I, the author, started at MoMA as a graduate intern, and in 2016, I became the first Andrew W. Mellon Fellow in Media Conservation. The on-site LTO-6 tape storage system was implemented at the end of 2015 by a vendor named *Arkivum*. The museum then began migrating the AIPs stored on the servers to the new tape storage system, as well as continuing with new ingests. In all of 2016, about 10 TB were ingested. Although this was not a large amount of data, it represented many artworks. Additionally in 2016, there were some significant staff changes. The museum hired a registrar experienced with time-based media collections, Coddington retired, and soon thereafter, Lewis succeeded him as chief conservator.

At this point, the museum had a fully operational, custom-built DRMC situated in

the Conservation Department to house the digital elements of artworks. It consisted of a staging area on the museum's servers for bagged SIPs and an on-site customized Archivematica instance with multiple automated transfer and ingest pipelines to convert those SIPs to AIPs, which were transferred to an on-site LTO-6 tape-based storage system managed by the storage vendor Arkivum. The AIP was written to three tapes, two online and one offline, stored in three separate geographic locations. Using Binder, time-based media conservators could see metadata and relationships, preview files, trigger retrieval, and search for AIPs and their contents. Each AIP corresponded to a component that was cataloged in the museum's collection management system (TMS) with its status, relationships, and provenance. The museum could now "road test" the system end-to-end. The incremental deployments had been largely successful, but there would be a few bumps in the road that could not have been foreseen. At the time of Fino-Radin's departure from the museum, there was an unresolved issue with the storage vendor system: a large amount of data was in a waiting state to be written to LTO tape, and its footprint was slowly growing. This mystery would not be solved until later.

Working day to day with the system, the bottlenecks started to come into focus. The cataloging and condition checking required for each component were, and remain, time intensive, so even though technically the system could theoretically ingest 1 TB per day, it was difficult to have the staffing capacity to keep the pipeline filled with material ready to be stored. This had the immediate effect of exacerbating the backlog. For example, it is possible to prepare four 250-GB disk images to be ingested into storage on one day, but much more time consuming to catalog, bag, and condition check a few hundred MP3 files that might add up to almost nothing in terms of total size. Just achieving a pace where the amount of digital material being deposited into storage matched or exceeded the amount being acquired or generated proved impossible. Another bottleneck was the comparatively slow performance of our relatively outdated custom version of

Archivematica, and the custom user inputs that it required. A staff member had to sit and run Archivematica to instruct it to send a file to Binder or to the DAM, and to review and store an AIP. And there were other issues: almost immediately we realized that raw DPX, the film scan file format we were receiving from the Film Department's preservation work, was taxing the software, taking literally days for Archivematica and Binder to process one AIP, sometimes failing during the process and causing hours of work to clean up and restart. Software troubleshooting on Archivematica and Binder were time consuming for staff.

In 2017, as Peter Oleksik, associate media conservator, and I continued to run the system following the departure of Fino-Radin, it was increasingly clear that the issue with data waiting to be written to LTO tape would not resolve itself with time, as we and our vendor partners had hoped. We know now that, unfortunately, an intermittent error was causing some AIPs not to fully transfer from our servers into the Arkivum storage system. For these AIPs, Archivematica showed no errors or failures because the software had created and handed off a valid AIP, and the Arkivum system showed a waiting state, as its hard disk waited to receive the full contents of the new AIP. Similarly, troubleshooting by our IT department revealed no network issues that could explain the problem. As such, it was some time before we were able to identify and diagnose the issue. The cause was a software integration problem, and, at the time, the responsibility for that aspect was situated in the Conservation Department by default. This has since been remedied by contracting with the vendors to take on those integration responsibilities. In the repair phase, to identify the affected AIPs, we needed to work with our storage vendor to manually review the storage status of more than 1,000 AIPs. Fortunately, because nothing is ever deleted from our servers until it is confirmed stored on three LTO tapes, the museum did not experience any data loss. New ingests into the repository were delayed for more than a year, and as a result, Conservation filled up 100 TB on museum servers and nearly 20 TB more on desktop RAID drives as we tried to keep up with the pace of new acquisitions. All of

this data was waiting to be transferred into the LTO tape system. The IT Infrastructure team, especially Senior Manager John DuFour and Technology Engineer Sergiy Petrychuk, worked closely with the Conservation Department throughout this challenging period to ensure that there was sufficient temporary storage space on the museum's servers to accommodate all of the artworks awaiting ingest.

The experience laid bare a few problems. First, that there was not adequate staff capacity to keep up with the cataloging, pre-ingest, and condition-checking workflow that had been developed. Some cataloging requirements were streamlined or eliminated, with the goal of backfilling them by importing AIP metadata automatically generated by Archivematica. Condition-checking procedures were reexamined and a tiered approach was applied that would set clear objectives and replace the one-size-fits-all condition report. Second, the conservation team was running a custom, purpose-built digital preservation storage system largely independent of the rest of the museum. With digital material representing a small fraction of the museum's total collection, it became clear that the system needed to integrate more substantially with existing systems, even if those systems were not purpose built for the task. It would require creativity, interdepartmental collaboration, and compromise to achieve the same functions as the system could in isolation. However, it would also afford new opportunities.

The following two-year period (2018–2019) proved to be a good one for the DRMC. The collection size was roughly 3,000 works and more than 600 TB. We were able to make additional refinements and realize efficiencies. In 2018, leveraging one of our small-group meetings as part of the Andrew W. Mellon Media Conservation Initiative, we focused on the problems of film scanning and digital preservation, resulting in the decision to use the open-source application RAWcooked by MediaArea (https://mediaarea.net/RAWcooked) to losslessly

package DPX for storage, which reduces the size by one- to two-thirds and also processes faster. Coupled with changes in curation and scanning project deliverables, we were able to reduce the current and projected digital storage footprint of Film Department material to a more manageable level. That year, the storage vendor upgraded our software so the conservation team had a new interface for browsing AIPs and triggering retrieval. Sarah Gentile started at MoMA as an assistant digital preservation specialist to focus on digital artwork storage, including cataloging, forensic acquisition from carriers, and retrieval.

In 2019, the museum upgraded to a mainline, non-customized version of Archivematica. The museum also discontinued its use of Binder. These two changes freed the conservation team from custom software development and operation to focus on the pre-ingest tasks required to prepare digital material for long-term storage. Moving to the mainline version of Archivematica will simplify software support and allow the museum to participate more fully in the user community for the mainline product. As for Binder, it had been a visionary attempt to pioneer an all-in-one software solution for management of a time-based media art collection but would have required more community adoption to truly sustain it. What was not known at the time of Binder's development was that other museums would ultimately adapt to accommodate the needs of time-based media artworks in such diverse ways that widespread use of one management application was unlikely. This open-source application remains freely available on GitHub (https://github.com/artefactual/binder) for others who wish to use it or build upon it, or learn from the tasks it was designed to perform.

The DRMC presently consists of the following (fig. 1): SIPs in bag format awaiting ingest on MoMA's servers, Archivematica running in a virtual machine in MoMA's data center to process SIPs into AIPs, the DAM (NetX) to preview works, Arkivum software to write and manage the LTO-6 tape libraries in on-site IBM TS4500 tape appliances, and the Arkivum graphical user interface for AIP retrieval.

Arkivum also performs fixity checks annually. MoMA uses TMS as a management application to store not only component cataloging, status, relationships, and provenance but also the AIP technical metadata, location, a Universally Unique Identifier (UUID), and preservation actions.

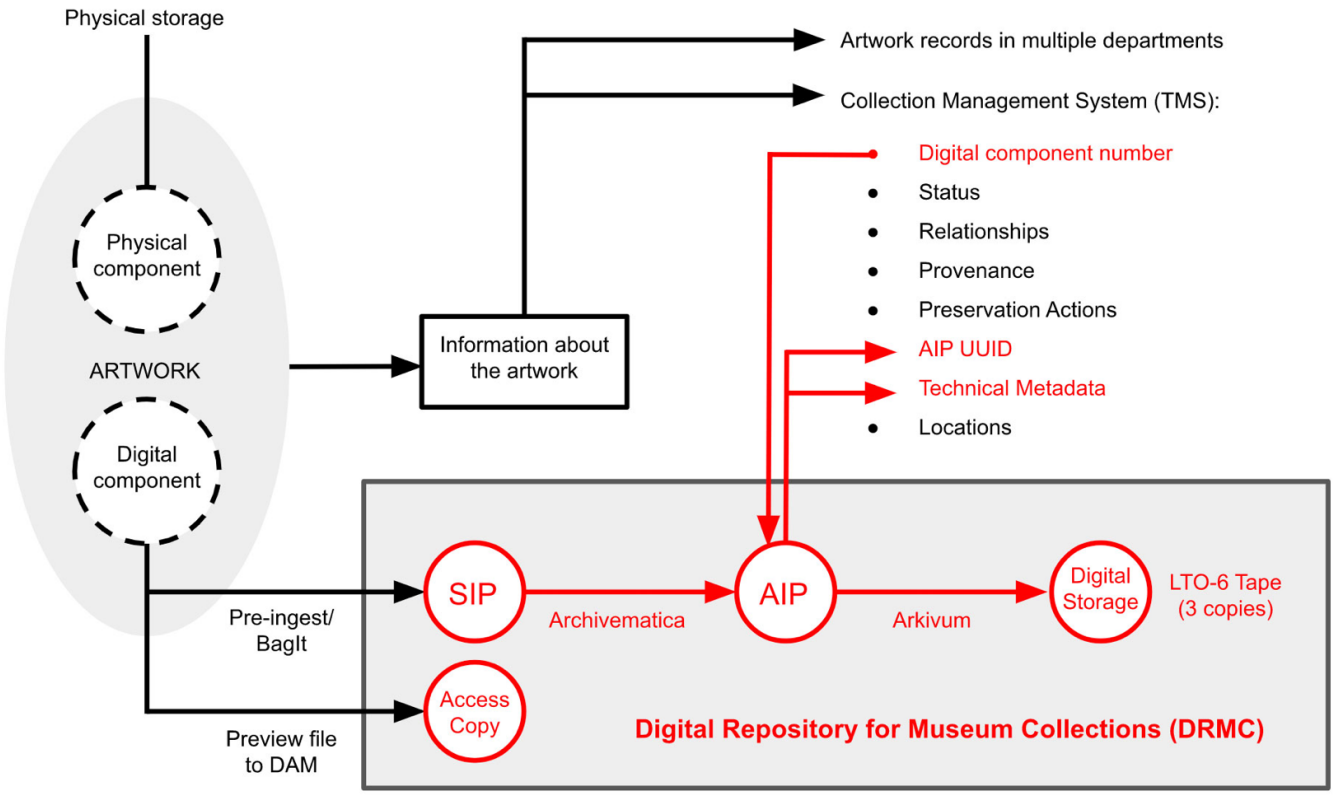


Fig. 1. MoMA DRMC, 2020

Of the approximately 600 TB in the collection, MoMA has stored more than 100 TB in the DRMC with roughly 500 TB being gradually ingested at the rate of up to 1 TB per day, while both acquisitions and film scanning continue apace. The projected size of the collection in a few short years is as much as 1.2 petabytes.

As described earlier, developing the management application Binder led the way in determining what management tasks were required for a collection of digital art objects. However, as the TMS and DAM teams at MoMA evolved over the years, it became clear that the conservation team could partner with them in new ways

to perform many of the same tasks that Binder did. This also widened the team that was doing preservation at MoMA and made the works in the DRMC—and information about them—more visible to more staff. The appendix lists Binder's full functional requirements and shows the systems performing those tasks in 2020.

Using TMS as a management application is, in some ways, like fitting a square peg into a round hole, but it has numerous advantages. When the DRMC was first launched, the museum's version of TMS and its limitations were such that what the Conservation Department could do with digital components in that system was quite limited. However, eventually it was decided that time-based media conservators could serve as catalogers for digital material, whereas the registrars would continue to catalog physical materials. Registrars did not have capacity for the increased cataloging at the time, so time-based media conservators were permitted to begin creating and describing components in TMS, using procedures cooperatively developed with the registrars and TMS team. When the TMS software was upgraded to the version TMS 2012, the conversation about entering AIP data into TMS was restarted. The TMS and IT teams were supportive of expanding the information in TMS about digital components. This also coincided with the migration of the Film Department's previously separate collection database to TMS. With curatorial departments across the museum collaborating as never before, using one system for object information became ever more crucial. Under the leadership of Chief Technology Officer Diana Pan and IT Director Helynsia Brown, the Enterprise Applications team began working regularly with the Conservation Department and TMS lead Ian Eckert, associate director of Collection and Exhibition Information, to enhance the usability of TMS as a management application for the DRMC. Steven Moore, developer and database administrator, set up various data imports from the Arkivum storage system into TMS. The following information is imported:

- *AIP UUID*: Archivematica generates a UUID for each AIP. The UUID is entered into TMS. No matter where the AIP is stored (e.g., server, LTO tape), it can be located by this UUID. These UUIDs correspond to a single component and a single AIP. This UUID is pulled into TMS automatically once the AIP is successfully stored on LTO tape and can be found at the component level in TMS.
- *AIP location*: Because there are multiple copies of AIPs, there will always be multiple locations. During the pre-ingest process, locations are granularly described in TMS (e.g., individual desktop RAID drives or specific staging servers). Once three copies of the data are stored on LTO tape, the location is "DRMC." Then, to add more precision, the LTO-6 tape numbers where each AIP can be found are also entered into TMS. These are the numbers of the tapes in MoMA's on-site tape libraries. MoMA owns the physical tapes as part of its on-site storage system, so even as media are migrated, these tapes will remain in the collection.
- *AIP technical metadata*: AIP technical metadata enables conservators to see important information about a digital component without retrieving it from the repository. For example, a video file does not have to be downloaded for a conservator to learn its codecs, aspect ratio, resolution, audio channels, bit rate, and so on to plan for an installation. This data is imported into TMS in two ways: one for readability and one for searchability. The entire METS file generated by Archivematica during AIP creation is kept on MoMA's servers, ready to be displayed in a human-friendly XML viewer developed in-house. This application, inspired by Tessa Walsh's METSFlask (https://github.com/tw4l/METSFlask), is called *Let's Go METS!* and was built by Ryan Sprott at MoMA on Ruby on Rails. There is a URL at the component level in TMS to display the AIP METS XML file in the METS viewer. The entire contents of the METS file is also pulled automatically into a Text Entry at the component level, so specific terms can be searched by TMS users and the TMS team through database queries.

Over the years, a number of expanded metadata schemas have been designed for MoMA for the DRMC, including by Barbra Mack, Chris Lacinak and Sarah Resnick, and Peggy Griesinger. Mack put forth an expanded PREMIS schema in 2009. Lacinak and Resnick worked on a metadata model and dictionary for the 2010 DRMC proposal. The latest data model by MoMA and AVPS was included in the DRMC system requirements document, and was influenced by the PREMIS and PBCore standards, the KEEP TOTEM project (Keeping Emulation Environments Portable project's Trustworthy Online Technical Environment Metadata registry), as well as research done by Mack (MoMa and AVPS 2012, 49–50). Griesinger, who held a National Digital Stewardship Residency at MoMA, proposed a metadata schema in 2016 for digitization and format migration that leveraged four different standards (Griesinger 2016). All of these models were essentially adapted from schemas into narrative form, due both to staff capacity constraints and a desire to create the most accessible, readable documentation and store it within our collection management system. Conceptually, information is captured as suggested in these schemas, but it is not bundled into the AIP metadata. The AIP contains only the metadata from Archivematica processing and the TMS identifiers and component number so that the narratives can be discovered in TMS. This is discussed in more detail in the next section.

In addition to having the AIP location and technical metadata widely available via TMS, the goal is to also have a preview file that is representative of the artwork available in the DAM (NetX) for curatorial review and research purposes. For some artworks, such as software-based installations, there is no preview file that can adequately represent the artwork in the DAM, so conservators in those cases provide the information and materials on an as-needed basis. Because departments across the museum are being encouraged to centralize their audiovisual materials in the DAM, the conservation team now has a NetX uploader tool, managed by the DAM team, to add preview files. The selection and preparation of preview files is done by the conservation team in collaboration

with the curatorial departments. No exhibition files are stored in the DAM, and the DRMC only contains digital materials that are considered artworks, including the exhibition files that are actually shown in the gallery. There is no documentation stored in the DRMC. Instead, this is stored on the museum's servers or entered into the applicable object records in TMS. With an LTO-tape-based system, editing AIP metadata on three geographically separated tapes is not feasible as part of a routine workflow. Therefore, AIPs for a particular artwork can accrue, but not be edited, although they can be deleted if necessary with significant effort.

Key Learnings

After this survey of the history and present state of the DRMC, it is useful to consider what the key learnings were over the past 10 years. Because MoMA is unique in terms of the size and composition of its digital collection, and because it was the first art museum to implement an on-site digital preservation repository, there were bound to be aspects that were surprising. In this author's view, the most surprising might be the degree to which contemporaneous developments within and outside the museum impacted the repository's evolution with ongoing frequency. These ranged from availability of new technologies and software to staffing changes and shifting internal mandates across departments. Although the strategic goals and business case remained largely consistent over time, the tactics, procedures, tools, and the roles of staff and vendors have undergone a process of continual testing, refinement, and change, and this is expected to continue. One way to think of the DRMC project is as the museum's effort to plan, build, and deploy a preservation system in which to store its digital collection. However, such a view suggests an endpoint, when in fact this process is ongoing. Moreover, that description is incomplete. The other, equally important part of the project was that the museum built a team to *do preservation*, and empowered that

team to flexibly and organically evolve its methods as the needs of the collection and the organization evolved.

Numerous staff at MoMA saw their roles shift to accommodate the needs of time-based media artworks. Some changed enthusiastically and some hesitantly. MoMA learned that it takes sustained commitment and advocacy to continually transform the institution to care for time-based media artworks. A 2018 paper by Vivian van Saaze, Glenn Wharton, and Leah Reisman titled "Adaptive Institutional Change: Managing Digital Works at the Museum of Modern Art" looked at these developments through the lenses of infrastructure analysis, sociomateriality, and institutional theory from organizational sociology (Van Saaze, Wharton, and Reisman 2018). They found that the challenges presented by caring for these artworks brought about a series of adaptations rather than radical changes within the museum. Their conversations with staff involved in this process of adaptive change suggested that this was possible because MoMA sees itself as a "change agent" with a "commitment . . . to innovate," and because of its financial and organizational resources (Van Saaze et al. 2018, 233). They identified "institutional entrepreneurs" among the organizational leaders whose "smoothing work" helped to "facilitate integration of new kinds of work in prior modes of practice" (Van Saaze et al. 2018, 226, 233). As Wharton explained, "I was surprised by the analysis of institutional change that we conducted. I went into the study thinking that radical change was happening all around me with the collection and management of digital art. Sociological theory on institutional change and Science and Technology Studies theory helped me understand that what we were experiencing was a process of institutional adaption—not disruption" (Wharton, pers. comm., September 20, 2020). The evolution of staff roles is an example of this process of adaption. As described in the chronology in this article, the DRMC began with a small launch team that subsequently grew to include more titles and leverage more internal expertise. Responsibility matrices at the time of the DRMC launch and in 2020 are presented in the appendix. As

these matrices show, roles and responsibilities became more defined and also more widely spread across the museum.

The changes required of staff members particularly impacted the IT department, and MoMA learned that working interdepartmentally involves overcoming fundamental communication issues through mutual patience and persistence. Digital preservation is a relatively young discipline practiced within a range of institutions that have valuable digitized and born-digital data designated for long-term—even permanent—retention. In an art museum context, this is not just data that needs to be permanently stored. It must also be permanently accessible, renderable, and usable as part of the daily work of the institution. What conservators are asking of the museum's data management team is an enormous shift that essentially makes the conservation and IT departments jointly responsible for the long-term preservation of high-value digital data, a role traditionally performed in archives departments. Fortunately, museum IT departments are expert at safeguarding critical data, including personnel, development, membership, retail sales, and other data types. Although there may be a significant communication gap to bridge (Prater 2018), once the unique requirements for digital preservation storage are understood by IT, and once conservators understand how their data storage needs fit into the overall data management landscape of the museum, conservation and IT departments can form a strong partnership. Similarly, decades ago, new interdepartmental partnerships were forged when conservators and building operations engineers began to collaborate on the monitoring and management of the museum environment. In both cases, departments performing quite different functions in the museum were able to partner effectively and evolve practices within the institution to improve collection care.

MoMA's key learnings about the technical aspects of implementing digital preservation storage may be best summarized through co-emergent learnings

from across the community of practitioners dealing with similar challenges during the same period. This includes the practitioners who formed the National Digital Stewardship Alliance in 2010, who authored and continue to refine the "Levels of Preservation" to help institutions advance their digital preservation efforts along a user-friendly spectrum. In 2016, the first year that the DRMC was fully implemented, the community-driven "Digital Preservation Storage Criteria" were first published by others also grappling with how to build preservation storage systems. And much of what the MoMA team learned in the development of the DRMC was written into the latest phase of the *Matters in Media Art* project, "Sustaining Media Art." This phase, which launched in 2016, added a "Digital Preservation" section to the web resource, collaboratively written by MoMA, Tate, and SFMOMA (Smith 2020, 48). Any institution approaching digital preservation storage planning and implementation today is able to leverage all of these learnings of the past decade—MoMA's included.

Last, there were significant lessons learned regarding the unique needs of artworks, and how they impact decisions around digital preservation and storage. One aspect of the DRMC that provides a rich basis for discussion is one that was never fully implemented, and that is the data model. As described earlier, MoMA decided to adapt the metadata models that would have added information to the AIPs stored in the LTO-tape storage system to narrative documentation stored in TMS instead. First, it is important to emphasize that the artwork itself, the actual preservation entity of concern to conservators, is not a record *per se*. Conservators manage the complete artwork record to conserve these unique artworks that can only be fully experienced as artworks when installed. The original design called for the DRMC to be not only secure storage for the digital elements of artworks but also a data management and conservation planning system. The AIP would consist of the artwork component and rich metadata, and selected information would be copied to other systems. However, many of these artworks are hybrids of digital and physical media, complicating the boundaries of the DRMC. The DRMC

system requirements document attempted to define those boundaries at the start, explaining that the following were to be excluded: audiovisual documentation of the artwork as installed in the gallery, "extensive" documentation about non-digital components of the work, and any dedicated hardware "not required by any of the files or executable aspects of the work" (MoMA and AVPS 2012, 9). Since then, the boundaries have been redefined by necessity.

At MoMA, there are practical reasons for privileging narrative documentation in TMS over richer metadata in the AIP. Boundaries within the museum are falling. The collection galleries at MoMA are no longer divided into territories belonging to individual curatorial departments. Conservators collaborate on acquisitions with greater frequency, as multimedia installations may cross any or all of our specializations in paper, photography, painting, sculpture, and time-based media. The artwork records on the museum's servers and in TMS require the Conservation Department to develop shared approaches so that collaboration within and beyond the department can be seamless. Information storage in the DRMC is shaped by this need to ensure that knowledge about an artwork is created and centralized in a consistent manner, and accessible and discoverable across the museum. To achieve this, in MoMA's workflow, descriptive information and preservation actions taken on the digital collection objects are documented narratively in TMS rather than in a metadata schema within the AIP. Each AIP in the DRMC contains only the technical metadata generated by Archivematica in the ingest process, and identifiers to associate the AIP with the proper TMS component. In this way, the AIP points back to the system of record that contains the rich narrative information about the artwork, along with that component's status, relationships, provenance, and the technical metadata that was imported from the AIP. Therefore, TMS is the hub for all information about the artwork.

Wharton and Engel elegantly navigated this boundary in their work on the Artist Archives Initiative. In speaking about their work at the 2020 AIC Annual Meeting,

Engel quoted Lev Manovich, who said, "Competing to make meaning out of the world, database and narrative produce endless hybrids. It is hard to find a pure encyclopedia without any traces of a narrative in it and vice versa" (Manovich 2007, 51, cited in Wharton et al. 2020). It is exactly this boundary that time-based media conservators at museums may struggle to delineate. Typically, most museum collection objects are still physical rather than digital. As Chief Conservator Kate Lewis emphasized in an interview about the DRMC, time-based media conservators must balance the unique needs of digital collection objects with the need to ensure that the digital collection remains integrated with the art collection as a whole. She stated, "How digital art is stored, and how we retrieve it is different than object-based collections. That's the edge where it is getting pioneered, but to the museum at large it shouldn't look different, it should feel like the same collection" (Van Saaze et al. 2018, 229).

The data model and the narrative approach can be equally detailed and standardized in terms of the information they contain. One is more machine-friendly, whereas the other is more human-friendly. In his article "The Role of the Technical Narrative for Preserving New Media Art," Mark Hellar (2015) described how SFMOMA developed a "technical narrative" to serve as a standard document to be created for all digital-based artworks (fig. 2). It is interesting to compare the "technical narrative" approach to the DRMC data model (fig. 3).

This is a standardized system for documenting digital artworks. The purpose of the technical narrative is to provide:

- A high-level functional description of the work. This is a general description of how the work functions and operates as a whole. This part of the narrative is a platform-neutral description of the work in a general and functional way.
- A modular examination of the individual components of the work and their specific functions. The intent of this section is to look at every individual

component of the work in detail. Additionally, a high-level examination is given to how all of the parts work as a complete system. This section attempts to map out a general technical schematic of the work.

- A detailed description of the artwork as it exists upon acquisition. This section is specific about the hardware, software, operating systems, languages, algorithms, video codecs, etc. These platforms, components and technologies are examined closely to inform an understanding of how they serve the operational requirements of the work. This section is closely tied to the technical documentation provided by the artist and engineers, describing the pragmatic requirements for operation and display.
- An analysis of the current technology platform and an evaluation of its longevity against the current state of technology. Here we consider the long-term stability of the piece upon acquisition. It calls out strategies and concerns in preserving the work over the long term and informs ongoing conservation and maintenance protocols including possible strategies for migration or emulation.

Fig. 2. Components of the technical narrative at SFMOMA (Hellar 2015)

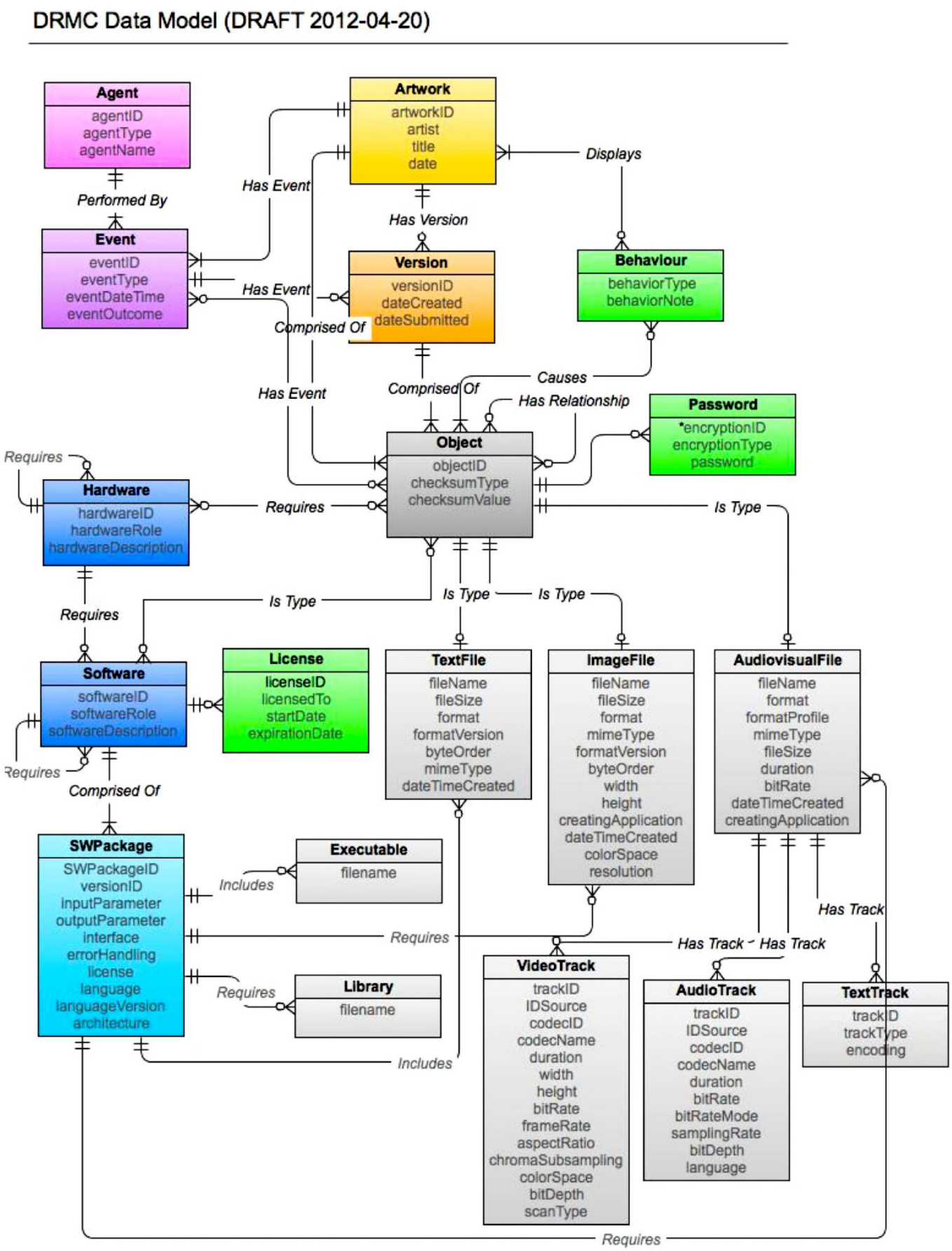

DRMC Data Model (DRAFT 2012-04-20)
Agent
agentID
agentType
agentName
Performed By
Event
eventID
eventType
eventDateTime
eventOutcome
Has Event
Has Event
Has Event
Artwork
artworkID
artist
title
date
Displays
Has Version
Version
versionID
dateCreated
dateSubmitted
Comprised Of
Behaviour
behaviorType
behaviorNote
Causes
Comprised Of
Has Relationship
Password
encryptionID
encryptionType
password
Object
objectID
checksumType
checksumValue
Requires
Hardware
hardwareID
hardwareRole
hardwareDescription
Requires
Requires
Is Type
Requires
Is Type
Is Type
Is Type
Software
softwareID
softwareRole
softwareDescription
Requires
Comprised Of
License
licenseID
licensedTo
startDate
expirationDate
TextFile
fileName
fileSize
format
formatVersion
byteOrder
mimeType
dateTimeCreated
ImageFile
fileName
fileSize
format
mimeType
formatVersion
byteOrder
width
height
creatingApplication
dateTimeCreated
colorSpace
resolution
AudiovisualFile
fileName
format
formatProfile
mimeType
fileSize
duration
bitRate
dateTimeCreated
creatingApplication
SWPackage
SWPackageID
versionID
inputParameter
outputParameter
interface
errorHandling
license
language
languageVersion
architecture
Includes
Executable
filename
Requires
Requires
Library
filename
Includes
Requires
Has Track
Has Track
Has Track
VideoTrack
trackID
IDSource
codecID
codecName
duration
width
height
bitRate
frameRate
aspectRatio
chromaSubsampling
colorSpace
bitDepth
scanType
AudioTrack
trackID
IDSource
codecID
codecName
duration
bitRate
bitRateMode
samplingRate
bitDepth
language
TextTrack
trackID
trackType
encoding
Requires

Fig. 3. Entity-relation diagram, draft DRMC data model (MoMA and AVPS 2012, 50), including entities, their attributes, and relationships. This diagram should be considered an early draft. It utilizes Information Engineering (IE)/Crows Foot notation.

Whether an institution chooses the database approach, the narrative, or a hybrid, it is clear that such decisions are highly context dependent. Robust data models may suit institutions with a different context or use cases for the data. The decision not to fully implement the data model at MoMA was not made consciously; rather, the staff has never had the capacity to carry out such detailed cataloging, especially when it duplicated information contained in the narrative documentation widely used by conservators and other museum staff.

Limitations of the OAIS Model

The foregoing section described how MoMA is packaging the digital elements of artworks as AIPs in accordance with the Library of Congress BagIt specification and storing them on LTO tape; however, the museum is not moving toward centralizing information about those elements within the metadata structure of the AIP, as would be typical in an archive following the OAIS model. In the course of daily work, time-based media conservators at MoMA are experiencing the conceptual tension that exists when a model developed for one context is partially adopted in another. Although the museum can function in this state of tension, it also creates an urgency to articulate the issues and structure a fully working model where presently there exists a set of practices in a conceptually liminal space.

The very models and software underpinning the DRMC were acknowledged from the start to be provisional. Over the years, a number of conservators and others working with time-based media have pointed out that the life-cycle model and OAIS reference model for an archive were never ideal for time-based media art

conservation. This is because there is no clear division between the active life of the digital components of artworks and their later archival phase. Fortunately, there are other models, including biographical approaches (Van de Vall et al. 2011) and continuum theory, where "records are not treated as an end product but as processes" (Laurenson et al. 2017), which may point the way forward.

The four-year European Union project PERICLES (Promoting and Enhancing Reuse of Information throughout the Content Lifecycle taking account of Evolving Semantics), which ended in 2017, considered "continuously changing environments such as for time-based media, where OAIS is less appropriate," instead basing the approach to preservation on a "continuum viewpoint" (Daranyi et al. 2015, 53). PERICLES drew on diverse examples of the creation and use of digital data at Tate and the Belgium Space Operations Centre for a series of case studies (Daranyi et al. 2015, 57–59) that led to the development of models and ontologies for the ongoing use and management of that data. To be clear, these are not models for conservation of time-based media artworks generally, but only for their digital elements. In its work on the digital video art case study, the team explained, "We are not aiming to model the domain at the artwork level, but rather the specifics of dependencies between digital things within a system which forms part of the artwork. It is therefore a partial model related to the dependencies of some of the components of the artwork. In this case, modelling enables the conservator to better understand the digital dependencies within the system and also helps identify areas where automation might be achievable. Modelling also facilitates communication with computer scientists and software developers who might provide tools to support activities related to long term digital preservation and the corresponding dependencies" (Lagos et al. 2018, part 4). The PERICLES team developed ontologies and tools for digital video art and software-based art, among other deliverables (CORDIS 2017, 40). Much in the way that Archivematica software was developed to create AIPs as conceptualized by the OIAS model, the PERICLES models led to the development of a suite of tools

that can be applied by diverse industries to dynamic preservation environments. The PERICLES team published the tools and research to its Preserveware Digital Preservation Hub for community use. As evidenced by MoMA's DRMC history, adapting tools into a preexisting information ecosystem at a museum is challenging. Specific tools aside, the continuum "viewpoint" that underpins this project suggests a way to reframe digital preservation storage for artworks to better fit the actual activity of preserving time-based media artwork records in the museum.

Records continuum theory was developed in Australia in the 1990s by records managers and archivists who took note of the fact that digital data did not follow a "life cycle" from record to archive. Instead, they observed that there was a dynamic, bi-directional relationship between their fields of practice, in which stages are blurred and recurring, and activities reverberate (McKemmish 2001, 340). In her history of records continuum theory, Sue McKemmish explained that "post-modern ideas about records view them as dynamic objects that are fixed in terms of content and meaningful elements of their structure, but linked to ever-broadening layers of contextual metadata that manages their meanings, and enables their accessibility and useability as they move through 'spacetime'" rather than on a linear timeline (McKemmish 2001, 349). The graphical representation of the model is shown in figure 4.

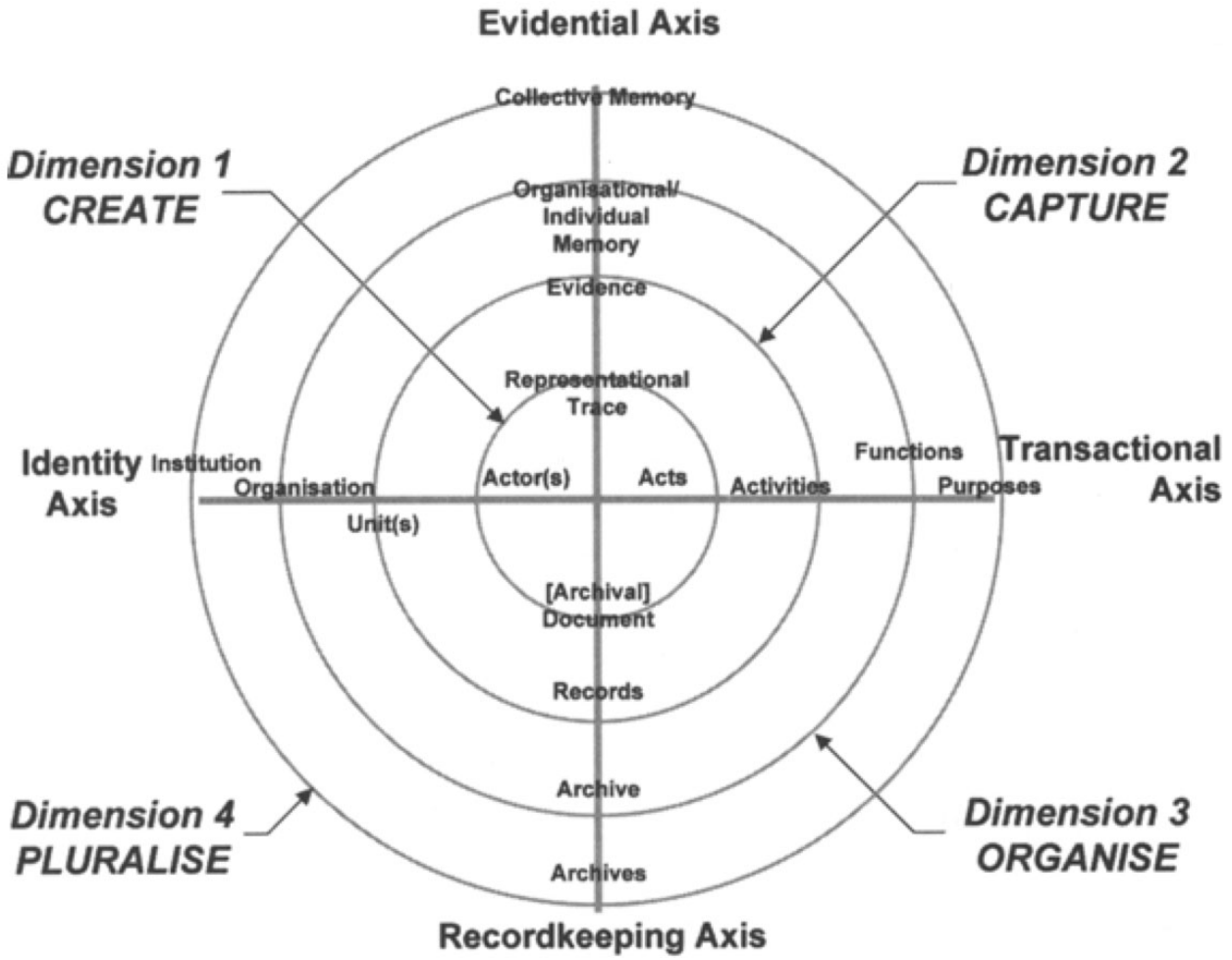


Fig. 4. Records continuum model (McKemmish 2001, 351)

This model does not exactly fit the management of digital art elements in an art conservation context either, but it opens up a way to contextualize and graphically represent how an archival document, such as an artist's video file, might fit within a larger preservation framework for data that has an ongoing, active life. Unlike in the OAIS model, the archival document in this framework is not a fully described static archival unit in storage in the center of the model, encircled by monitoring, extraction, and management processes. Instead, it is situated within a living record that spans a dynamic web of continual activity.

The DRMC was originally designed to account for this need for change over time. The system requirements described creating the ability to "version" AIPs, where

the term *version* referred to an AIP version or artwork version and was defined as follows: "A new version is created when any change happens to the state of the current AIP. This may include the addition of metadata, documentation, new components, or changes to the content of parts of the artwork (e.g., the artist delivers a new dataset for an installation, resulting in a new artwork version)" (Van Malssen and Fino-Radin 2013, 13). This would have been one way that the DRMC could allow AIPs a "life" through editing and supplementing AIP metadata, but as described previously, moving to LTO tape storage rendered the editing and updating of existing AIPs impractical. Using the AIP for safe file storage and TMS as a complementary documentation system and management application is meeting that need in a different way. Version control systems such as Git and MediaWiki provide another way to manage works that change over time (Barok et al. 2019). Conservators at MoMA have trialed these with selected artworks; however, they are used in addition to, not in place of, TMS and AIP storage.

Adapting the records continuum model to data management in a time-based media art conservation context can provide a useful technology-agnostic thought map. For example, placing cataloging and storing in the third dimension prompts consideration of the life of an AIP in the other three dimensions. The recordkeeping axis can be redefined so that the innermost level is a single digital component and the outermost level is the museum's complete artwork record, encompassing all information about the artwork in all of its forms, from its digital elements to documentation and embodied knowledge. Even so, there are inherent limitations to modeling. As McKemmish writes, "The records continuum worldview envisages an inclusive, multidimensional archival place . . . where there is recognition that the richness, complexity, diversity, and idiosyncrasies of the contexts in which records are created, managed, and used cannot be fully represented in our models, systems, standards, and schema, but that recognition does not detract from their significance and strategic importance to practice" (McKemmish 2001, 358–359).

In terms of implications for the DRMC project, the ways in which the museum has pulled back from pursuing an AIP structure that aligns ever more closely with the OAIS model is a direct result of the limitations of applying an archival model in an art museum context. The potential applicability of continuum theory and other models to the art conservation context and how they can be adapted to guide the daily work of time-based media conservators in the museum should continue to be explored. Records continuum theory is one example of how other models may help resolve the conceptual tensions that conservators experience in their daily work with preservation storage systems for the digital elements of artworks.

Conclusions

To accommodate a rapidly growing collection of digital acquisitions, the David Booth Conservation Department at MoMA pioneered methods to safely store the digital elements of artworks. After several years of planning and research, the DRMC was first proposed in 2010. Over the following decade, many factors would impact its evolution, including contemporaneous developments in the digital preservation field, new technologies and software, staffing changes, and changes in other museum departments and systems. Along the way, there were growing pains during implementation, as theory was put into practice and as practices were made practical. Passing the 10-year milestone provides a moment to reflect on the museum's key learnings and future direction. Not only did MoMA build a team to gradually transform the institution to care for time-based media artworks, but it also empowered that team to *do preservation* of these artworks on a continuing basis. Now running a mature storage system, time-based media conservators at MoMA are experiencing firsthand the conceptual and practical tensions resulting from the limitations of the underpinning models and software adapted from the archives world. The first decade of the DRMC is an important case study that offers a full spectrum of lessons, from stumbles and struggles to

visionary goals and clear successes, but what also stands out is the uniqueness of MoMA's project. As more museums join MoMA on this journey, a dynamic community of practice around digital storage for artworks will continue to grow and advance the field toward collaboratively established norms.

ACKNOWLEDGMENTS

The author is deeply grateful to the following current and former members of the DRMC team and external partners for reviewing the text: Jim Coddington, formerly Agnes Gund Chief Conservator, MoMA; Justin Simpson, Managing Director, Artefactual Systems, Inc.; Kara Van Malssen, Managing Director, Consulting, AVP; Glenn Wharton, Lore and Gerald Cunard Chair, UCLA/Getty Program in the Conservation of Archaeological and Ethnographic Materials and Professor of Art History and Conservation of Material Culture, UCLA; and the MoMA Media Conservation team: Kate Lewis, Peter Oleksik, Sarah Gentile, and Lia Kramer. I am also indebted to the entire DRMC core team:

Media Conservation
Kate Lewis, Agnes Gund Chief Conservator; Peter Oleksik, Associate Media Conservator; Sarah Gentile, Assistant Digital Preservation Specialist

Information Technology
Diana Pan, Chief Technology Officer; Helynsia Brown, Director; John DuFour, Senior Manager, Infrastructure; Sergiy Petrychuk, Technology; Rik Vanmechelen, Manager of Enterprise Applications; Steven Moore, Developer and Database Administrator; Ryan Sprott, Developer, Apps

Exhibitions & & Collections
Ian Eckert, Associate Director, Collection & Exhibition Information; Jen Sellar,

Digital Assets Manager, Imaging and Visual Resources

With support from Curatorial, Legal, Finance, Film, and Registration teams

APPENDIX

2012 Requirements	2020 Status
I. Staging Area	
1. BR-1 SIP Staging Area SIP components MUST reside in a staging area until the completed SIP is ready for ingest, which occurs once all components have been received, the SIP requirements checklist is completed and approved by Manager, minimum required metadata created and input, virus check is complete, and checksums generated. The SIP staging area should be in an environment external to the primary DRMC storage area.	
FR-1: Store files in staging area for SIP components until ready for ingest*	Bags stored on MoMA staging server
FR-2: Generate SIP requirements checklists	SIPs structured using BagIt/Archivematica
FR-3: Generate SIP requirements automatically based on established SIP classes	SIPs have uniform structure
FR-4: Allow users to generate SIP requirements manually based on a set of options	N/A

FR-5: Track and display the arrival of new submission files*	Archivematica
FR-6: Indicate state of SIP completeness	Archivematica
FR-7: Assign unique identifier to each digital object*	Archivematica
BR-2: Pre-ingest processing Preliminary file addition, virus check, condition reporting, and metadata creation/addition MUST be performed on SIP in the staging area before the work is ingested into archival storage.	
FR-8: Verify SIPs according to requirements	Archivematica
FR-9: Create checksums*	Archivematica
FR-10: Extract technical metadata from files (e.g. characterize files), index, and store*	Archivematica
FR-11: Detect password protection*	Manual process
FR-12: Log all events that occur in an object's lifecycle*	Archivematica logs PREMIS events for all preservation actions executed by Archivematica. No other lifecycle events are logged in the AIP.
II. Ingest	
BR-3: Manage the original object AIPs should be created in a way that preserves the content, structure and integrity of the original artwork without modification of the underlying file or filesystem.	

FR-13: Package AIPs	Archivematica
FR-14: Retain original filenames	Archivematica (except forbidden characters)
FR-15: Retain original file structure	Archivematica
BR-4: Add to AIPs Artworks often change over time, such as through new documentation, or new exhibition components. Users MUST be able to add new metadata, artwork components, and other new elements over time. Additions or other modifications will result in new versions of an AIP.	
FR-17: Allow users to add to existing AIPs	No; AIP is not edited. New AIP is added.
FR-16: Enforce metadata requirements	No
FR-18: Allow users to add documentation and metadata	No; AIP is not edited
BR-5: Manage versions of AIPs Versions should be managed so that the original AIP elements are retained in addition to the new elements. Examples of new versions include the addition of new documentation or metadata, the deposit of a new dataset from the artist, or a migration performed by conservation. Versions could be managed through a variety of approaches, such as delta versioning, forward-delta versioning, or reverse delta versioning.	
FR-19: Submit full or partial versions*	No; AIP is not edited. New AIP is added.
FR-20: Manage versions of AIPs*	No; AIP is not edited
FR-21: Retain all versions*	All AIPs are retained
FR-22: Prompt users to indicate active	

status of files when new versions are added	Status changes are entered into TMS
FR-23: Assign version identifiers*	AIP UUID
III. Storage	
BR-6: Archival storage Artwork files MUST have a dedicated storage area. Archival storage MUST be a replicated, highly secure environment. AIPs in archival storage are considered permanent, and cannot be deleted after ingest except by a super administrator. Modifications are allowed to existing AIPs through the addition of new files and metadata, resulting in new AIP versions.	
FR-24: Backup AIP according to storage hierarchy policies*	Storage vendor
BR-7: Manage file integrity Files MUST be checked periodically for corruption by generating checksum values and comparing stored checksum hashes. Corrupt files MUST be repaired with good backups.	
FR-25: Run checksum validation according to a set schedule*	Storage vendor
FR-26: Report integrity check outcome to System Administrator*	Storage vendor
FR-27: Report file repair event results*	Storage vendor (this has not yet occurred)
IV. Description and Access	
BR-8: Provide a database and application layer for metadata management, search and retrieval Users will need to interact with data about the artworks stored in the DRMC through a user- friendly interface. Users MUST be able to perform searches, input metadata, and create reports through this interface.	

FR-28: Index metadata for storage in database*	AIP METS XML
FR-29: Allow users to add metadata received through application layer to AIP*	No; AIP is not edited
BR-9: Integration between DRMC, DAM, and TMS The DRMC MUST push and pull data between itself and TMS according to mapping and business rules. DAM should house proxies of artwork files stored in the DRMC, and should be used to store and retrieve any exhibition copies that are created. TMS will contain information about the artwork, artist, accessioning, location of physical copies, etc. TMS will also need to receive a core set of metadata from the DRMC to support end-user needs.	
FR-30: Pull data from TMS to DRMC: Accession Number, Title, Artist, Date(s), ObjectID, and ComponentID*	Component number (dc:identifier); Object ID and Component ID from SIP name
FR-31: Push data from DRMC to TMS: Technical metadata*	AIP UUID, AIP METS XML
FR-32: Pull data from DAM to DRMC*	No
FR-33: Push data from DRMC to DAM*	No; DAM file submission is a separate, manual process
FR-34: Have a status indicator for different areas of a record	No
BR-10: Description Users MUST be able to enter comprehensive metadata about the artwork, individual files, documentation, relationships, etc, through a user-friendly application layer.	

FR-35: Allow data entry using multiple data types*	TMS
FR-36: Enable documentation of file relationships and dependencies*	TMS
FR-37: Identify file relationships and dependencies	TMS
FR-38: Provide read-only access to AIP directories*	Storage vendor GUI
FR-39: Display documentation	TMS
BR-11: Comprehensive search and browse Users MUST be able to perform complex search queries based on technical, administrative, and descriptive parameters through the GUI application layer.	
FR-40: Facilitate search queries based on technical metadata*	TMS
FR-41: Facilitate search queries based on lifecycle metadata*	No
BR-12: Reporting service Users MUST be able to produce a variety of reports from the system.	
FR-42: Generate report based on user requested data*	TMS, vendors
FR-43: Visualize data queries for reports	Conservation staff (manual process)
BR-13: Check files out/in Users may need to check out files for various purposes, including code documentation, creation of exhibition versions, virtualization,	

migration or other access or conservation activity. New files may be checked in to new AIP versions.	
FR-44: Check out part or all of an AIP*	Storage vendor GUI
V. Conservation	
BR-14: Obsolescence monitoring The DRMC MUST have mechanisms for monitoring and documenting obsolescence of artwork components, whether automated or manual. Automated obsolescence monitoring will be an important future function of the DRMC, as artworks are accessed infrequently and cannot be monitored manually. Also, DRMC manager will not be able to know the obsolescence status of every format and language housed in the DRMC. Format registries such as UDFR should be integrated using the API to support part of this function. Hardware obsolescence status must also be tracked.	
FR-45: Allow users to manually flag risk status*	No
FR-46: Automatically check formats for obsolescence risk	No
FR-47: Automatically flag risk status when detected*	No
BR-15: Support for conservation planning The DRMC SHOULD contain a conservation planning service in order to model, budget and plan for conservation action. While this is a recognized research area, and not likely to be available at the time of implementation, it should be enabled for future deployment. Conservation planning tools may be from projects such as Plato Planets23 or the Open Planets Foundation Scape project.	
FR-48: Provide a mechanism for users to create, save and execute	No

conservation plan	
VI. Administration	
BR-16: Permissions-based user roles The system MUST support various users, who will need different permissions for the system, including upload, edit, create reports, administrate system, administrate users, etc.	
FR-49: Allow administrative users to create user roles and assign permissions*	Access to servers and AIPs is controlled by MoMA IT
FR-50: Enforce system access based on user roles and permissions*	Access to servers and AIPs is controlled by MoMA IT
BR-17: Vocabulary/Ontology Management Data elements, values, and relationships MUST be defined and managed through an ontology or vocabulary management service. Ontology may be expressed in RDF/OWL, possibly based on the TOTEM data model developed by the KEEP project.	
FR-51: Manage ontology	TMS
FR-52: Allow users to edit ontology	TMS
BR-18: DRMC upgrades, maintenance, and development It should be relatively easy to introduce a new service to the DRMC architecture and define workflows and business rules associated with that service. The DRMC MUST support the addition, removal and configuration of services by MoMA staff or external developers (see section 8.1 on technical requirements for SOA or micro-service architecture).	
FR-53: Allow users to add micro-services*	Limited flexibility within Archivematica
FR-54: Allow users to remove micro-	Limited flexibility within

services*	Archivematica
FR-55: Perform tests of services or workflows without affecting data*	Archivematica and storage vendor development environments
Non-Functional Requirements	
Architecture	
Technical (storage, database, hosting, server, OS, client)	
Policy (SIP requirements, adding new components, metadata minimums, migration, commenting artist-supplied code, backups, file audit)	
Security (user ID/authorization, access control and permissions)	
Licensing (overall DRMC license MUST be compatible with the licenses of the underlying components, open source components, etc)	

DRMC Business (BR) and Functional (FR) Requirements 2012
** = Mandatory*

NOTES

1. Like MoMA, Tate records status for digital elements of artworks. Claudia Roeck's master's thesis lists the status types or component "classifications" used at Tate. They are "archival master (AM), an artist's supplied master (ASM), an artist's verified proof (AVP), an exhibition format (EF), a research copy (RC), a duplicating copy (DC) or documentation (DOC)" (Roeck 2016, 25).
2. Tate has done extensive research in recent years on how to create and describe relationships in the digital preservation ecosystem. The Tate digital preservation system requirements developed for digital video art during the PERICLES project described an "AIP network" that would be created for each

artwork, comprised of AIPs whose relationships to the "root AIP" (Artist Supplied Master) and to each other could be described and maintained within the system (Hedges et al. 2015). The PERICLES team also explored applying a Linked Resource Model to the digital preservation of video and software-based artworks. The Linked Resource Model defines an ecosystem through resources and dependencies, and has both a static and dynamic schema. For the static schema, the team created terms for relationships that encompassed preconditions, intentions, specifications, and impact (Kontopoulos 2016). The static and dynamic schemas were published on GitHub at the end of the project (https://github.com/nikolaosLagos/Linked_Resource_Model).

3. MoMA is working toward leveraging recent research by fellows Flaminia Fortunato and Caroline Gil at MoMA and by others, including the Guggenheim's Conservation of Computer-based Art (CCBA) team. The NYU-MoMA Conservation of Computer-based Art working group was a one-year project that ran from 2009 to 2010, and the academic-museum partnership between MoMA and NYU ended with the departure of Wharton from MoMA in 2013. However, an entirely separate initiative with the same name started in 2014 when Joanna Phillips, time-based media conservator at the Solomon R. Guggenheim Museum in New York, began collaborating with Deena Engel, clinical professor in the Department of Computer Science at the Courant Institute of Mathematical Sciences, and her students at NYU (Dover 2016).

REFERENCES

Barok, Dušan, Julie Boschat Thorez, Annet Dekker, David Gauthier, and Claudia

Roeck. 2019. "Archiving Complex Digital Artworks." *Journal of the Institute of Conservation* 42 (2): 94–113 . Accessed July 17, 2020. https://doi.org/10.1080/19455224.2019.1604398.

Community Research and Development Information Service (CORDIS), Publications Office of the European Union. 2017, March. "A New Approach to Digital Content Preservation." *Research*eu Results Magazine* 60: 40–41. Accessed August 26, 2020. https://ec.europa.eu/information_society/newsroom/image/document/2017-19/research_eu_0E0D8065-D64E-6D4C-400852DD349D3CBA_44528.pdf

Daranyi, Sandor, John McNeill, Ioannis Kompatsiaris, Panagiotis Mitzias, Marina Riga, Fabio Corubolo, Efstratios Kontopoulos, Nikolaos Lagos, Christian Muller, Simon Waddington, Mark Hedges, and Jean-Yves Vion-Dury. 2015. "PERICLES – Digital Preservation Through Management of Change in Evolving Ecosystems." In *The Success of European Projects Using New Information and Communication Technologies (EPS Colmar 2015)*. 51–74. Accessed August 26, 2020. https://pdfs.semanticscholar.org/703d/0ae883571e725c55f9469f07d0a1711064e0.pdf.

DOCAM Research Alliance. 2010a. *DOCAM Documentation Model*. Montreal: The Daniel Langlois Foundation for Art, Science, and Technology. Accessed July 17, 2020. https://www.docam.ca/en/presentation-of-the-model.html.

DOCAM Research Alliance. 2010b. *Glossaurus – Hierarchy*. Montreal: The Daniel Langlois Foundation for Art, Science, and Technology. Accessed July 17, 2020. https://www.docam.ca/glossaurus/hierarchy.php?lang=1.

Dover, Caitlin. 2016, October 26. How the Guggenheim and NYU Are Conserving Computer-Based Art—Part 1. *Solomon R. Guggenheim Museum Blog*. Accessed August 20, 2020. https://www.guggenheim.org/blogs/checklist/how-the-guggenheim-and-nyu-are-conserving-computer-based-art-part-1.

Engel, Deena, and Glenn Wharton. 2014. "Reading Between the Lines: Source Code

Documentation as a Conservation Strategy for Software-Based Art." *Studies in Conservation* 59 (6): 404–415. Accessed July 15, 2020. http://glennwharton.net/wp-content/uploads/2015/07/Engel_Wharton-Reading-Between-the-Lines.pdf.

Griesinger, Peggy. 2016. "Process History Metadata for Time-Based Media Artworks at the Museum of Modern Art, New York." *Journal of Digital Media Management* 4 (4): 331–342. Accessed July 8, 2020. https://www.henrystewartpublications.com/sites/default/files/Griesinger.pdf.

Hedges, Mark, Pip Laurenson, John McNeill, Anna Henry, Katarina Haage, John Langdon, Patricia Falcão, Louise Lawson, and Madeline Betts. 2015. Arts and Media Application Domain Requirements: Subdomain Requirements – Digital Video Art. *PERICLES Wiki*. Accessed July 16, 2020. https://projects.gwdg.de/projects/pericles-public/wiki/subdomain-requirements-digital-video-art.

Hellar, Mark. 2015. "The Role of the Technical Narrative for Preserving New Media Art." *Electronic Media Review* 3: 38–47. Accessed August 20, 2020. http://resources.culturalheritage.org/emg-review/volume-three-2013-2014/hellar/.

Kirschenbaum, Matthew, Richard Ovenden, Gabriela Redwine, with research assistance from Rachel Donahue. 2010. *Digital Forensics and Born-Digital Content in Cultural Heritage Collections*. Washington, DC: Council on Library and Information Resources. Accessed July 16, 2020. https://www.clir.org/pubs/reports/pub149/.

Kontopoulos, Efstratios, Marina Riga, Panagiotis Mitzias, Stelios Andreadis, Thanos G. Stavropoulos, Nikolaos Lagos, Jean-Yves Vion-Dury, Georgios Meditskos, Patricia Falcão, Pip Laurenson, and Ioannis Kompatsiaris. 2016.

Ontology-Based Representation of Context of Use in Digital Preservation. Paper presented at the 1st Workshop on Humanities in the Semantic Web (WHiSe 2016), Heraklion, Crete, Greece. Accessed March 17, 2021. http://doi.org/10.5281/zenodo.344824.

Lagos, Nikolaos, Marina Riga, Panagiotis Mitzias, Jean-Yves Vion-Dury, Efstratios Kontopoulos, Simon Waddington, Georgios Meditskos, Pip Laurenson, and Ioannis Kompatsiaris. 2018. "Dependency Modelling for Inconsistency Management in Digital Preservation – The PERICLES Approach." *Information Systems Frontiers* 20: 7–19. Accessed August 27, 2020. https://core.ac.uk/download/pdf/77065183.pdf.

Laurenson, Pip, Kevin Ashley, Luciana Duranti, Mark Hedges, Anna Henry, John Langdon, Barbara Reed, Vivian van Saaze and Renee van de Vall. 2017. *The Lives of Digital Things: A Community of Practice Dialogue*. London: Tate. Accessed July 15, 2020. https://www.tate.org.uk/about-us/projects/pericles/lives-digital-things.

Manovich, Lev. 2007. "Database as Symbolic Form." In *Database Aesthetics*, edited by Victoria Vesna. Minneapolis: University of Minnesota Press.

McKemmish, Sue. 2001. "Placing Records Continuum Theory and Practice." *Archival Science* 1 (4): 333–359.

Museum of Modern Art (MoMA) and AudioVisual Preservation Solutions (AVPS). 2012. MoMA Digital Repository for Museum Collections: System requirements. Unpublished manuscript. Conservation Department, Museum of Modern Art, New York.

Oleksik, Peter. 2015. "Wrangling Electricity: Lessons Learned from the Mass Migration of Analog and Digital Media for Preservation and Exhibition." *Electronic*

Media Review 3: 38–47. Accessed August 20, 2020. http://resources.culturalheritage.org/emg-review/volume-three-2013-2014/oleksik/.

Prater, Scott. 2018. "How to Talk to IT About Digital Preservation." *Journal of Archival Organization* 15 (3): 90-101. Accessed August 24, 2020. http://digital.library.wisc.edu/1793/78844.

Roeck, Claudia. 2016. Preservation of digital video artworks in a museum context: Recommendations for the automation of the workflow from acquisition to storage. Master's thesis and appendix, Bern University of the Arts. Accessed July 16, 2020. https://www.academia.edu/41381787/Preservation_of_digital_video_artworks_in_a_museum_context_Recommendations_for_the_automation_of_the_workflow_from_acquisition_to_storage https://www.academia.edu/41381788/Masters_Thesis_Appendix.

Rubin, Nan. 2010. Final Report: Preserving Digital Public Television. *National Digital Information Infrastructure and Preservation Program, Library of Congress.* Accessed July 17, 2020. https://www.thirteen.org/ptvdigitalarchive/uncategorized/final-report-preserving-digital-public-television/.

Smith, Madeline Page. 2020. Caring for the moving image in art museums: *Matters in Media Art* and the stewardship of time-based media artworks. Master's thesis, Moving Image Archiving and Preservation, New York University. Accessed September 12, 2020. https://miap.hosting.nyu.edu/program/student_work/2020spring/20s_thesis_Smith_deposit_copy_y.pdf.

Van de Vall, Renée, Hanna Hölling, Tatja Scholte, and Sanneke Stigter. 2011. "Reflections on a Biographical Approach to Contemporary Art Conservation." In *ICOM Committee for Conservation preprints. 16th Triennial Conference, Lisbon*. Paris: ICOM. 19–23. Accessed August 31, 2020. https://pure.uva.nl/ws/files/1262883/115640_344546.pdf.

Van Garderen, Peter, P. Jordan, T. Hooten, C. Mumma, E. McLellan. 2012. "The Archivematica Project: Meeting Digital Continuity's Technical Challenges." *Proceedings of the UNESCO Memory of the World Conference*. Vancouver, Canada. 1–11. Accessed July 15, 2020. http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/mow/VC_Van_Garderen_et_al_26_Workshop1.pdf.

Van Malssen, Kara, and Ben Fino-Radin, with Glenn Wharton, Jim Coddington, Juan Montes, James Heck, Chris Lacinak. 2013. Digital Repository for Museum Collections: Management application requirements. Unpublished manuscript. Conservation Department, Museum of Modern Art, New York.

Van Saaze, Vivian, Glenn Wharton, and Leah Reisman. 2018. "Adaptive Institutional Change: Managing Digital Works at the Museum of Modern Art." *Museums & Society* 16 (2): 220–239. Accessed July 15, 2020. https://www.researchgate.net/publication/327402175_Adaptive_Institutional_Change_Managing_Digital_Works_at_the_Museum_of_Modern_Art.

Wharton, Glenn. 2009. Digital Collections Conservation Repository (DCCR) project justification. Unpublished manuscript. Conservation Department, Museum of Modern Art, New York.

Wharton, Glenn, and Deena Engel. 2015. "Museum and University Collaboration in Media Conservation Research" *Electronic Media Review* 3: 111–117. Accessed

August 20, 2020. http://resources.culturalheritage.org/emg-review/volume-three-2013-2014/wharton/.

Wharton, Glenn, Deena Engel, and Barbara Clausen. 2020. Developing the Joan Jonas Knowledge Base: An Open Access Digital Resource. Paper presented at the Electronic Media Group–Contemporary Art Network Joint Session of the AIC 48th Annual Meeting, Salt Lake City, UT.

Wharton, Glenn, and Barbra Mack. 2012. "A Case for Digital Conservation Repositories." *Electronic Media Review* 1: 23–44. Accessed July 15, 2020. http://resources.culturalheritage.org/emg-review/wp-content/uploads/sites/15/2016/07/Vol-1_Ch-5_Mack_Wharton.pdf.

Wharton, Glenn, Kara Van Malssen, Sydney Briggs, Deena Engel, Jeri Moxley, Cara Starke, Ramona Bannayan, and Jim Coddington. 2010. Design for a digital repository for museum collections. Unpublished manuscript. Conservation Department, Museum of Modern Art, New York.

FURTHER READING

Goethals, Andrea, Nancy McGovern, Sibyl Schaefer, Gail Truman, and Eld Zierau. 2018. Digital Preservation Storage Criteria. Accessed August 24, 2020. https://osf.io/sjc6u/.

National Digital Stewardship Alliance. 2018. Levels of Digital Preservation (Version 2.0). Accessed August 24, 2020. https://ndsa.org/publications/levels-of-digital-preservation/.

Phillips, Megan, Jefferson Bailey, Andrea Goethals, and Trevor Owens. 2013. The

NDSA Levels of Digital Preservation: An Explanation and Uses. *National Digital Stewardship Alliance*. Accessed August 24, 2020. http://www.digitalpreservation.gov/documents/NDSA_Levels_Archiving_2013.pdf.

Amy Brost
Assistant Media Conservator
The David Booth Conservation Department
Museum of Modern Art, New York

Electronic Media Review / Proudly powered by WordPress