TRANSLATING THEORY TO PRACTICE:

DEFINING DIGITAL PRESERVATION PLANNING IN MUSEUMS

A thesis submitted to the faculty of
San Francisco State University
In partial fulfillment of
The requirements for
The Degree

Master in Arts

In

Museum Studies

By

Emma Palakika James

San Francisco, California

January 2016

CERTIFICATION OF APPROVAL

I certify that I have read *Translating Theory to Practice: Defining Digital Preservation Planning in Museums* by Emma Palakika James, and that in my opinion this works meets the criteria for approving a thesis submitted in partial fulfillment of the requirement for the degree Master of Arts in Museum Studies at San Francisco State University.

Victoria C. Lyall

Dr. Victoria Lyall
Professor of Museum Studies

Dr. Edward Luby
Professor of Museum Studies

TRANSLATING THEORY TO PRACTICE:

DEFINING DIGITAL PRESERVATION PLANNING IN MUSEUMS

Emma Palakika James
San Francisco, California
2016

Digital preservation is an emerging activity in museums today. The development of technology as a tool for work, research, information capture, and artistic expression, as well as the increasing percentage of important cultural materials created only in digital form, argues that museums must begin to focus on digital preservation. In this thesis, digital preservation in museums is examined, specifically the development, planning, and implementation of digital preservation initiatives. First, a literature review of digital preservation basics, cross-disciplinary collaboration, and digital preservation policy is presented, followed by case studies of three best-practicing museums. Four key themes are discussed, including defining digital preservation, integration of digital preservation technology, collaboration, and policy development. Finally, several conclusions and recommendations are presented, most notably that digital preservation in a museum context must be viewed and implemented from a collections management perspective.

I certify that the Abstract is a correct representation of the content of this thesis.

Victoria P. Lynn

Chair, Thesis Committee

12/09/15

Date

ACKNOWLEDGEMENTS

I would like to thank my thesis committee of Dr. Victoria Lyall and Dr. Edward Luby for their generous support and guidance through this process. Your feedback has helped me develop into a stronger writer and a better museum professional. Thanks are owed to Layna White, Jenny Choi, Dan Lipcan, and Ben Fino-Radin for allowing me to interview them for this thesis. Finally, I would like to thank my family and friends for their continued support and encouragement, without which I would never have been able to write this thesis.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF APPENDICES

Chapter 1: Introduction

Museum collections management and care has always been a necessary focus of the museum profession. As object-centered institutions, museums collect, as part of their general mission, the unique and rare objects that are important to our human history and legacy, and protect them for the public good. However, today in the 21st century, it can now be recognized that traditional collections management and care has shifted in the face of the increased use of computational and digital technology in society. The normativity of technology as a tool for work, research, information capture, photography, film, and other artistic expression is palpable in our 'postdigital' society today. There is an increasing percentage of important cultural materials created only in digital form, many of which will be deemed important to our cultural history that deserves to be collected and stewarded within museums. However, currently, digital preservation remains a new, and not-broadly practiced activity in museums. The practice of digital preservation will therefore become increasingly important to the museum field, and should be considered with the same responsibility and effort as traditional museum collection management. If museums are going to continue their role as well-equipped stewards for the cultural heritage of today and of our future, then digital preservation will need to be adopted within the broader scope of museum work.

The museum field will also need to understand that digital collections require a different level of care and attention than traditional collections - the practice of *digital preservation* and *digital stewardship* to be explained and discussed in this thesis. Significantly, digital objects by nature are malleable, unfixed, immaterial, and often without stable physical manifestations. An oil painting may not be collected by a museum for many decades after its creation, but there is essentially no fear that the painting and its historical meaning will be inaccessible as long as it physically exists. However, an important piece of digital film is a series of source code formatted using a particular video codec and requires various levels of technology mediation; as a result, it cannot be expected renderable for viewing decades from its creation because by that time its particular set of technology will be obsolete. As we know, the technology industry is motivated by futuristic goals of advancement, change, and improvement. For example, as this commercial industry changes the smartphone every 6 months, the smartphones of three years ago

are now barely functional, and most certainly not supported by its manufacturer. However, this reality is inherently very problematic for cultural heritage, museum collecting, and also for collections management -- museums may not be able to keep up with the fast-pace of the technology industry, which controls the availability of digital materials, even that of the not-so-distant past.

In some form or another, eventually all museums will adopt digital technology into their institutional assets, museum archives, and museum collections, all of which will continually be expected to be cared for and preserved just as long as any analog collections. The practice of digital preservation is very much aligned with the theoretical practices of museum collection management, but involves a very different set of tools and procedures. As collections continue to change, so must our policies, procedures, and protocols for responsible collection management of the digital age.

How will museums prepare for caring and stewarding digital collections when digital formats, software, hardware, and media carriers will be constantly in flux? How can we ensure that museums are well-equipped for digital stewardship? The answers can be found within the digital preservation best practices and procedures put forth by a community that has been manifested in the library, archive and scientific research fields. This thesis strives to use the models set forth by the library and archive fields to relate the needs of museums and collections management within the context of digital preservation.

Digital preservation is also an important practice for museums who do not collect objects in digital formats. Institutional assets and investments are also important to protect. For example, the museum profession has seen over a decade of striving for digitizing material beginning with text, and then moving into photographic material, three-dimensional objects, audio-visual materials, motion pictures, and sound recordings. What has resulted is the development of huge collections of digital surrogates for museum objects or forms of expression, that are themselves becoming assets in their own right. As more and more people rely on them for use and access, they become as critical in understanding a museum's mission as the originals they stand in place for. Therefore, it is also necessary for museums to ensure ongoing access to

these digital assets. While digitization was the very beginning of increased public access to collections, digital preservation is simply the flip side of ensuring ongoing access -- providing consistent entry to information that is already manifested in digital form. If access to collections is becoming a mainstreamed part of the Museum's responsibility, the ongoing access to born-digital institutional assets is also certainly worthy of consideration.

A final case for good practice in digital preservation, is that cultural heritage institutions, including museums, are the only ones with distinctive preservation mandates for the public trust. If cultural heritage institutions do not deal with digital preservation of important materials, it is likely that no one else will. In many more cases than museums, Libraries and Archives are the current pioneers in digital preservation, and they are already acutely aware of the importance of digital preservation mandates, plans, and policies. Museums will likely come into the need for digital preservation as digital collections increase in tandem; however, when dealing with digital objects, time is of the essence. Technology will continue to change, digital media will continue to decay, and data will continue to be lost if museums do not commit to some form of action sooner than later. This thesis strives to help elucidate for the museum field advocacy for the integration of digital preservation practice with the familiar practice of collection management, records management, and collections care.

Making digital preservation accessible to the museum field is surely not as simple as understanding the need, although this is certainly an important starting point. Levels of education, technology, and institutionalization will need to occur in order to enable a field-wide effort towards digital stewardship in museums. Long term digital preservation is also not solely driven by technology, and is most truly an issue that management and governing bodies of museums also need to recognize, without which, the personnel and funding requirements for digital preservation will remain unsupplied. Organizations whose mission encompasses the preservation of cultural heritage and information will have high reliability requirements for their digital preservation systems. Museums will need policies and plans that can define those high standards, staff roles, and protocols. Fortunately, the literature and research that can alleviate the need for education on the qualities, requirements, and technological recommendations of digital preservation exists in a well-formed digital preservation community. In addition, although

museums are currently behind in achieving digital preservation, there are some key, good-practicing museums from which other museums can look to for models and community. This thesis will explore the issue of digital preservation planning, implementation, and policy in the museum context to provide advocacy for this kind of collection management, as well as increase accessibility to key information that is relevant to the museum field. The major themes prevalent to digital preservation planning and policy are explored including: the nature of digital objects; threats to digital collections; definitions, standards, and strategies for digital preservation; collaboration; and digital preservation policy. This thesis is an endeavor to bridge the gap between current standards and new practices, and offers conclusions and recommendations to the field.

As a note to the reader, digital preservation is an important practice for museums of all types, art, science, history included. Currently, much digital preservation practice in museums has been focused around art museums, and the three case study institutions chosen for this thesis are art museums. As such, although the topics discussed will be applicable and relevant to art, history, and science museums alike, the analogies and examples used throughout this thesis were set within the parameter of art museums. However, these parameters are not intended to limit the reader to apply the same practices and principles to the context of other museum fields. The following section of this chapter will outline the organization of this thesis.

The Organization of This Thesis

This thesis is organized into four main sections: (1) the literature review; (2) the methodology; (3) the case study analyses; (4) the discussion and conclusions. The first section of this thesis is a review of relevant literature from the digital preservation community, as it pertains to the museum context. The literature review can be found in chapters two through five. Specifically, chapter two reviews the nature of digital objects and the unique vices that threaten their ability to be preserved. In chapter three, the definitions, history, terminology, theories, best practices, and practical steps and strategies of digital preservation are reviewed to summarize the basics of digital preservation. The fourth chapter speaks to the collective mission of libraries, archives, and museums and why collaborative models from the library and archive fields are

worthy models for future digital preservation in museums. Finally, the fifth chapter reviews the importance of policy in museums, and the implications that digital preservation policy will have for institutionalizing digital preservation in museums.

The second section of this thesis is found in chapter six, which is a description of the specific methodology used to conduct research for this thesis. The selection process for the literature review and conducting case studies is outlined in this chapter. The third section of this thesis, chapters seven through nine, highlights the good work in digital preservation being done by three key case study institutions: The Metropolitan Museum of Art, The San Francisco Museum of Modern Art, and The Museum of Modern Art. The case study chapters include a review of relevant information of each institution's practices, information from interviews with key museum personnel, and a comprehensive analysis of the institution's practices within the scope of digital preservation in the museum field.

In the last section of this thesis, the discussion and conclusions are presented in chapters ten and eleven. Chapter ten is a discussion of findings and themes between the three case study institutions and the literature review that elucidate the key challenges and successes of digital preservation in the museum field today. Chapter eleven presents the final conclusions of this thesis, along with practical recommendations for digital preservation that can be achieved by museums of any size.

Scope

The scope of this thesis includes information and resources from a variety of fields including scientific research, computer science, library and information science, archival studies, and museum studies. The wide range of resources available in the digital preservation community means that the scope of this thesis is inevitably limited, and there may be sources, initiatives, tools, and vendors not mentioned in the body of the chapters. The field of digital preservation involves many layers of technology, and similar to the technology industry, may change quickly, and new tools, vendors, and initiatives are forming constantly. This thesis is a reflection of many of the useful, well-mentioned, relevant resources in the digital preservation field upon the writing of this thesis, but it is not intended to encompass the entire scope of the

digital preservation field, nor is it possible for the scope of this thesis to include the entirety of good work in digital preservation within the museum field.

A Case for Good Practice

Museums are of course well aware of the importance of protecting and preserving their collections, institutional assets and investments. Digital collections and digital assets are no different from the analog materials museums already work so hard to protect. In light of the changing landscape for museum work and collections, it is time to advocate for digital preservation as an integrated, necessary, and responsible activity within collection and records management, and collections care. With this mission in mind, this thesis seeks to answer these important research questions: How are U.S. museums handling the long-term accessibility and preservation of their many digital assets? Furthermore, are U.S. museums well-equipped to be prudent stewards of digital cultural heritage objects, records, and data by way of digital preservation plans and policies?

Chapter 2: Threats to Digital Objects

According to the law of entropy, every physical system naturally decays from order to disorder, and the loss of data is expected over time. The Museum, as a community and as an institution, has fought a long battle with entropy -- beginning with the very origins of the Museum, which has been represented through history as archive, protector of objects, and steward of our social memory. The concerns of decay may not always be discernable within a single human lifetime, but when considering the mission of institutions like museums, entropy commands a myriad of ruin. This same fight to slow entropy for physical museum collections is equally applicable to the contemporary conundrum of the management of digital collections and assets in museums.

Although in our modern era most people understand the fragility of digital materials, there has been a misconception that the breakthroughs of modern technology are the cure for the normal inefficiencies of analog materials: "Flawless computer memories! Lightening-fast chips! Fat fiber optics! Massive storage facilities! Bits not atoms! It's immaterial so it needs no preserving; it's escaped from the python coils of history; time harms it no more..."(Sterling 2003, 14). These common misconceptions about anything "digital" fail to recognize that digital objects have their own set of entropy-driven vices that will in time threaten any and all information, history, culture, and social memory that we choose to house in digital formats. Perhaps the singular difference between traditional museum preservation practices for analog objects, and that of digital objects, are the preservation concerns that go beyond any physicality of the digital medium (disks, tapes, CD's, etc). To further this point, consider how the significance of a physical artifact is indivisible from its material properties (Brown 2013). For example, it would be nonsensical to think of Leonardo da Vinci's *Mona Lisa* in an abstract form -- the artistic, historical, and cultural significance is ineluctably bound to its unique physical entity. The message and the medium are inseparable. However, in the digital realm, precisely the opposite is true. Digital media can be considered only the mere carrier of source code, which can be reproduced and represented on a plethora of digital media types without compromising its 'message.' Digital preservation therefore, addresses the special variability that digital objects

possess (Brown 2013, 195). This chapter seeks to identify the special nature of digital objects, which in turn elucidates the specific threats and inherent vices that endanger digital objects.

This chapter will also provide a basis for the argument towards the relevancy and importance of digital preservation for the museum community. Ultimately, within our digital world today, digital objects *do* form a part of all museum (library and archive) collections, whether that is in the form of accessioned objects or museum records. These digital objects have their own set of unique characteristics and conditions of care, which because of their immaterial nature are counterintuitive to traditional methods of collections management. Therefore, additional education is necessary to define digital objects and identify potential threats, which formulates the scaffolding needed to understand next how to approach digital preservation strategies.

Understanding Digital Objects and Their Conditions of Care

Digital information, objects, and collections are inherently immaterial (made of a series of bits composed of 1's and 0's), and therefore creates a content paradox that may be difficult to comprehend within the context of typical museum collection management. Since digital information is immaterial, it resists fixation; and thus defies traditional collection management which focuses on preservation of materials in their original, or most stable state (Smith 2004, 108). Even more confounding to traditional museum preservation, we cannot rely on preserving digital information or objects when it is simply manifested on physical mediums (Smith 2004, 108). For example, if one were to properly house a 17th-Century German lithograph in an archival box within an environment-controlled room, museum professionals can be confident that in 50 years time, the lithograph will remain physically viable. However, store a Compact Disc (CD) or portable hard drive on a shelf for 50 years, and the story will be entirely different; by this time the sensitive digital media may be succumbed to demagnetization or other forms of degradation. Even if the data survives intact, the technology required to read the storage medium will be long gone. Because of our fast-paced technological world, obsolescence becomes of imminent concern when maintaining current physical digital media, software, hardware, and software applications (Brown 2013, 206). An excellent example of this is the preservation tactic

known as "digital curation" which promotes the concept that preservation of digital data relies on an *active* and *iterative* management of the bitstream and digital environment (DCC 2015). Therefore suspending the lifespan of digital information within a static state on physical medium is merely a temporary solution, and avoids the other necessary activities required to make digital materials viable for the long term (DCC 2015). Recognizing the need for continuous and active intervention to preserve digital materials is a major step in understanding how the nature of digital objects shapes museums' approach to digital preservation.

When we talk about a digital object, it is crucial to identify whether the significant elements to the object are its hardware (storage) or the bitstream (object), the information itself (Ippolito and Rinehart 2014, 25). Within the scope of this research, the preservation of the bitstream and format will be the main focus; whereas the collection management of hardware manifests an entirely separate discussion, and perhaps could be a topic large enough for an entirely separate graduate thesis. Preservation of media storage carriers will only be discussed within this chapter in reference to the threats to digital assets.

Since a bitstream is not inherently 'human-readable', it contains no intrinsic meaning until it is extracted through the correct interpretation of that bitstream in accordance to some pre-existing program (Brown 2013, 200). For example, a digital image in TIFF format can only be rendered as an image using software that has been programed to interpret the bitstream in accordance with the TIFF format algorithm. It is important to recognize the distinction here: the bitstream is a *data object*, while its realization as a meaningful entity by way of the appropriate digital environment is termed as an *information object* (Brown 2013, 195). The process of transforming a data object into an information object is complicated; usually requiring the mediation of many levels of technology to access the source code, decipher the code, and present it to the user on a digital interface. Some combination of hardware and software is always needed to gain physical access to the bitstream, and therefore digital objects are entirely co-dependent on these entities (Brown 2013, 195). The Open Archival Information System Reference Model (to be discussed in depth in the subsequent chapter) refers to these various digital environment requirements as *representation information* (Brown 2013, 196; Magenta

Book 2012, 1-14). Understanding the specific 'representation information' needed to support digital 'information objects' is a fundamental digital preservation activity.

The separation of message from medium makes clear another unique property of digital materials: it is possible for the same digital object to be represented by more than one data object (Brown 2013, 196). Referring back to the example of a digital photograph in TIFF format, a second version may be created in JPEG format for access purposes. The actual image, and hence the conceptual information object, remains unchanged, although the JPEG format is encoded utterly different, so the technical representation has indeed changed. The TIFF and JPEG versions of the image are two different data objects, but they both generate the same information object. Furthermore, the same data object can be housed in a variety of storage mediums, as long as that medium supports the particular format that the object is encoded to. There is no one media that a data object must be stored within, so its very existence has a high threshold for variable existence.

Museum scholars John Ippolito and Richard Rinehart categorize the above quality of digital objects as *variability* (Ippolito and Rinehart 2014, 47). The very nature of the concept of variability goes against the more traditional doctrine of thought for museum collection management when we consider the preservation of digital museum objects. The goal for typical museum collection management revolves around sustaining the integrity and authenticity of the *original* object. However, in digital preservation we must recognize that the essence of a digital object is in tandem with its variability (Ippolito and Rinehart 2014, 47). It is the information object which we must preserve; we can change the sources and processes used to render it (representation information), as long as the essential performance can be replicated over time (Brown 2013, 196). While it is considered a standard for museum professionals to preserve any and all original components of a digital object, digital archivists also encourage the precept that this digital museum collection must be thought of as a variable, not static, entity. The familiar topic of variability in digital archiving will need to be adopted within the museum world to prevent its own 'digital dark age' (Harvey and Mahard 2015).

Significantly, digital objects call for the museum field to rethink its notions of originality (Ippolito and Rinehart 2014). Variability highlights that, for many types of digital objects, the

performance and information properties of a digital object, which a museum intends to have endure through technological change, are more fundamental and are more important than the specific technical properties of that object. Preserving the technical requirements without any plan for variability may have short-term value in how we approach preservation, but this tactic is transitory, much like the media itself (Brown 2013, 209). Most cultural institutions seek long-term preservation to care and manage all of its collections and institutional legacy, so to achieve this for the digital objects and assets in our institutions, we must look to the field of digital preservation, which will continue to be the focus of this thesis throughout the literature review chapters.

Indeed one can define the basic act of preservation as being concerned purely with preserving the quintessence of an object, which some may argue manifests in its original representation information (Brown 2013, 199). Museums in particular have believed that the continued survival of such "original" properties is fundamental to its authenticity (Harvey 2014, 18). However as discussed here, digital information has its own unique qualities within our material world. The nature of the immaterial objects belonging to cultural institutions will come with its own set of unique threats that, unless considered within the context of digital preservation theory, can be quickly lost within as short as a decade. With the basic understanding of the ephemeral nature of digital objects underway, it is equally important to fully understand the endangering factors of digital materials in order to better understand digital preservation tactics. The endangering factors that threaten digital objects to be discussed below are: *diffusivity, data obsolescence, physical degradation,* and *extrinsic threats.*

Threats to Digital Objects

Storage has been the default preservation strategy used by museum professionals as an expression to steward traditional and unique artifacts and archives. Using the appropriate best practices, stored cultural materials are bound in a form of suspended animation, protected from the elements. However, whereas storage is the longest-term strategy for traditional museum materials, it is the shortest-term solution for new media (Ippolito and Rinehart 2014, 8). The reason why simple storage solutions are only a short term solution for digital objects (whether

they be art objects, or digital photos of museum collections) is due to three *inherent* vices: diffusivity, data obsolescence, and physical degradation (Fino-Radin 2011, 8). These threats to digital material can manifest itself in a variety of ways, and are not considered mutually exclusive, for any digital object may experience more than one of these issues (Fino-Radin 2011). In addition to the inherent vices, there are many threats to digital objects that are *extrinsically* applied, including institutional managerial failure, human tampering, lack of recorded metadata, loss of human knowledge, and natural disasters.

Diffusivity

The first of the inherent vices, *diffusivity*, refers to works whose data is not contained simply within one object, but that references external databases or any dynamic and real-time data sources, such as an active web-crawling program (Fino-Radin 2011, 8). Diffusivity can also be problematic for preserving digital objects that do not exist solely in one location, but as a series of actions over a variety of locations and platforms (Fino-Radin 2011, 9). Of course there are plenty of examples of Internet based assets that are completely self contained, such as a domain name that points to a single page website. However, with the contemporary practices of linked open data, it is becoming more and more common for structural complexity that can create problems for museum archiving (Fino-Radin 2011, 9). For example, if a museum were interested in archiving an active public forum on its web-page, the many external hyperlinks and databases on that webpage that originate from a variety of sources can pose problems for archiving. One possible solution is to capture this internet-based asset via screenshots saved as .png or .jpeg formats. To what extent this suffices is contingent to one's definition of authenticity and experience. This particular concern is especially relevant for artworks that are digitally diffuse, in which authenticity of the experience and integrity of the artwork's conceptual significance is key to its preservation. A work that is diffuse presents a data structure that is antipodal to singular authority and ownership.

Obsolescence

Technological obsolescence is perhaps the most pervasive threat to digital objects and relates to both hardware and software components. It is inherent in all forms of digital assets and variable media artworks (Fino-Radin 2011, 10). The fact that digital material is mediated by technology poses a great problem in making data accessible unless there is appropriate hardware, and associated software which will make it intelligible. Contrary to typical museum objects, digital materials are always interdependent, and no element of digital materials are autonomous (Fino-Radin 2011). Because of the rapid rate of technological change, electronic materials may become inaccessible just a few years after they are created; formats become outdated and content may not be readable using new software (Corrado 2014). After CD-ROM technology was introduced in 1984, only four years later in 1988, CDs finally outsold vinyl records (MIT Libraries 2012b). Within that same year, many proprietary file formats proliferated in the tech world, which led to many competing word processing software and pushed many file formats into rapid obsolescence, such as .moo, .mic, .jbig, .cpx, .flan, etc. By 1990 most 2-inch videotape machines became obsolete; and by 1992 CDs outsold cassette tapes (MIT Libraries 2012b). Technological history continues on with this trend of file formats, media, and software rapidly changing, constantly being improved or newly reincarnated, and leaving old formats and media behind not even within a decade's time.

In addition, as we upgrade our hardware to newer, faster, larger-memory computers to replace poorly-working, slow, or defunct hardware, any media that runs on old software will be incompatible with the contemporary computers in our world. Because of our dependence on technology as more information (and even museum materials) goes digital, a growing volume of museum information is at great risk of loss if digital preservation is not taken seriously (Corrado 2014). For better or for worse, all digital artists and social memory institutions recording information digitally relies on the legacy of the technology industry (Corrado 2014). Without a call to action, technological obsolescence can quickly create a sense of mythos for any institutional legacy -- our materials can become an inaccessible history.

Many library institutions are already very attuned to this major threat, such as The Cornell University Library which offers a digital preservation management tutorial that includes

a resource called the "Chamber of Horrors: Obsolete and Endangered Media." Now hosted by ICPSR and MIT Libraries, this web document highlights a timeline of the digital hardware formats that are endangered or already obsolete (MIT Libraries 2012). Also Stanford University's Video Preservation Website offers a reference list that categorized digital media similarly to an endangered species list, from extinct media to vulnerable media (VPW 2015). The timeline on this "endangered species list" only goes back to 1956 when video recording became a viable technology; most of the digital file formats before the 1970s are now extinct (VPW, 2015). Although these tools focus on physical hardware, the issue of obsolescence is nonetheless a point well made. It is important to remember that software is just as likely to become antiquated, perhaps even more quickly than hardware in the face of the constant technology updates occurring within today's industry. Obsolescence in the digital world can risk the integrity and usability of information, two major goals of any preservation program.

Most individuals may have experienced the devastating realization that a 3.5" floppy disc found in one's home or office is no longer easily accessible on today's computers. However there are still many companies that can retrieve data from obsolete media for a fee. Even though such "digital paleography" exists, obsolescence should be taken extremely seriously, especially when critical data is at risk. For example, in the 1980's the British Broadcasting Corporation (BBC) created a project to collect fragments of life and culture from across the U.K. into a single collection to honor the 900th anniversary of William the Conqueror's *Domesday Book*, which housed the records of 11th Century life from over 13,000 towns in England. This new project, called the Domesday Project, eventually became the central repository of over a million British contributions. In addition to having many statistical databases, there were tens of thousands of digital photographs and interactive maps (Cohen and Rosenzweig 2005). Since this huge multimedia collection required a high-density, fully modern format to capture the entirety of its data, the BBC decided to encode the collection on two special videodiscs, accessible only on specially configured Philips LaserVision players with a BBC Master Microcomputer or a research Machines Nimbus. Of course by the late 1990's, the LaserVision, the BBC line of computers, and the Nimbus had all become obsolete; and this rich historical collection faced the imminent threat of being unusable except on a few rare functioning computers with the correct

hardware and software translators. Ironically the original Domesday Book vellum has withstood nine centuries intact and perfectly readable (Cohen and Rosenzweig 2005). In the end, some programmers from the University of Michigan and the University of Leeds were able to figure out how to reproduce the necessary computing environment on a standard PC by 2003, and so the Domesday videodiscs have gotten a reprieve, at least for a few more years or decades if lucky. However this solution did not come without panic and a considerable expense to safeguard it after almost realizing it could be too late to save the data.

A similar project was conducted through the US Census Bureau to ensure continued access to the 1960 census, which was recorded on long-outdated computer tapes; while the government can surmount such major engineering challenges, an individual archivist, museum data manager, or even some major museums will probably not foot similar bills for their own digital collections (Cohen and Rosenzweig 2005).

Physical Degradation

The third inherent vice of digital objects, degradation, can be considered for both the immaterial data object and for the physical storage media. Physical degradation refers to the deterioration of any physical component of a digital object (Fino-Radin 2011, 12). As quoted from an early digital preservation study from the Research Libraries Group in 1998:

> *"Digital materials are especially vulnerable to loss and destruction because they are stored on fragile magnetic and optical media that deteriorate rapidly and that can fail suddenly from exposure to heat, humidity, airborne contaminants, or faulty reading and writing devices"* (Hedstrom and Montgomery 1998, 1).

Many are familiar with the gradual process of decay for most physical objects; for analog materials the process of loss of content is a slow erosion, such as a manuscript fading slowly over time. As the Digital Preservation Coalition asserts, the rate of degradation is quite different for digital media; even though such media is made with more industrial materials, it is quite delicate compared to other museum objects and archive materials. Typical materials and artifacts in museums are not destroyed in a single moment, except for catastrophic disaster such as fire. However, instantaneous and complete loss is the norm for digital data. Since the decay profile for digital data is considered binary, it typically has two possible states: "readable" and

"unreadable" (Brown 2013, 200). Digital Archivist Adrian Brown succinctly explains this binary relationship between digital data and its medium:

> *"Although the physical medium on which the data is stored may decay gradually, there will typically be a single point along that path of physical degradation at which the information content will flip from being completely readable to utterly lost"*(Brown 2013, 200).

There are of course some exceptions; in some cases data recovery experts may be able to retrieve some intact data from a damaged disk or drive, however there is usually some amount of loss in the integrity of the information regardless (Harvey 2014, 62). An example of the fragility of digital media can be found within one of the more common media formats still used today: Compact Discs (CDs). In 2003, NIST researcher Fred R. Byers estimated a variable of 20 to 200 years lifespan for media like the CD or DVD, and even the low end of this estimate may only be possible under ideal environmental conditions (Byers 2003). A significant fraction of collections from the 1980's of audio CDs may already be unplayable. For example, the Library of Congress, which has more than 150,000 audio CD's, is able to store its digital media in conditions far better than those of smaller cultural institutions; however the Library of Congress still estimates between 1 and 10 percent of the discs in their domain already contain serious data errors that render it unreadable (Cohen and Rosenzweig 2005). With only a few exceptions, digital formats tend to require an exceedingly high degree of integrity assurance in order to function properly. In an ironic way, the perfection of digital media is also its imperfection: they are encoded in a precise fashion that allows for unlimited identical copies, but any minute amount of loss of the original can mean disaster (Cohen and Rosenzweig 2005).

Extrinsic Threats

When considering the threats to digital materials, it is equally important to acknowledge that digital preservation is more than just a technological challenge and has many extrinsic threats. The organizational and social issues associated with digital preservation are just, if not more, important than the technology (Hirtle 2003, 135). Institutional managerial failure is therefore a major extrinsic threat to our digital collections. This is becoming more and more recognized within the field.

In the beginning of digital preservation efforts, most of the attention was given to technology as both the root of the problem and the basis for the solution (Kenney and McGovern 2003). This emphasis is undeniably important, but does have its downsides. Much energy has gone into advocating for one technology over another, notably evidenced in the data migration vs. emulation debate (Kenney and McGovern, 2003). The focus on technology has led to a equating technology with solution, which is inherently problematic when technology is constantly in flux. Even when convincing technology solutions are at hand -- D-Space, for example, is being characterized as a sustainable solution that "enables easy and open access to all types of digital content including text, images, moving images, mpegs and data sets" -- it is still important to maintain that technology is only part of the real solution. One can say that the focus on technology has mimicked computational methods that simplify things to an *on* or *off* status; either you have a solution or you do not. However this either/or type assessment gives little room to consider the effort required to reach the *on* state, nor to differences in institutional settings (Kenney and McGovern 2003). It also does not take into account that a partial program at one institution may represent a fully mature program at another. Some organizations may only ever need to preserve a limited range of formats, or may progress in stages to expand its capabilities to all formats. Unsurprisingly, in light of this fact many organizations have been left uncertain as to how to proceed (Kenney and McGovern 2003).

While there may be no true universal solution for cultural institutions, if an organization cannot even imagine how to start, this may explain why so few museums have done so. However it is extremely unwise to continue any postponing of the development of a digital preservation program, for many vital digital resources will be sacrificed at the interim. Consider a study conducted at UC Berkeley 12 years ago that estimated 93% of the world's yearly intellectual output had been produced in digital form (UC Berkeley 2003). Considering how outdated this survey is today, one can only imagine how this percentage has increased by 2015 and therefore the sheer increase on our dependency of digital materials. Despite the increasing evidence on the fragility and ubiquity of digital content, cultural repositories have been slow to respond to the need to safeguard cultural heritage materials in digital formats (Kenney and McGovern 2003). In the end, productive work in digital preservation will need to start with the

commitment of our cultural institutions to develop and maintain a program. So much of this need is centered on the fact that in many ways digital preservation is a *management* issue. As stated within the Trusted Digital Repository Audit and Certification Criteria and Checklist, the first and foremost elements of a trustworthy digital repository are concerned with governance, financial sustainability, and legal issues -- all of which are management related (RLG 2007). Without higher-level institutional commitment for plans, policies, overhead budget, and staff, the technological considerations need not be an issue for there will be no long-term digital preservation possible in the first place. Digital preservation is not something that can be done once and then be forgotten, but requires consistent, and ongoing follow-through by the parent organizations of digital collections who will be the providers of many of the basic elements that allow a trustworthy digital repository to exist (Corrado 2014, 5).

After a comprehensive survey in 1998 through the RLG/OCLC Working Group, digital archivist and librarian Margaret Hedstrom has spoken of a real issue for effective digital preservation: there is a "gap between current guidance on digital preservation and institutional capacities to follow through"(Hedstrom and Montgomery 1998, 29). While work has certainly been done to reduce this gap since Hedstrom's analysis, one can also argue that the gap has not closed much at all for the museum field. The reason for this lag in institutional take up? One must understand the organizational impediments to digital preservation practice in order to recognize what museums are lacking. These impediments are typically lack of knowledge within an institution, lack of funding, lack of personnel, and lack of institutional mandate; all of these issues will add up to institutional managerial failure to safeguard digital assets if action is not taken sooner than later. Thus any institution who owns digital material that requires long-term preservation will need a clear administrative mandate to lead such activities, and the financial sustainability to continue. (Hirtle 2003).

Two other important extrinsic threats to note are that of recorded metadata and human knowledge (legacy knowledge). Descriptive metadata (literally "data about data") is essential for identifying and retrieving digital assets that would otherwise appear as foreign entities to any user. For example a digital photograph saved in the Tiff format is automatically assigned a

generic name upon creation or ingest: e.g, 0145897.tiff. This photograph is really only easily identifiable to a user if someone were to administer a more descriptive name to it: e.g. museumlaunchparty_2015.tiff. Besides simply naming digital objects, metadata is a far more complicated and robust concept that includes descriptive data about who created an object, keywords that identify the object, artist name, historical context, place names, date taken, institution, copyright, etc. In addition metadata can be recorded as a kind of "guide book" for an asset's technical requirements such as software, digital architecture, applications, hardware specifications, etc. This is called Technical Metadata (Corrado 2014, 114). Since creation of metadata can be extremely variable, entire schemas have been created to help standardize and guide what metadata to capture, and to provide a uniform language that defines entries and modifiers so many institutions can understand and share data interchangeably. Much of metadata capture can happen automatically, such as that captured by modern digital cameras, by way of software. However for cultural institutions, much of this work requires a human touch in order to identify persons, places, or things that may be represented by the digital object (Corrado 2014, 114). Thinking in the long-term, without the necessary descriptive, administrative, and technical metadata recorded, users of the future may find it virtually impossible to identify, render, or use a data object to any effect. More on the function and importance of metadata capture as a roadmap to accessing and using digital materials will be discussed in the subsequent chapters.

Oftentimes the maintenance of digital materials can be completely reliant on limited, or sometimes on a single person's knowledge. Thus the loss of human knowledge is always an imminent threat to any digital asset. For example the Ivar Aasen Centre of Language and Culture, a literary museum in Norway, lost its ability to use a large and expensive electronic digital catalog after the death of one administrator who was the sole keeper of two sequential passwords required to access the system (Cohen and Rosenzweig 2005). The catalog was an invaluable research tool stored in an encrypted database format, had taken the museum four years to create, and contained over 11,000 entries. The Centre desperately and unsuccessfully tried to break into the system themselves, but had to resort to an expensive open call to computer experts to hack the system (Cohen and Rosenzweig 2005). Although ultimately the problem was remedied, the panic and fear of lost data could have been avoided with proper documentation

procedures in place at the museum. Such procedures are essential in any digital preservation policy or plan, and without which staff may be granted unnecessary autonomy (CHIN 2013). In addition, metadata capture can be utilized as a vehicle for recording certain exclusive and unique administrative information to prevent the immediate threat of human knowledge loss.

As a parting topic, the more familiar threats of natural disasters and human tampering are just as applicable to digital collections as normal physical museum collections. Just as museums have well-thought out plans for safeguarding physical archives and collections during a natural disaster, a similar plan ought to be in place for digital assets (Corrado 2014, 21). Such plans often incorporate some combination of maintaining a master copy, or offsite backup system that mitigates the location-specific loss of data during a disaster (Harvey 2014, 313). Backing up data is an excellent method for mitigating human tampering as well; if data were to be accidentally deleted or modified, the hope is that a master or backup copy would be retrievable in order to recall that lost data. Checksums are also effective tool to use in tandem with a backup system to aid in identifying loss of data integrity. Checksums are essentially algorithms programmed to sum the binary code of a digital object; should even one bit be missing or out of place, the binary summation would not match the "master," allowing the system to identify data errata that could be the result of human tampering, or sometimes bitrot (Corrado 2014, 130). It is certainly more common for cultural institutions to have a regular backup protocol for its computer servers, and while this is effective for mitigating certain threats (namely natural disaster and human tampering), it is still important to recognize that backup systems are only one, single level of digital preservation.

Conclusion

With the massive increase of dependency and use of digitized or "born-digital" materials in our world, a fundamental challenge facing cultural institutions today is to preserve the accessibility and authenticity of digital objects over time, various domains, and changing technical environments. Cultural institutions, while less equipped with the financial and personnel stability of the business and tech world, need to accept the inevitability of change and separation of logical information objects from its physical environment in order not lose any

important history, art, data and money invested on digital formats. The causes for data loss ranges from the inherent risks that all digital materials possess, to causes that are in our control, or that are extrinsically applied. Ultimately for the museum field, digital objects exist within their own category for collection management, with their own set of recommended strategies and requirements in order to be considered safeguarded for future access and use. The collection management and conservation work that can be employed for digital collections will be discussed in the following chapter.

Chapter 3: What is Digital Preservation? Digital Preservation Basics

This chapter will outline a literature review and discussion on the basics of digital preservation within the following broad categories: *defining digital preservation, history of digital preservation, fundamental digital preservation concepts, common steps in digital preservation, digital preservation methods,* and *useful tools*. This chapter provides foundations for implementing digital preservation as well as for understanding the digital preservation activity of the three case study institutions.

Introduction

The preservation ethic and mandate at many museums, along with libraries and archives, naturally endorses the American Institute for Conservation of Historic and Artistic Works' statement that "every institution has a responsibility to safeguard the collections that are entrusted to it. That responsibility includes incorporating preservation and conservation awareness into all facets of the institution's activities so as to ensure the long-term preservation of its collections" (AIC 2002, 1). Within museology, similar concepts are considered a central and immutable ethical obligation and a philosophical approach to prudent collections care. For example, museum scholar Marie Malaro insists that the Museum's central goal, as a part of the nonprofit sector and as an authority in cultural stewardship, needs to uphold the public trust in regards to service to the community and ethical handling of its assets (Malaro 1994, 3-15). When considering Malaro's assessment of the museum's role in society, it is clear that devoting time, finances, and plans for digital preservation will be a core expression of key responsibility as museums increasingly acquire digital materials.

In order to prevent the many problems revolving around the sustainability and viability of digital museum materials, whether that includes archives, databases, exhibit materials, or collections, active digital preservation tactics should be accessible, manageable, and realistic solutions. Indeed, a primary issue that needs to be evaluated is how many digital materials are created in museums that are worth preserving for the long-term. These numbers will vary from institution to institution, but there will undoubtedly be an increasing number in our not-so-distant

future. Since the advent of accessible personal computing and the World Wide Web, the world has undergone an immense paradigm shift towards dependence on technology. In an effort to stay relevant in our changing society, museums have in turn embraced this paradigm shift by modernizing museums with technology, whether that is interactive technology in galleries, showcasing film, creating phone apps, using electronic databases, documenting collections with digital photography, recording oral histories, scanning slides, etc. In addition, many art museums are acquiring art that is "born-digital" that requires special preservation considerations in order to be maintained for the art historical canon. For the museum world, the digital materials can range from that used for the public-facing museum, that used for academic research, and that used to streamline museum collection and employee processes behind the scenes. Whether the museum assets in question are important unique objects or digital tools, within a very short period of time, museums have acquired a burgeoning collection of vital materials that have been created, stored and transmitted in digital form, yet they are often housed in cultural institutions that are ill-equipped to uphold their duty to preserve them.

Defining Digital Preservation

One rudimentary problem in understanding what digital preservation entails are the numerous definitions that can be found within the field. At the same time, digital preservation is a relatively new discipline, and as such is a fertile breeding ground for a specialized nomenclature that has yet to mature and settle (Brown 2013, 12). For members of the museum community new to digital preservation, the terminology can appear foreign and confusing as a number of alternative terms are often applied to the same, or similar, concepts. Furthermore, digital preservation amongst cultural memory institutions bridges many long-established fields that have been traditionally kept separate from each other, each with their own unique vocabularies (Brown 2013, 12). Currently there is no available definition for digital preservation within the museum context. This makes delineating or defining digital preservation in the museum field that much more difficult, although as discussed throughout this thesis, the similarities between the mission of library science and museum studies makes definitions found within the library field well suited for the museum context.

Through a literature review, three main definitions from the digital preservation community were chosen: the Library of Congress, the Digital Preservation Coalition, and the JISC *Beginner's Guide to Digital Preservation.* These three definitions are general enough to be applicable to the museum context. In addition, the discussion of defining 'digital preservation' warrants some explanation of the terminology 'digital curation' and 'digital stewardship,' and how these terms relate to the greater umbrella of digital preservation.

Library of Congress, Digital Preservation Coalition, and JISC

The Library of Congress defines digital preservation as "the active management of digital content over time to ensure ongoing access" (LoC 2015a). The simplicity of this definition can appeal to a broad range of disciplines including museums, universities, research centers, etc. Although this definition comes specifically from the library field, which has long been committed to promoting access of information to the public, one can argue that museums similarly strive for continued access in the form of *object-level* preservation for the people, and thus aligns with similar preservation goals of the library field. Ultimately, digital content in museums is just as important to preserve for ongoing access as library materials. In as such, the definition proposed by the Library of Congress can be recognized as a universally relevant definition to all cultural memory institutions, and for the specific purposes of this paper, museums.

Similarly, the Digital Preservation Coalition (DPC) from the UK defines digital preservation as:

> *"[a]series of managed activities necessary to ensure continued access to digital materials for as long as necessary. Digital preservation is defined very broadly for the purposes of this study and refers to all of the actions required to maintain access to digital materials beyond the limits of media failure or technological change. Those materials may be records created during the day-to-day business of an organization; "born-digital" materials created for a specific purpose; or the products of digitization projects"* (DPC 2015).

This definition nicely names the variety of digital content that cultural institutions share an interest in preserving. These materials worth maintaining for viability and accessibility can

range from content that helps promote and support the institution itself, all the way to unique digital collections that are kept under the public trust. A few keywords to be noted between the DPC and LoC definitions are "management" as well as "ongoing access." These key phrases are singularly important in truly understanding the necessary environment for effective digital preservation (Corrado and Moulaison 2014, 10).

Another important digital preservation definition comes from the JISC (Joint Information Systems Committee) in the UK, which elaborates on the DPC definition above by highlighting and explaining five aspects. These five keywords that distill the definition are *managed*, *activities*, *necessary*, *continued access*, and *digital materials* (JISC 2012). The JISC uses these five keywords as an effective way to dissect and make tangible the elements of digital preservation, as explained below.

The concept of "managed" is certainly the most important, and the most shared element in any digital preservation definition (Corrado and Moulaison 2014, 7). According to the JISC *Beginner's Guide to Digital Preservation*, digital preservation at its core is a *managerial* problem (JISC 2012). If our institutions are to take preservation of digital materials seriously, all digital preservation projects need to have its activities (planning, resource allocation, use of technologies, etc) to be properly managed, and require support from upper administrators in order to be successful (JISC 2012). The term *managed* stresses the need for policy, which is a major focus for this research project. The *activities* that need proper managing refer to the certain activities that need to take place in order to ensure ongoing access, such as ingest, migration, fixity checks, checksums, normalization, etc. These activities should be broken down to individual tasks that can be performed in well-defined ways. In addition, these activities and their corresponding tasks should be well documented so that someone else can perform them if necessary (JISC 2012).

The term *necessary* highlights the act of prioritizing what needs to be done (JISC 2012). Not all objects will need preserving, and those that do will inevitably require some kind of prioritizing system to determine which materials are addressed first. In addition, not all digital content will require the same degree of preservation which is another essential part of

understanding what "needs to be done" (JISC 2012). Such selection protocols are typically outlined within an institution's digital preservation action plan (Corrado 2014).

Continued access is an especially vital element in any preservation efforts. This key term is given more importance simply because in order to have continued access, inevitably there are a series of *activities* that require ongoing *management* in order to keep digital content *alive* in the face of our rapidly changing digital world. How long access is needed will be an ongoing topic as well, and surely will vary from digital object to object; such specifications should also be defined within one's policy (JISC 2012). The expanded definition provided by JISC exemplifies how closely linked and in some cases dependent, are the elements of digital preservation.

The last term, *digital materials* is a broad way to encompass "the stuff" you are preserving, whether this is also coined as digital materials or digital objects, digital content, data, etc (JISC 2012). For the sake of this thesis, the terms digital materials, digital objects, and digital content will be used interchangeably. Some examples of categories of digital materials/objects/content relevant to the museum field include images, datasets, audio recordings, videos, scanned archives (such as books, newspaper articles, primary sources, etc), databases, emails, websites, digital documents, institutional records, and digital artwork. Most museums can soundly claim to have the need to preserve at least one of these types of digital objects. While many of these categories may require specialized considerations within the preservation process (especially digital artwork), the foundations of digital preservation can still be applied to all; this includes the applicability of any basic policy.

The JISC Beginner's Guide to Digital Preservation expands its definition to quantify preservation within three different lengths of time: long-term, medium-term, and short-term preservation (JISC 2012). Each has its own requirements of a preservation repository. Long-term preservation is when continued access to digital objects is required indefinitely, or at least to the information contained within them. This is the most challenging of the three. Medium-term preservation is when continued access to digital materials is desired beyond changes in technology within a defined period of time, but not indefinitely. Short-term preservation is when continued access to digital materials is needed, but does not extend beyond the foreseeable

future, and/or until it becomes inaccessible because of changes in technology (JISC 2012). For libraries, archives, and museums, most of the digital assets will typically require long-term preservation, or at least medium-term preservation depending upon whether the digital object is part of the mission to preserve cultural memory, versus objects that are more utilitarian (Corrado and Moulaison 2014, 6).

Three Key Terms: Curation, Stewardship, and Preservation

In addition to these basic understandings of digital preservation, in place of the word *preservation* we may often see the terms *curation* or *stewardship*; the seemingly interchangeable use of these three key terms in digital preservation literature furthers the confusion around what digital preservation exactly entails. Is it the same or different from digital curation or digital stewardship? (Harvey and Mahard 2014, 7). *Curation* and *stewardship* are less associated with physical objects compared to the connotation of *preservation*, which works appropriately with the intangibility of digital materials (Harvey and Mahard 2014, 7). More importantly however, these terms encourage a wider view of digital preservation as not just a set of technical processes, but also services, policies, and stakeholders from across disciplinary boundaries -- such as libraries, archives, and museums combined into a trifecta of cultural memory institutions (Harvey and Mahard 2014, 8). As will be discussed later in this chapter, certain standards within the digital archiving and long-term preservation field, such as the OAIS Reference Model, were concerned with providing definitions that can apply to a wide range of disciplines, and thus sought to select terms that were not already heavily entrenched in any one discipline. However, despite the recognized importance of standardized, clear definitions, there is currently a lack of consensus (Harvey and Mahard 2014, 8).

Curation is commonly used today, although its meaning is not widely agreed upon. When referring to data curation or digital curation, what should automatically accompany is a life-cycle model that describes how digital objects are managed over time to ensure preservation; hence a unique quality for curation is the emphasis upon *cyclical, iterative* activities associated with creating digital objects, selection and appraisal, and enhancing digital objects for use and reuse (Harvey and Mahard 2014, 8). A major player in the development of the concept for digital curation is the DCC or Digital Curation Centre from the United Kingdom,

which has created the concept of the digital curation lifecycle model (See Figure 1). The DCC was created in 2004 with the goal to respond to the needs of managing large quantities of scientific research data from major universities in the UK (DCC 2004a). The science and e-research communities that formed the DCC have been more driven by immediate re-use of data, as opposed to the concerns of longevity that museums are primarily striving towards (Lazorchak 2011).

Although the target groups of the DCC are not cultural institutions, the general concepts produced by the DCC have been accepted as useful tools within the greater digital preservation community. According to the DCC, digital curation composes of eleven steps: conceptualize, create, access and use, appraise and select, dispose, ingest, preservation action, reappraise, store, access and reuse, and transform (DCC 2004b). In essence, digital curation is the very process of enacting the many *preservation activities* necessary to achieve the final goal of digital preservation, but it is not another term to be used interchangeably with "digital preservation," which has more of a high-level, all-encompassing meaning. Curation emphasizes adding value to data, for example through metadata annotations to enhance reuse. Significantly, the cycle of data curation may not always prove appropriate for some of the digital assets that museums may be interested in preserving. For example, digital art collections require a more diligent preservation process that does not alter or compromise the original artist's intent. For this reason it makes sense then that the DCC's Curation Lifecycle Model is usually associated with science and social science data. Ultimately, *curation* is a useful concept that clarifies the evolving "whole-life view" of digital preservation, but its concentration on the underpinning activities of building and managing collections of digital assets does not fully describe a broader approach to digital materials management (Lazorchak 2011) that many museums will need.

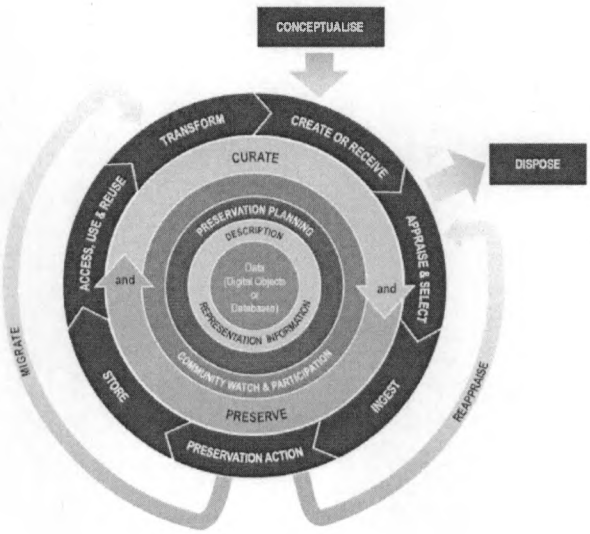


Figure 1: Digital Curation Lifecycle Model (DCC 2015)

Digital stewardship is a broader concept than curation, and even that of preservation, which includes both the technical processes and overarching elements such as services, policies, and stakeholders (Corrado and Moulaison 2014, 11). For others, the term preservation is viewed *only* as technical processes, and therefore as isolated from underlying managerial and big picture elements (Harvey and Mahard 2014, 7). *Stewardship,* on the other hand, originally evolved out of the environmental community, but has been adopted by the nonprofit sector and cultural institutions to instill the idea of holding resources in trust for future generations (Lazorchak 2011; Malaro and DeAngelis 2012, 7).

The museum field takes the concept of stewardship of humankind's collections for the public trust very seriously (Malaro and DeAngelis 2012, 8). Stewardship and trust are some of the very backbones for modern museum collection management theory within both an in-house and governance domain (Malaro and DeAngelis 2012, 8). As the range of institutions using digital formats to store information increases, the concept of ethical stewardship equally becomes a shared idea between disciplines. Priscilla Caplan of the information science field notes: "Institutions exercise stewardship, individuals curate or manage data…If you have stewardship of something, you don't dump it in the bit-bucket when your funded research project ends" (Harvey and Mahard 2014, 8). Similar if not identical to stewardship within a

museological context (Malaro and DeAngelis 2012), a strong sense of duty is associated with the practice of stewardship, in that it is a necessary duty of everyone involved in managing objects, in this context: digital objects. Stewardship is truly the responsibility of everyone in the community -- from the creator of the digital object to the curator, the user, and everyone in between (Harvey and Mahard 2014, 8).In light of these terms, *digital stewardship* is becoming more accepted as a term that incorporates the concepts of preservation and curation together, including the technical processes, as well as the lifecycle approach, while also emphasizing duty, preservation, and management as core components of action (Lazorchak 2011).

Information scientists Ross Harvey and Martha Mahard view the terms preservation, curation, and stewardship not as interchangeable but within a hierarchical structure (Harvey and Mahard 2014, 9). From this perspective, *digital stewardship* encompasses the "cultural, public policy, and ethical questions about how and what we remember and forget (relation to information in digital form). Digital stewardship [includes] the full range of preservation practices and issues applied by information professionals, who have the obligation of keeping collections and the objects in them in trust for future generations" (Harvey and Mahard 2014, 9). As such, digital stewardship subsumes digital curation, and digital curation subsumes digital preservation (Harvey and Mahard 2014). This may be a useful way of categorizing these commonly used terms in the field. While the term *digital stewardship* may closely resonate with the museum field, it is still much more common to hear and use the term *digital preservation* to mean *digital stewardship*. For the purposes of this thesis, the more common term *digital preservation* will be used here to emphasize the concepts of trust, duty, management, and technical processes that can also be associated with the term *stewardship*.

History of Digital Preservation:

The history of digital preservation is important to consider in understanding the development and importance of digital preservation as it pertains to cultural memory institutions, such as museums. This section is only meant to be a succinct exposure to the many pioneers and projects that have promoted and built the digital preservation field.

In evaluating the history and state of digital preservation in the United States in 2008, librarian and archivist Peter Hirtle noted that the earliest reference that he could find in English of the digital preservation of data was within the context of the research that Anne Kenney and Lynne Personnius undertook in 1990 at the Cornell University Library in conjunction with the Xerox Corporation (Hirtle 2008, 125). In this research, "digital preservation" meant using digital technologies to reformat analog media as a way to preserve those media (Hirtle 2008, 125). As Hirtle points out, the earliest concepts of digital preservation were generally focused on the digitization of collections. While it is accurate to think that creating a digital surrogate of an object, whether that is a high-resolution photograph or a more complicated 3D data set, is a way to preserve the documentation of that object's existence, this is not truly "digital preservation" as understood in the field today (Corrado and Moulaison 2014, 4).

The prevalence of digitization projects in museums is certainly increasing, but what digital preservation is *truly* concerned with is the process of maintaining the digital asset after it has been created in order to keep it usable and retrievable for as long as needed. For example, if a museum has a series of high-resolution digital photographs that have been taken of a collection, or a library of references that has been scanned to digital PDFa files, these materials can only be considered "preserved" if they can be usable beyond technological and format obsolescence (as quickly as 5-10 years) if the bitstream has been maintained, along with its digital environment. Digitizing museum collections is often a first step in digital preservation, but is not in of itself enough to truly preserve our cultural memory within the digital age. As noted in the previous chapter, the major threats to digital objects can include technological obsolescence, lack of metadata, or loss of human knowledge to run the necessary digital architecture. It is therefore useful to think of digital preservation as the "preservation of digital information," which the Society of American Archivists explains "...is not so much about protecting physical objects as about specifying the creation and maintenance of intangible electronic files whose intellectual integrity is their primary characteristic" (SAA 1997). There has been a short, but very progressive history that has led up to this way of thinking about digital preservation, as outlined below.

Early Digital Preservation

Interest in the maintaining the longevity of digital information have been evident since the early years of the digital information age in the 1960s. Even as computers were first being integrated into society, the fear of a "digital dark age" was already recognized (Brown 2013, 9). The first data archives were established in the 1960s and were designed for the scientific research data field with the goal to make research data accessible to the scholarly community (Brown 2013, 9). Archives such as the Inter-University Consortium for Political and Social Research at the University of Michigan (1962), the UK Data Archive (1967), and the Machine-Readable Records Branch at the National Archives (1960s) laid much of the groundwork for management practices of digital assets (Brown 2013, 9; Hirtle 2008, 124). At this early stage of digital archiving, these programs mostly did not use the term "digital preservation,", but referred to the archiving of electronic records or data sets (Hirtle 2008, 124).

A New Digital Age: 1980s - 1990s

The advent of personal computers occurred in earnest in the 1980s and with the emergence of the World Wide Web in the 1990s both an explosion in the creation and use of digital materials was triggered that has only increased ever since. Using computers and the internet had gradually migrated from being only used by big business and major research data institutions to becoming a fact of everyday life worldwide (Brown 2013, 9). As technology has grown more accessible, most jobs also came to involve use of a computer and the internet in some form or capacity. To facilitate work and to stay current in the world, cultural institutions have also slowly adapted to the shift from using paper to "going digital" for many aspects of their work.

A plethora of new kinds of digital information have also come out of this major change in society; we now use computers for everything from office documents, to multimedia, to web pages, to 3D models, to databases, to emails, to ebooks, etc. The sheer increase in the dependence and number of digital materials used in our world was perhaps the original stimulus for digital preservation as we know it today. Reformatted information, or what became known as "re-born digital" objects, were the very beginning of early digital preservation initiatives (Hirtle 2008, 124). These "re-born digital" materials were originally developed in libraries (not archives

or museums) as a way to address ongoing analog preservation efforts, especially when materials like discs or CD's were submitted to collections (Hirtle 2008, 125).

A turning point in global awareness about the fragility of digital information crystallized in the formation of the 1994 Task Force on Archiving of Digital Information (Corrado and Moulaison 2014, 97). After two years of deliberation, this U.S-based group distributed a seminal report in 1996 on the concerns and future of digital viability. Significantly, this document has laid the foundations for most of the subsequent work in the field (cultural, government, and business alike) and continues to shape the agenda even today (Brown 2013, 9). According to the final report, the Task Force sought to frame the key problems (organizational, technological, legal, economic, etc.) where resolutions were needed in order for "technology refreshing" (such as fixity checks or migration) to be considered an acceptable approach to ensure continuing access to digital records indefinitely (Task Force 1996). Based on this analysis, the Task Force recommended actions and alternatives to technology refreshing. Perhaps one of the most important conclusions made by the Task Force was that around the concept of establishing trustworthy digital repositories/stewardship:

> *"The Task Force sees repositories of digital information as held together in a national archival system primarily through the operation of two essential mechanisms. First, repositories claiming to serve an archival function must be able to prove that they are who they say they are by meeting or exceeding the standards and criteria of an independently-administered program for archival certification. Second, certified digital archives will have available to them a critical fail-safe mechanism. Such a mechanism, supported by organizational will, economic means and legal right, would enable a certified archival repository to exercise an aggressive rescue function to save culturally significant digital information"* (Task Force 1996, iii).

According to its charge, the Task Force identified an imminent need for a digital repository certification process that would address the range of activities, functions, and responsibilities associated with repositories, while providing layers of trust for all involved. Furthermore, the need for official plans and policies for digital preservation was an important conclusion from the Task Force, and the report urged that official processes be established within institutions in order to guarantee long-term preservation (Task Force 1996). In order for

digital preservation to not be approached like a popular fad, the Task Force called 'for full institutional commitment'(Task Force 1996).

Another seminal group in the beginning of widespread digital preservation awareness was the Commission on Preservation and Access/Research Libraries Group (CPA/RLG) Task Force on Archiving Digital Information, which also came together also in 1994 (Hirtle 2013, 125). Unlike the Task Force on Archiving of Digital Information, the CPA/RLG Task Force did not include any members from the museum field, and was primarily associated with libraries and archives. The concepts and concerns that both of these mid-1990s groups addressed, such as certification of trusted digital repositories, format registries, cost models, and authenticity, remain a major focus of discussion even today (Brown 2013, 10).

Building on the recommendations of the Task Force on Archiving Digital Information, digital preservation today is the focus of a large, active, and collaborative community. Yet it is still considered an emerging discipline. Interestingly, the Task Force recommended two strands of future activity: first, the development of strong theoretical underpinnings and standards, and second, the establishment of a diverse and active pool of practitioners to advance and expand the theory through practical application (Task Force 1996, 40). Much of the literature review in this chapter represents success of the first recommendation. An example of success of the second, is the development of the PDF/A format. Since the Portable Document Format (PDF) was being used frequently as the de facto preservation format, Adobe led an effort to have an ISO committee develop a PDF specification for archival needs known as PDF/A (with the "A" standing for "archive") (Hirtle 2013, 133). While there will always be other proprietary formats in the tech world, the creation of PDF/A came from an already widely practiced habit in the archiving community. As a result, a step towards a more global-wide recognition of digital archiving standards was taken, which in turn makes digital preservation efforts more widely applicable across many different fields.

Another example of an important standard that developed after the Task Force report was published is the Open Archival Information System (OAIS) Reference Model (Magenta Book 2012). Originally developed for the space science community in the 1990s, and released as a recommendation by the Consultative Committee for Space Data Systems in 1996, the OAIS

Reference Model quickly became the accepted de facto standard for a conceptual framework for digital preservation (Corrado and Moulaison 2014, 44). The OAIS Reference Model was formally published in 2002, and was later issued as an international standard (ISO 14721:2003), and most recently updated in 2012 as the Magenta Book 2 version (ISO 14271: 2012) (Brown 2013, 10). To develop OAIS, the CCSDS conducted many open discussions with a variety of stakeholders (including social memory institutions), which is where the "Open" part of the name came from (Ockerbloom 2008). It is important to note that OAIS does not require a repository have "open access" or "open architecture," and it has no direct relation to the similarly-acronymed Open Archives Initiative (OAI) (Ockerbloom 2008). In addition, the use of the term "archival information system," or "archive," can be thought of a way of defining any entity that is responsible for long-term preservation of the information it manages; it was not created with archives specifically in mind (Ockerbloom 2008). The OAIS is ultimately a conceptual model for what digital repositories should do and can be adapted to many different operational digital preservation services/software. Importantly, however, is the presence of a universal system for digital preservation, which was strongly lacking before 1996.

Recent Digital Preservation History

Moving into the 2000s through today, there is a growing collaborative community from which many important tools and standards have emerged. By 2003, the RLG/OCLC Working Group had consolidated much of the preservation-specific metadata work done by many internationally recognized digital preservation groups like CEDARS, Pandora, and NEDLIB, among others, to complete a framework for the PREMIS (PREservation Metadata: Implementation Strategies) schema that was made to align within the OAIS Reference Model (OCLC 2002). With work on standardized processes for digital preservation underway by the early 2000s, the first major digital preservation repositories began popping up around the world. Most of these first repositories were built by large national cultural memory institutions such as the National Library of Australia (2000), the Koninklijke Bibliotheek, the National Library of the Netherlands (2002), and the UK National Archives (2003) (Brown 2013, 11). Since the early 2000s many major research projects have advanced the field, such as those funded through the

Library of Congress' National Digital Information Infrastructure and Preservation Program (NDIIPP) (2000). Through this important initiative from the Library of Congress, major tools and services have been developed including JHOVE2, LOCKSS, and the MetaArchive, all which will be discussed later in this thesis (Brown 2013, 11). In December, 2002, Congress accepted the planning report from the NDIIPP and released over $100 million for the program. The recognition of the importance of the NDIIPP's mission to "develop a national strategy to collect, archive, and preserve for current and future generations the burgeoning amounts of digital content..." provided the funding that the NDIIPP needed to become positioned as a leader in the field of digital preservation (Library of Congress 2015a). The U.S. National Archives and Records Administration (NARA) is another leading group within the field today. NARA's work has been heavily focused on the technology and infrastructure necessary to build sustaining digital archives. It has established strong partnerships with some of the leading research institutes and initiatives, including the San Diego Supercomputer Center, U.S. Army Research Laboratories, the National Initiative for Standards and Technology, and the National Aeronautics and Space Administration (Hirtle 2013, 127).

Even more recently, in 2007, a Blue Ribbon Task Force on Sustainable Digital Preservation and Access (BRTF-SDPA) was formed with funding from the National Science Foundation and the Andrew W. Mellon Foundation (Corrado and Moulaison 2014, 73). The BRTF-SDPA has created a powerful partnership with leaders in the field including the Library of Congress, JISC, the Council on Library and Information Resources (CLIR), and NARA (Corrado and Moulaison 2014, 74). This task force seeks to analyze any previous and current models for sustainable digital preservation and identify current best practices among existing collections, repositories, and analogous projects (Blue Ribbon Task Force 2008). Also extremely useful to the field, the BRTF-SDPA has the goal to develop a set of economically viable recommendations to make digital preservation strategies more achievable and reliable (Blue Ribbon Task Force 2008). These goals were achieved within the BRTF-SDPA's final report, *Sustainable Economics for a Digital Planet: Ensuring Long-Term Access to Digital Information.* The report identified three imperatives for any digital preservation stakeholder:

1. *Articulating the value of digital preservation*
2. *Providing clear incentives for preservation in the public interest*
3. *Defining roles and responsibilities among stakeholders to ensure ongoing and efficient flow of resources for digital preservation throughout the digital lifecycle* (BRTF-SDPA 2010, 14).

Still often referenced among the digital preservation community, the BRTF-SDPA report has also detailed five conditions necessary for digital preservation, which has in turn become a tool used as an intellectual backbone for digital preservation programs:

1. *Having decision makers recognize the benefits of digital preservation*
2. *Selecting digital objects that have long-term value*
3. *Having incentives*
4. *Having appropriate organization and governance for digital preservation activities*
5. *Ensuring financial security* (BRTF-SDPA 2010, 73-74).

The notion of long-term *sustainability* (with a focus on economics) addressed in the BRTF-SDPA report is a unique approach to digital preservation compared to those by RLG/OCLC or the 1994 Task Force; much of the research conclusions of the BRTF-SDPA has become an integrated part of most digital preservation planning best practices since the 2010 report was published (Corrado 2014).

As history tells us, the creation of the digital preservation field has been primarily an activity of the library community; as a result, some issues that are important to archivists and museums may have initially received less attention. For example, early library-based digital preservation initiatives focused on capturing and preserving the information found in documents, whereas archivists (and similarly museums) are also interested in preserving the integrity, authenticity, and reliability of original records (Hirtle 2008, 126). Today, while the archiving community has seen a participation spike in digital preservation, the museum-specific field has only engaged mostly within the realm of art collections-focused initiatives such as DOCAM, the New Media Initiatives Group, International Council of Museums CIDCO-DP Working Group, and the Smithsonian Time Based and Digital Art Working Group.

Innumerous digital preservation initiatives exist worldwide today. This is excellent news regarding the promotion and development of digital preservation tools and educational materials. However, museums have generally been 'missing in action' when taking part in such initiatives. This could be due to a lack of knowledge within the museum field, likely pared with lack of leadership and financial resources (Yeung 2004). Currently within the United States, *art* museums are piloting the field of digital preservation because many contemporary digital artworks are directly threatened by obsolescence. Museum ethics has been heavily focused on the treatment of collections (Malaro 1994, 54), which explains why digital preservation efforts thus far has been primarily occurring only in institutions that steward accessioned digital objects. However, it is also important that the field recognize that digital preservation concerns *all* cultural memory institutions, and will in fact become more of a reality in our everyday life as the world's dependence on technology increases. What about history and science museums that produce many academic papers, proceedings, and digitized collections? What about small cultural museums and historical societies that collect oral histories? What about museum digital photographs that document our artistic, political, and cultural world today?

As the emerging topic of digital preservation gains the attention and prevalence it needs in the museum field, more institutions will come to realize that any and all its valuable digital materials (including records, databases, library materials, etc.) should be elevated within their concerns and priorities.

Fundamental Digital Preservation Concepts

The history of digital preservation shows us where the intellectual foundations for digital preservation came from and the many pioneers that contributed to its development. Stemming from the recent past, the *OAIS Reference Model* and the *Trusted Digital Repository Model* are of particular significance when understanding the fundamental process of digital preservation. Since the Open Archival Information System Reference Model is a high-level tool, it is widely accepted by digital preservationists as a key standard for any digital repository (Corrado and Moulaison 2014). This will be the starting point for understanding the more technical processes of a digital preservation system. In the section below, OAIS will be discussed, followed by the

Trusted Digital Repository Model, to highlight how model approaches used today supply the basic infrastructure for digital preservation.

OAIS Reference Model

In essence, the OAIS Reference Model describes how digital objects should be preserved for a certain group of users from the point the objects are acquired to the point when they are disseminated, including ongoing preservation and administrative activities in between (Corrado and Moulaison 2014, 43). Fortunately, because the OAIS Reference Model is meant to be applicable to a variety of collection circumstances, it does not have any specific mandate for the needs of any one specific "designated community,"and so it is designed to be as context neutral as possible (Corrado and Moulaison 2014, 43). While the model deliberately avoids jargon from both IT and the archival professions, it does introduce its own vocabulary to define terms related to digital preservation within its own context (Corrado and Moulaison 2014, 43). In addition to a unique digital preservation vocabulary, the OAIS also provides a *data model* (or some refer to it as an information model), a recommended *functional model* to actually "[carry] out the archive's required responsibilities,"(Ockerbloom 2008) and a detailed set of those responsibilities. The JISC Standards Catalogue notes that the OAIS "documentation is quite long and complex and this may prove to be a barrier to smaller repositories or archives"(Allinson 2006). It is true that the actual 2012 Magenta Book ISO standard is about 148 pages, but the basics of the OAIS model can be understood without having to be conversant with the entire reference document. What the model supplies for museums is a better understanding of what one needs to be doing in a theoretical way if one plans on maintaining digital media/information for the long term.

The vocabulary created for the OAIS Reference Model assumes that digital preservationists will need their own language in order for different stakeholders to communicate effectively between themselves and with IT (Corrado and Moulaison 2014, 43). While there are many definitions to be found in the first section of the OAIS Magenta Book, only some of the most commonly referred to terms need to be well understood in order to understand the model, and therefore, how most digital preservation systems work. "Designated community" has

already been mentioned earlier in this chapter, but it is an important term for the OAIS as it addresses the wide range of people involved in the preservation model. Designated community is defined as: "an identified group of potential Consumers who should be able to understand a particular set of information. The Designated Community may be composed of multiple user communities. A Designated Community is defined by the Archive and this definition may change over time"(Magenta Book 2012, 1-11). For the museum field, the OAIS definition of Designated Community is equivalent to what the field simply calls "users," and it also seems to include what we call "stakeholders," in addition to "users."

Central to the OAIS Reference Model is the concept of "packaging" information; in the computer science world this is nearly synonymous with using what's called a "wrapper."According to the 2012 Magenta Book, an Information Package is: "a logical container composed of optional Content Information and optional associated Preservation Description Information (PDI). Associated with this Information Package is Packaging Information used to delimit and identify the Content Information and Package Description information used to facilitate searches for the Content Information" (Magenta Book 2012. 1-12). This definition is rather confusing, but essentially the Information Package is the central entity within an OAIS archive. The noted Content Information is the actual data object which the archive is trying to preserve, plus any accompanying Representation Information, which is the equivalent of the information that "maps a Data Object into more meaningful concepts" (Magenta Book 2012. 1-14). An example of Representation Information for a bit sequence is JPEG software that is needed to render a JPEG file; rendering the JPEG file as bits is not meaningful to humans, but the software, which embodies an understanding of the JPEG standard, can map the bits into pixels which can be rendered as an image for human viewing. Also within the Information Package is Preservation Description Information or PDI, i.e. all the information needed to preserve the digital object together with any Representation Information, which will be needed in order for the object to be understood. The PDI is likely to include provenance, context, reference codes (like unique identifiers such as accession numbers), and fixity (Alan 2008). In addition, one's OAIS software system of choice can include Packaging Information at its own discretion. This Packaging Information includes any information like file structure or

directory structure that the system recognizes the data to have (Alan 2008). There is often a separately-stored metadata file that houses the Descriptive Information, which enables the whole Information Package to be searchable within the OAIS Archive after it has been ingested. It might just be the title of the package, or a full set of searchable attributes. These elements are what make up the central product of the OAIS. Now that the general concept of the Information Package is understood, there are an additional three types of IP that are rendered throughout the OAIS Reference Model: Submission Information Package (SIP), Archival Information Package (AIP), and Dissemination Information Package (DIP)(Magenta Book 2012, 2-7).

Submission Information Packages are the first step in any OAIS model, and perhaps one of the more important terms to deeply understand. SIPs are information packages that are delivered by a Producer to the OAIS, of which the data within the SIP can be used in the construction or update of one or more Archival Information Packages that may or may not already be in the Archive (Magenta Book 2012. 1-15). So some of the submissions of original data objects, which the OAIS is calling an SIP, will have insufficient Representation Information or PDI to meet the stringent AIP requirements, which is why they are differentiated (Alans 2008b). Most SIPs will have some Content Information and some PDI, but may require many submissions to form a final AIP. Ideally there should be a submission agreement between the Producer and the OAIS that specifies criteria like file formats, subject matter, ingest schedule, access restrictions, verification protocols, etc. (Magenta Book 2012, 2-9). The data submission formats, procedures, and deliverables must be documented in the OAIS's data submission policies in order to streamline the process (Magenta Book 2012, 4-12). The Ingest entity in an OAIS software accepts SIPs, performs some quality assurance checks, and then generates an AIP. If there are errors in the SIP submission, then Ingest will request a resubmission. So, upon adding an SIP, Ingest then transforms the SIPs into AIPs, which can include file format conversion, reorganization, transfer to different media, or create a unique identifier (Magenta Book 2012, 2-8).

The next step in the process, the creation of the Archival Information Package, consists of the "Content Information and the associated PDI, which is preserved within an OAIS"(Magenta Book 2012, 1-9). One may now ask, but how is the Content Information and

PDI of this Information (data) Object preserved? It is important to remember that this model is still extremely high-level, so the specific technical processes are something IT or a developer would more greatly understand. Essentially the AIP "[provides] a concise way of referring to a set of information that has, in principle, all the qualities needed for permanent, or indefinite, Long Term Preservation of a designated Information Object...the specification of the AIP as a container that contains all the needed information to allow Long Term Preservation and access to archive holdings remains valid"(Magenta Book 2012, 4-36). Tangible examples of the kinds of information that an AIP would contain may include system architecture, necessary software or APIs, vendor information, hardware specifications, bitstream orientation (big-endian vs. little-endian), other compatible file formats, etc. The AIP packages the source code of an Information Object, along with a very specific platform-independent set of instructions for how to use and view that information in a human-readable way. Without a system that can create such a roadmap, there would be no way for people of the future to know how to run a certain set of source code to make it usable, nor would they be able to search for data objects that are in threat of becoming obsolete due to imminent changing hardware and software to save them. To make this process even more tangible, Figure 2 supplies a visual of the OAIS Reference Model:

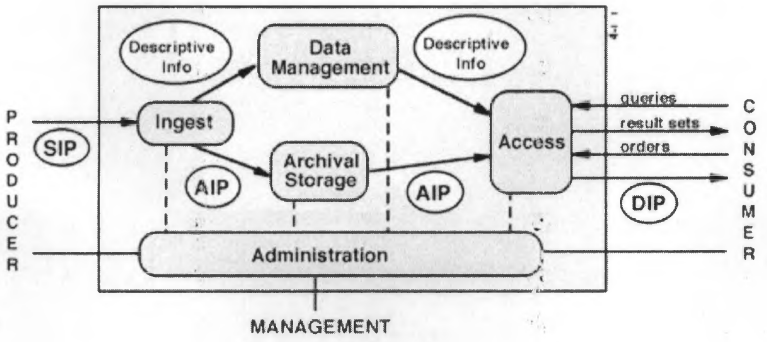


Figure 2: OAIS Reference Model (Sawyer 2000)

Keep in mind the various entities in the model are not always solely controlled by preservation software, but can involve a combination of staff and computer tools. For example, the AIPs are

managed within the OAIS by an Archival storage entity (Magenta Book 2012, 4-1). The Archival Storage entity functions can include managing the storage, refreshing the media, performing routine and special error checking, and providing disaster recovery capabilities (Magenta Book 2012, 4-2). These necessary functions can be performed by staff, or by way of preservation software tools, or a combination of both. The OAIS simply requires that these functions are performed within that step in the preservation process. These functions outlined in the model, along with preserving a specific set of instructions for that data object, are how digital materials conceptually are preserved.

The Dissemination Information Package is the most simple element of the system and is defined by the OAIS as "an Information Package, derived from one or more AIPs and sent by the Archives to the Consumer in response to a request to the OAIS" (Magenta Book 2012, 1-11). The DIP is the "use-copy" and is only created should a data object from the system be needed for use by someone in the Designated Community. If the DIP is somewhat of an equivalent to a "use-copy," then the AIP can be thought of as a similar concept to maintaining a "master-copy." The OAIS in many ways aligns with normal archival concepts, but it is overall much more technical since its goal is to protect materials that are made of the immaterial source code.

The process outlined by SIPs, AIPs, and DIPs makes up the OAIS Functional Model. However as highlighted by Figure 2, other functions also occur in the Functional Model, such as Data Management and Administration. Part of any OAIS compliant system, will include a Data Management entity that provides services and functions related to populating, maintaining, and accessing descriptive and administrative metadata. These can include maintaining schemas, performing database updates, performing queries, and producing reports (Corrado and Moulaison 2014, 46). The Access entity makes the Archive's holdings visible to Consumers, so it provides functions to support end users. It allows users to search the Archive's contents, create DIPs, and monitors their delivery (Alan 2008c).

Underlying all of these entities is the Administration function, which essentially monitors the OAIS's operation, looks for ways to optimize the system, and negotiates submission agreements with Producers. Additionally it provides systems-engineering functions and is

"responsible for establishing and maintaining Archive standards and policies, providing customer support, and activating stored requests"(Magenta Book 2012, 4-2).

The very last component of the OAIS Reference Model is the responsibility of the Preservation Planning entity. The people involved in the Preservation Planning step provide recommendations and preservation plans to ensure that the information stored in the OAIS remains accessible to, and understandable by the Designated Community over the long term. These recommendations can include archival information updates, migration of holdings, and documenting Archive standards and policies (Alan 2008c; Corrado and Moulaison 2014).

While the OAIS Reference Model is a significant element of digital preservation practices today, it is important to note that to be considered compliant with OAIS, one need only fulfill the required responsibilities (Administration, Preservation Planning, Data Management), and support the basic OAIS data model of information packages (Ockerbloom 2008). Nonetheless, it can be very useful to fully understand the functions, both to make sure that an institution is doing everything it needs to do, and to see how the big problem of digital preservation can be broken down into smaller, more manageable workflows.

Trusted Digital Repository Model

While the OAIS is arguably highly technical, it has also been used as a model to help construct detailed criteria for "trusted repositories," as well as to audit and to create certification checklists (Ockerbloom 2008). *Trust* is a crucial fundamental for digital preservation (Corrado and Moulaison 2014, 95). Digital repositories provide services to both those who deposit content for preservation, and to those who consume that content sometime in the near or far future (Brown 2013). As such, establishing a high level of *trust* revolving around institutional responsibility, and trust in the authenticity of its objects, is the fundamental backbone for any cultural institution's authority as stewards of objects, in this case including digital materials. The trust concept here resonates with museum practices and approaches, including those outlined by museum scholars such as Marie Malaro, who has published extensively on museum code of ethics, stewardship, and collections (Malaro 1994). Just as Malaro has asserted for years in the museum field that the effectiveness of an institution

depends on the extent to which it is trusted by its community, the same concept is equally transferable to preservation and management of digital collections and records.

How can a museum operating a digital preservation system or repository establish this *trust*? One way to answer this question is for repositories to establish procedures based on current best practices that have been recognized by the digital preservation experts, and then to document that they are following said practices rigorously (Corrado and Moulaison 2014, 95). In fact, in 2002 the RLG/OCLC published a report, *Trusted Digital Repositories: Attributes and Responsibilities*, that established the actual attributes of *trust* for organizations (mainly research institutes), all the while incorporating the ISO standard of the OAIS Reference Model. The report functions to define the specific characteristics and responsibilities of trustworthy digital repositories within cultural organizations (RLG-OCLC 2002). The report even includes museums within its audience by using a museum as an example of a cultural institution that can follow due diligence to become a Trusted Digital Repository:

> "*Scenario 3: A museum with a growing collection of digital materials, including surrogates of museum objects, surrogates created for online exhibitions, and original digital art. The museum serves a very diverse community comprising students, researchers, artists, the general public, and organizations seeking digital material for commercial use...The museum uses a content management system to provide day-to-day access to the digital collections, but the system was never intended to facilitate archival storage. Because the museum lacks technical infrastructure and qualified staff, it will contract with a third-party archiving service so that its materials will be professionally managed, controlled, and backed up to meet its long-term management responsibilities. The commercial service is OAIS-compliant...*"
> (RLG-OCLC 2002, 6-7)

As evidenced in the RLG/OCLC 2002 Report, taking action to acquire attributes of a trusted digital repository can be as minor as using an OAIS-compliant third-party service. However, museums should also follow the other recommendations in the RLG/OCLC report and consider the process of formal certification, or that of self-auditing to meet standards and criteria of a Trusted Digital Repository (TDR). Referring back to the 1996 Task Force on Archiving of Digital Information, it was also recognized in this primary document that some kind of formal process of certification would aid in creating an overall climate of value and trust about the viability of preserving digital materials (Task Force 1996).

Since the publishing of both of the 1996 Task Force report and the 2002 RLG/OCLC report, a number of substantive initiatives have emerged, including:

- **TRAC:** The Trustworthy Repositories Audit & Certification Criteria and Checklist (TRAC) published in 2007, has become an international *de facto standard.* TRAC is considered a *formal* certification process, but has self-audit options as well (RLG-NARA 2007).
- **ISO 16363:** Primarily based on TRAC, this ISO standard published in 2012 formalized previous TDR initiatives as "*de jure* standards" (Brown 2013, 85).
- **Nestor**: The Network of Expertise in long-term STORage (Nestor) is the German trusted digital repository certification initiative. It is coordinated with the TRAC standard, but also focused particularly on the requirements of libraries, archives, and museums in Germany. Nestor is characterized as an *extended* certification level of TDR (Corrado and Moulaison 2014). It is now published as an official DIN standard (DIN 31644) (Brown 2013, 85).
- **DRAMBORA**: The Digital Repository Audit Method Based on Risk Assessment (DRAMBORA) was also published in 2007 as a toolkit that provides a risk-based methodology for TDR audits. It was developed by the UK Digital Curation Center and the DigitalPreservationEurope project. This toolkit draws on and complements the TRAC and Nestor, but it additionally focuses on practical application of audit methods, based on pure self-assessment (McHugh et al. 2008).
- **Data Seal of Approval**: The 2010 Data Seal of Approval (DSA) provides a much lighter weight assessment process than the other initiatives and is therefore considered more of a *basic* certification. It is seen as a distilled version of both DRAMBORA and TRAC checklists. This TDR certification uses 16 criteria that a repository assesses itself to determine if it fits within four simplified compliance levels, and this is subject to external review by the DSA Board in order to receive rights to use the DSA seal (Corrado and Moulaison 2014, 99).

While these various levels of basic-formal level certifications are available, it is important to note that in many cases repositories will operate perfectly well without satisfying every single criterion or checklist item; therefore, it is common for institutions to forego the formality of certification and opt for using these standards for self-auditing, as well as for evaluating its needs in lieu of selecting the best third-party vendors or products (Ockerbloom 2008).

These methods of certification and audit that help establish prudence and verification as a trusted digital repository, but also are most accessible once an institution has *already* established some kind of digital preservation system. One can argue that beginners can use the Trusted Digital Repository Model to help determine the requirements to be considered a TDR eventually, so it is certainly true that these models may seem out of reach if an institution is just starting their journey. Outside of the TDR certifications, the Library of Congress' National Digital Stewardship Alliance has created a model for digital preservation called, "Levels of Digital Preservation" that any institution can use regardless of maturity of its digital preservation program. The "Levels of Digital Preservation" is a tiered set of guidelines for how organizations should begin to build or enhance their digital preservation activities (NDSA 2015). While always a work in progress, it is intended to be an easy-to-use grid that can walk any institution through planning and enhancing their concepts for existing systems or workflows (NDSA, 2015). These guidelines are organized into five functional areas that are thought to be at the heart of digital preservation systems: storage and geographic location, file fixity and data integrity, information security, metadata, and file formats (NDSA 2015). See Appendix B for a closer look at the NDSA "Levels of Digital Preservation."

Mapping the Trusted Digital Repository Model to OAIS

While the concept of becoming a Trusted Digital Repository is focused on the administrative aspects of building and managing a digital repository, the OAIS takes the role of outlining the specific functions and processes. Many in the digital preservation community rely on these two foundational documents, and they can even more relatable to one another when mapped together. Nancy McGovern and Anne Kenney from the University of Michigan Library

have mapped the OAIS model to the TDR framework as shown in Figure 3 (Kenney and McGovern 2011).

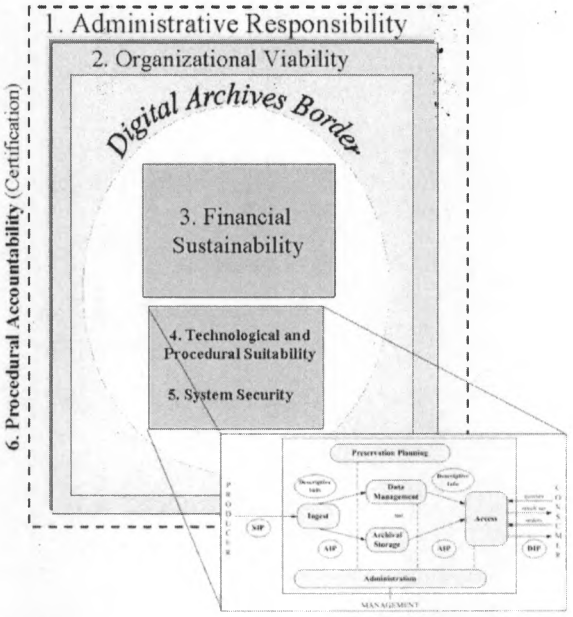


Figure 3, Mapping the OAIS to the TDR (Kenney and McGovern 2011)

The OAIS and the TDR together are a more meaningful and comprehensive model for digital preservation planning and development. Whereas, the components separated, or alone, inflicts a disconnection between the technical requirements of the OAIS from the conceptual model of trustworthy archiving, or collection management. In reality, they are very much related to each other, but as suggested by McGovern and Kenney, only overlap in certain areas. Figure 3 supplies a visual for how the layers of requirements posited by the Trusted Digital Repository model relate to each other; for example, the outer layers of Administrative Responsibility and Organizational Viability hold together the entities that make up a digital archive like financial sustainability and the technological procedures, which is where the OAIS model fits nicely within the overall model. Kenney and McGovern's mapping of the two foundational documents in digital preservation also emphasizes the importance of institutionalization of these processes (Kenney and McGovern 2011). Institutionalization thus requires an explicit acceptance by the organization of responsibility for and commitment to a digital preservation program. An

encapsulated view of the infrastructure for successful digital preservation, as shown here, incorporates institution-wide, ongoing planning to establish a program, including development of fundamental policies and guidelines, as well as the allocation of core funding over time. These concepts will return during the discussion of digital preservation policy in Chapter 5.

Common Steps in Digital Preservation Systems

From a general perspective provided by the OAIS Reference Model, or the requirements of a Trusted Digital Repository, the overall functions of any repository can therefore be specified as: capture of content, preservation management, and access (Brown 2013). With this in mind, one can certainly model a system's functional requirements after the elements of the OAIS: ingest, data management, archival storage, preservation planning, administration, and access. But since these functions may not be entirely self-explanatory, especially beyond the digital preservation practitioner community, it is practical to break down the functional aspects of digital preservation to make the high-level concepts understandable and in the form of a practical workflow. While workflows certainly vary from institution to institution, the common steps of a digital preservation system to be discussed below are: *appraisal/selection, identification, transfer, ingest, quarantine, characterization, integrity and authenticity checks, and metadata encapsulation.*

Appraisal and Selection

One of the first practical aspects of any digital preservation program stems from archival practices: *appraisal.* According to the Society of American Archivists, *appraisal* is simply the process of identifying what materials have permanent (archival) value (SAA 2015). While this concept has been widely applied throughout the archival and even museum field when regarding accessioning of objects/collections, the same process is necessary when determining what digital objects need to be added to an archival system (SAA 2015). For a digital preservation program or system, the appraisal step occurs before the technical function of "ingest" (Brown 2013, 110). In this first stage of any preservation workflow, the institution makes a decision to acquire a specific collection of material, performs any preparatory activities, and then physically transfers that material into its custody (Brown 2013, 109). The

analogous museum field practice for this process is "accessioning." The selection and transfer process can be initiated in a many different ways, but the most common are planned deposits (single occurrence) and periodic transfer (regular, recurring) (Brown 2013, 111). For example, many cultural memory institutions have programs to digitize all or parts of their analog collections. The digital surrogates created represent a significant investment of time, money, and expertise, and may be difficult or impossible to recreate. To preserve these assets, one could consider periodic transfer (in the case of ongoing digitization), or planned deposit (after the program is already complete), in which the selection decision will be made only once.

Each institution will have its own well-defined approach for the appraisal and selection of collections and content, which also applies in the digital world. Museum curators and collection managers have always performed these duties for analog collections. Thus the same fundamental rationale for museum collection management is shared equally in the digital world: the institution should have some form of documented collection policy that defines the types of materials it chooses to collect. These written policies are key in providing staff guidelines and rules when assessing potential acquisitions to determine whether or not they should be selected. Such guidelines may occur under various guises including collection management policies, acquisition and deaccession policies, or information management policies, but they are something that every museum, library, or archive should have (Brown 2013, 113). The very beginning of any digital preservation process will be outlined and attainable through such documentation. For example, every institution should establish standards that govern the acceptability of digital objects into a repository like file formats, minimum documentation standards, transfer media, etc. (Brown 2013, 118). More details and discussion on digital collection and preservation policies can be found in Chapter 5.

Identification

The process of assigning *identification* in the form of metadata or descriptors is another fundamental technical function of digital preservation within the appraisal step. It is what digital preservationists and other stakeholders need in order to organize and retrieve digital content so that it does not become lost or unidentifiable (Corrado and Moulaison 2014, 111) The initial identification of a digital object can be as simple as creating a standardized file name, but

probably will involve additional metadata capture if the information is readily available (Corrado 2014, 111). Identifying a digital object with as much descriptive, technical, and administrative metadata up front is always recommended. After identification has been initiated, some form of evaluation of the digital content has been assessed against a collecting policy, and with some additional considerations of legal and technical issues, a formal agreement between the depositor and repository will culminate, and transfer can finally be made (Brown 2013, 110). The outcome of the selection and appraisal process, as agreed between the depositor and the repository, should be clearly documented (Brown 2013, 124). Transfer agreements often include descriptive metadata, agreed documentation to be provided by donor, agreed methods and timeframes for transfer, level of preservation required, and any conditions on use and access by the repository and its end-users (Brown 2013, 125).

Transfer

Once a transfer agreement has been made, the repository producer may now prepare the digital object for *transfer*. These preparations may include assembling, creating, or transforming documentation into acceptable formats, migrating content to approved transfer formats, or setting up/configuring transfer mechanisms, such as FTP sites or File Transfer Protocol sites (Brown 2013, 125). After preparation for transfer is completed, the actual physical transfer of the digital content and associated documentation to the repository can occur (Brown 2013, 111). At this stage, the digital object is simply moved to a storage environment controlled by the receiving institution; accession and ingest will come later. After having completed the transfer, the repository system should acknowledge receipt to the depositor in some form of a report or documentation (Brown 2013, 111). From this point, is where the *ingest* process begins, and the production of a SIP and then a AIP to be submitted into archival storage.

Ingest

Ingest of a Submission Information Package, then generation of an Archival Information Package, will likely account for the majority of activity when preparing digital objects for long-term preservation, so it is especially important to have an efficient and robust process (Brown 2013, 129). It is especially helpful to standardize as many of the accession activities as possible, although this may or may not be feasible depending on the nature of the materials being

ingested. The first stage of the ingest process follows closely to the workflow recommended by the OAIS Reference Model; the content and metadata will be assembled into a "package" suitable for long-term preservation. Although the OAIS defines the idea of a SIP, there is no actual standard format for these information packages, and so practice varies among digital repositories. The SIP is a conceptual entity, but some real-life embodiments can be in the form of a physical package, such as a Zip file, but this truly varies from institution to institution (Brown 2013, 131). Many institutions involved in digital preservation already use a container format as an efficient means for managing SIPs (Brown 2013, 132). There are a series of tools developed in recent years for this very purpose, such as software that works to package content according to the BagIt format, which is growing in popularity for creating SIPs (Brown 2013, 132). These tools include Bagger, BagIt Library, and BagIt Transfer Utilities to name a few. On the other hand, many commercial preservation systems will use their own SIP formats, and provide proprietary SIP creation tool which means this step can be automated (Brown 2013, 132). It is important to remember that a SIP does not need to be complicated if the object is not complicated. A SIP format can be as simple as a standard folder structure using typical naming conventions. For example a SIP can be a folder that contains two subfolders: one for the original data/content, and the other for the metadata. For those starting to create or understand a digital preservation workflow, the key point here is that the actual format of an SIP is less important than the fact of having one (Brown 2013, 132).

Quarantine

After the SIP has been created and before ingest, another important practical aspect to consider for a preservation workflow is the concept of *quarantine*. The purpose of quarantine is simply to prevent any malicious software from being ingested into the digital repository (Harvey and Mahard 2014, 312). Similarly, in museums, all new accessions are inspected for signs of mold or insect infestation, which could spread to other parts of the collection once stored. It is of course important to identify such problems up front, and treat them in a quarantine until neutralized for deposit into long-term storage. In the digital environment equivalent threats are posed by viruses, worms, Trojans, and other forms of malware (Brown 2013, 134). These too can spread and infect other systems inside an institution's computer network, thus malware

detection, virus-checks, and any other treatment must take place before the content is allowed to make direct contact with the repository itself (Brown 2013, 134). To quarantine a newly transferred digital object, one needs a quarantine environment, which for traditional museum objects can take the form of a freezer or anoxic chamber; for the digital side of collections this can simply take the form of a dedicated laptop or workstation that is physically isolated from other systems (Harvey and Mahard 2014, 312). It is conventional to leave new SIPs in quarantine for four to six weeks, with a final virus-check rerun just before ingest (Harvey 2014, 312).

Characterization

As with any library, archive or museum collection, it is imperative that all aspects of its holdings be well understood to allow the objects to be properly managed, preserved, and accessed in the future. Ideally in practice this requires the capture of metadata describing the content in enough detail. However for digital objects, in the real world it is unrealistic to expect the depositor to be able to provide all this information. The types of metadata required for long-term preservation by a digital repository is often highly technical or specialized and most administrators of a digital repository will lack the knowledge or wherewithal to provide it. The ingest process of *characterization* can be used to fill this gap (Brown 2013, 136). *Characterization* is a series of processes used to identify, extract, and record the nature of digital objects. A simple characterization function may be determining the format of a file, and a more advanced function might include the automated extraction and construction of metadata for the object (Hutchins 2012, 8). Characterization usually includes these three main activities: identification of the format, validation of the object to confirm that it is correctly formed, and metadata extraction to acquire additional descriptive or technical information (Brown 2013, 136).

Characterization relies on automated tools, many of which are either free or at least very inexpensive. A well-known format identification tool that was made specifically for digital preservation is DROID, a free tool developed by The National Archives; it performs automated batch characterization of file formats (DROID 2015). Another well known open-source tool is FIDO (Format Identification for Digital Objects) which was developed by the Open Planets Foundation (Hatchins 2012, 16). Both DROID and FIDO cull from the PRONOM database of

file formats, which in of itself is an important tool for digital preservation (Hutchins 2012, 16). More specialized metadata extraction tools are JHOVE and JHOVE2, which are software tools developed in conjunction with JSTOR and Harvard University to create an open-source platform to perform the three processes of characterization (DCC 2015c; JHOVE2 2015).

Authenticity and Integrity Checks

Similar to checking for an artist's signature or documenting the condition of a museum object, checking the *integrity and authenticity* of a digital object is a necessary function before, during, and after ingest into a digital repository (NDSA 2014). Immutability of a digital object can prove that the content within an archive, or the content being accessed from an archive is in fact the expected quality for a trustworthy document or object (Interpares 2015). The terms *integrity*, *authenticity*, and *fixity* are often interchanged seamlessly; while these terms all have similar meanings and goals, the term *fixity* is most accurately applied to the technical functions of a digital preservation workflow. Fixity is the property of a digital file or object being fixed or unchanged and can be viewed as synonymous with bit-level integrity (NDSA 2014). The PREMIS data dictionary defines fixity information as "information used to verify whether an object has been altered in an undocumented or unauthorized way" (PREMIS 2008, 46). So fixity checks generate data used to verify whether an object has been altered or degraded at the point of access. There are many opportunities for change and fixity threats within a digital object's lifecycle including: submission, retrieval, migration, transfer to media, network transmission, or simply the passage of time (Novak 2006, 1).

The most widely used tools for establishing fixity are checksums (such as CRCs [cyclic redundancy check]) and cryptographic hashes (such as MD5 and SHA algorithms), but there are other methods such as expected file size and file count that can provide basic fixity confirmation (NDSA 2014, 1; Duryee, 2014). The process of using checksums, hashes, etc for fixity checks throughout the lifecycle of a digital object is critical for what is termed "bit-level preservation" (as opposed to the full scope of digital preservation which also includes format preservation). While only one aspect of digital preservation, fixity is a significant aspect for maintaining authentic digital objects in a trusted repository (NDSA 2014).

Metadata Encapsulation

As described in the OAIS Reference Model, inclusion of various types of *metadata* in the Submission Information Package to Archival Information Package creation process is extremely key for long-term preservation of digital objects. Making the case for documenting any and all functional requirements and human identifiers in the form of metadata essentially acts as a set of instructions, or a road map, to keep a digital object "alive" when our future successors need to access the material. In traditional collections, a museum conservator can run tests on a painting to determine the age and chemical makeup of the paint used to create it. Of course, for something made with source code, to unlock its meaning or renderability, only the catalogued metadata can describe what system architecture, hardware, etc., is required to run it. In addition, metadata enables a resource "to be understood by both humans and machines in ways that promote interoperability"(NISO 2004, 2). This is key for long-term preservation since interoperability enables "multiple systems with different hardware and software platforms, data structures, and interfaces to exchange data with minimal loss of content and functionality"(NISO 2004, 2).. Thus the topic of metadata is extremely important and enormous in the field of digital preservation; and this thesis will will only be able to briefly discuss this topic. Listings of more metadata standards for digital preservation are available on the Web. Preservation authorities, Digital Curation Center and the Library of Congress both provide lists of metadata standards that are important to the field (Library of Congress 2015; DCC 2015d).

It has been implied that there are various categories of metadata that can be recorded. Generally speaking, there are four fundamental kinds of metadata: (1) descriptive metadata, (2) administrative metadata, (3) technical metadata, (4) structural metadata (Corrado and Moulaison 2014, 113). While descriptive metadata is the most intuitive (involving attributes like title, author/creator, name, date, etc), technical and/or structural metadata is much more specialized and can require a greater understanding of computer science (involving attributes like file format, file size, software, hardware, relationship to other files, etc) (Corrado and Moulaison 2014, 113). In order to know what metadata attributes of a digital object one should be recording, standards known as metadata schemas have been created in the field; one can pick a particular schema that best fits their needs, and also adjust those standards to more customized

options as necessary. According to NISO (National Information Standards Organization), metadata schema "are sets of elements designed for a specific purpose, such as describing a particular type of information resource"(NISO 2004, 2). Some of the most common metadata schemas used among cultural memory institutions are Dublin Core, VRA Core (Visual Resources Associate Core), PBCore (Public Broadcasting Core), CDWA (Categories for the Description of Works of Art), among others (Corrado and Moulaison 2014, 120). The OAIS Reference Model recommends that metadata either be stored as a separate but connected entity, or embedded with the data object. Metadata files are often, but not exclusively, formatted with a header and body using Extensible Markup Language (XML) as an encoding language (Corrado and Moulaison 2014, 117). XML is derived from, Standard Generalized Markup Language (SGML)(ISO 8879) and is maintained by the World Wide Web Consortium (W3C) (W3C 2015). XML has become such a widely used standard and has the ability to exchange data with the Web and store data in digital preservation systems, which makes it a very appealing choice for metadata formats (Corrado and Moulaison 2014, 118). An example of a non-XML markup language that is also often used in the digital preservation field is MARC (MAChine-Readable Cataloging)(ISO 2709) (Corrado and Moulaison 2014, 118). For institutions needing a less robust format for metadata entry, using the Unicode compatible, human-machine readable ASCII encoding language is also considered very durable for long-term preservation (NINCH 2003, 84, 88, 201).

Even more importantly for the purposes of this research are the metadata schemas that have been created specifically for digital preservation purposes. The reason that digital preservation metadata is considered different from standard approaches is because of the specialized nature of the digital preservation process (Corrado and Moulaison 2014, 125). The OAIS Reference Model confirms that long-term digital preservation requires that the digital object and its metadata be maintained in tandem (Corrado and Moulaison 2014, 126). While in the end, metadata that supports digital preservation is just metadata, it does need to include three broad categories:

1. Descriptive metadata necessary for retrieval and storage
2. The Preservation Description Information (provenance, reference, fixity, context, access rights as described by the OAIS Reference Model)
3. Metadata (usually technical schemas) about digital objects and their digital environments (Corrado 2014, 131-32).

While the first two parts of any digital preservation metadata package can be related to normal descriptive metadata standards (such as Dublin Core or VRA Core), the third category of technical metadata needs to be more specific to preservation attributes. The primary standard used for this purpose in digital preservation systems is PREMIS (Preservation Metadata: Implementation Strategies). Since its publication by the OCLC/RLG Working Group in 2003, PREMIS has gained widespread adoption and is considered the *de facto* international standard for preservation metadata (Brown 2013, 166-67). In particular the PREMIS Data Dictionary is an excellent resource for defining the semantic units which are associated with the OAIS preservation data model (Brown 2013, 168).

In order to combine all the required metadata together in a way that is easier to store with the original data object in a digital archive, the file packaging format BagIt can be used, in which a 'bag' container enables easier transfer across a network or to physical media; this method is used by the Library of Congress and the California Digital Library (Brown 2013, 171). As iterated earlier, BagIt is increasingly used as an ingest package format (SIP), a means of transferring content (DIP), and also as a convenient format for storage (AIP). The other metadata container formats commonly used in tandem with PREMIS is MPEG-21 and METS (Metadata Encoding and Transmission Standard) (Harvey and Mahard 2014, 119; Brown 2013, 171). While MPEG-21 is a complex, sophisticated container format most suited for audiovisual resources (Brown 2013, 171), METS is a much more common standard developed and endorsed by the Library of Congress (Library of Congress 2015). METS is expressed as an XML document that encodes embedded descriptive and administrative metadata (METS 2015). Because METS schema keeps together a hierarchical order for different parts of a data object, the provided structure and flexibility has made it the most widely adopted generic metadata scheme. One may often hear about the 'METS wrapper' because of the enveloped metadata

structure that METS schema provides. It is therefore a common best practice to use a metadata scheme like PREMIS, VRA Core, or MODS, in tandem with METS which acts as the structural metadata wrapper that holds it all together (Habing 2007).

Conclusion

Digital preservation workflows are an extremely variable topic and thus the practical elements discussed above must be considered from a general point of view. However, the practical functions of appraisal/selection, identification, transfer, ingest, quarantine, fixity, characterization, and metadata encapsulation are best practices when formulating any institutions unique workflow. To fulfill the functions of these workflow steps, many various preservation techniques and tools must be employed along the way. For more information on the practical implementation of a digital archiving and preservation workflow, Ricky Erway and Julianna Barrera-Gomez for the OCLC published a paper titled "*Walk this Way: Detailed Steps for Transferring Born-Digital Content from Media You Can Read In-House*" (Erway 2013). This paper offers eleven steps for how to begin the preservation process, resources, and sample workflow charts (Erway 2013).

Digital Preservation Methods

Whereas the practical steps described above outline how to prepare content for submission into a digital repository, there are additional methods, or strategies, that can be used to keep a digital object and its source code viable for future use, exhibition, or access. Such digital preservation strategies can sometimes be used in lieu of having an established digital preservation repository. For museums, some of the methods to be outlined below are important considerations for access and exhibition of digital collections. The methods and tools of digital preservation discussed here do not encapsulate a full view of the resources available, but is a basic overview of the some commonly discussed methods in the field today. The practical methods for long-term digital preservation always seek to maintain the authenticity of an object which can be validated by three essential objectives: reliability, integrity, and usability (MIT Libraries 2012c). The methods of *refreshing*, *migration*, *emulation*, and *replication* are commonly employed in digital preservation procedures (CHIN 2013). It is important to note that

many of the methods of digital preservation are not just concerned with preservation workflow (OAIS) and storage, but also with long-term usability, or renderability.

Refreshing

According to the Digital Preservation Workshop developed by Cornell University Library and the National Endowment of the Humanities, *refreshing* is a preservation method that transfers data between two versions of the same storage medium to prevent any "change whatsoever in the bitstream (e.g. from a decaying 4mm DAT tape to a new 4mm DAT tape, or from an older CD-RW to a new CD-RW)" (MIT Libraries 2012d). In addition, there is an alternative method called *modified refreshing*, in which one copies that data object from a medium to a "similar enough type" that still maintains no change in the bit-pattern that would affect the application or operating system used to run the data (MIT Libraries 2012d). For example, one can copy the bitstream from a 100 MB Zip disk to a 750 MB Zip disk. While perhaps not an end-all solution for digital preservation, refreshing is certainly a necessary component of any successful digital preservation program (MIT Libraries 2012d). Refreshing helps to address both decay and obsolescence issues related to media for more immediate or short-term contexts.

Migration

The refreshing strategy is often combined with what is called *migration*, especially when the software or hardware required to read the data is no longer available or unable to understand the current format of the data (MIT Libraries 2012d). Migration is the process of copying or converting data from one technology to another, whether that is hardware or software, preserving the essential characteristics of that data (MIT Libraries 2012d). Instead of transferring the data to the same or similar media, migration is concerned with migrating the data to current standards, to avoid maintaining data on hardware or software that is or near obsolete (Brown 2013, 209). For example, one can convert documents created with an obsolete word processor (such as WordPerfect, .wpd) into a format that is accessible using contemporary software (Microsoft Word, .docx). Migration requires the use of software tools capable of converting data objects from one file format to another. While this may seem like a simple

solution for applicable types of digital collections, there are a number of potential challenges (Brown 2013, 209). With any transformation process there is a risk of potential information loss, which according to digital preservationist Adrian Brown can arise from two sources:

- *" The target format may not support the full range of significant properties required to preserve the performance of the original. For example, converting a Word document to plain text will lose much of the formatting.*
- *The migration process may not be capable of transforming all the properties of the original. As an example, the MS Word 97 filter for converting Wordstar documents did not correctly interpret how WordStar used the 8th bit of each ASCII character, resulting in the insertion of incorrect typographical characters..."* (Brown 2013, 210).

To avoid such problems, it is vital to choose the appropriate formats to migrate to, and suitable migration tools. The best practice in migration combines source format, migration process, and target format, which together is commonly known as the migration pathway (Brown 2013, 210). Some migration pathways may require multiple migration processes through various intermediate formats in order to achieve the ultimate pathway between a source format and target format that are not directly compatible (Brown, p. 210). While migration is a broader and richer concept than refreshing, as directed through the concept of digital curation, migration still needs to be employed periodically in order to preserve the integrity of digital objects and retain its renderability in the face of constantly changing technology. Migration certainly has its benefits; however, the process of migrating complex file formats has not been widely tested. Many criticize migration on the basis that neither authenticity nor confident integrity can be assured (MIT Libraries 2012d).

Emulation

When migration is not appropriate, digital preservationists often turn to the method of *emulation*. Migration is considered unsuitable to the preservation of software or for any kind of digital object that has complex display behaviors (Brown 2013, 212). This method utilizes the concept of maintaining the object in its original form, and instead developing ways to access it

within the current technology environment. This can be done by way of an emulator software that mimics the original (often obsolete) digital environment but runs on current hardware, applications, and software (Fino-Radin 2013, 110-12). Of course for many museum digital collections, especially concerning original works of art, the most ideal preferred environment would be using the authentic period-specific (vintage) hardware/software (Fine-Radin 2013, 110-12). Preserving the technical environment including operating systems, original application software, media drives etc. is sometimes referred to as the "computer museum" solution (MIT Libraries 2012d; Brown 2013, 208). However ultimately this is a dead-end strategy, since eventually vintage machines cannot be kept functional indefinitely or will one day cease to exist (Fino-Radin 2013, 110).

The practice of *emulation* is growing in popularity within the diverse digital preservation community, and has already been used extensively among gaming enthusiasts. In as such, emulation depends on the active creation of emulators, which "translate code and instructions from one computing environment so that it can be properly executed in another"(MIT Libraries 2012d). While most emulators available today were specifically developed to allow video games written for obsolete hardware to run on modern computers, fortunately there are a growing number of free, open-source platforms which can be adopted across a variety of communities (MIT Libraries 2012d; Fino-Radin 2013, 110). Emulation is considered a superior solution when fidelity to the original environment is a key factor for object preservation. For many museum collections with original digital objects (such as video installations, video games, internet art, etc), with the passage of time the original software provides important cultural and aesthetic context for those objects. As such, emulation provides a method of higher fidelity and authenticity than migration, and arguably more sustainable than the maintenance and migration of source code for complex digital materials (Fino-Radin 2013, 112). On the other hand, emulators can be rather complicated to implement and may require a large amount of monetary investment, including extensive staff who can create, maintain, and understand antiquated computer systems (Corrado and Moulaison 2014, 51). It is also important to make sure that emulators are also digitally preserved since they too are pieces of technology. Another challenge to emulation can be possible patents or copyright restrictions that may cover the systems to be

emulated (Corrado and Moulaison 2014, 51). Essentially the question of emulation versus migration is dependent on the specific needs and contexts of a digital object.

Replication

The method of *replication* is in of itself a simple and self-explanatory concept: to create duplicate copies of data, although it can be used to mean multiple things. Replication typically refers to both the concepts of bitstream copying and the consortial form of replication known as LOCKSS (Lots of Copies Keeps Stuff Safe). In each case, the intention is to enhance the longevity of digital materials through copying and the use of multiple storage locations (MIT Libraries 2012d). Digital objects that exists as only a single copy in one location is very vulnerable to risk such as software/hardware failure, intentional or accidental alteration, and environmental catastrophes. The data is more likely to survive if it is replicated as a "backup" or stored as copies in several locations. Bitstream copying is commonly referred to as "backing up your data," and simply is the process of making an exact duplicate of a digital object. It is important to keep in mind that bitstream replication is not a long-term technique, but is a necessary component of all digital preservation strategies (MIT Libraries 2012d); it can also be an entry-point for institutions just beginning to use digital preservation strategies (Brown 2013).

The LOCKSS system was originally developed by Stanford University for the preservation of electronic journals, but has since been expanded to many types of digital collections (Harvey and Mahard 2014, 311; Hirtle 2008, 17). LOCKSS is a more secure version of replication that incorporates bitstream copying along with access to the Internet to keep multiple copies of files in distributed locations (across multiple servers/networks) to safeguard against loss. In addition a LOCKSS system regularly compares copies of files to identify any errors (as a checksum), and automatically repairs errors in files from a clean copy held at another participant in the consortia (Harvey and Mahard 2014, 311).

Replication is a simple, yet important method of digital preservation, however inevitably only covers bitstream level preservation and not format preservation. Therefore it should be considered more of a minimum maintenance strategy compared to that of migration or emulation.

Important and Useful Tools from the Digital Preservation Field

Much of the work and functions in a digital preservation program can be automated using specialized software, especially that of metadata, SIP creation, AIP creation, and DIP transfer. Such software can be found as both third-party vendor products or free, open-source options. While paying for vendor services in digital preservation is ultimately very expensive, the setup, maintenance, and user-interface is significantly easier (Brown 2013). Popular preservation-as-as-service (PaaS) softwares include Preservica, Chronopolis, DuraCloud, DSpace, and Portico (Digital POWRR 2014). If one's institution has a much more limited budget, luckily the digital preservation community is dedicated to making such tools accessible to any and all cultural memory institutions, by way of open-source, free software. Important open-source digital preservation software used in the field includes Archivematica, BitCurator, Cinch, Curator's Workbench, DROID, Fedora, and the Internet Archive (Digital POWRR 2013). Two important references that list, describe, and evaluate current digital preservation tools include the IMLS-funded POWRR (Preserving digital Objects With Restricted Resources) Tool Grid and COPTR (Community Owned digital Preservation Tool Registry) (Digital POWRR 2013; COPTR 2014). The POWRR Tool Grid in particular offers insight from the cultural heritage field, and from the perspective of a small-budget institution.

The standards also put forth by major digital preservation initiatives are important to note. Some such standards that have not yet been mentioned include the Library of Congress and their *Recommended File Formats* project, the Digital Preservation Coalition's *Digital Preservation Handbook*, the National Digital Stewardship Alliance *Digital Preservation in a Box* toolkit, and the InterPARES internationally recognized policies and proceedings (Library of Congress 2015b; DPC 2015; NDSA 2015; Duranti and Preston 2008). All of these resources are freely available online and accessible for any individual seeking further help or education about digital preservation.

Conclusion

Managing digital collections in any cultural institution inevitably involves a series of technical considerations when formulating an implementation protocol for the preservation

process. While many museum professionals may find the technical side of digital preservation unfamiliar, the vocabulary and comfort level with understanding the technical requirements can be learned over time. However, the essential steps and concepts behind digital preservation are in many ways analogous to many of the aspects of museum collection management as we know it today: accession, quarantine, records management, conservation, and maintaining an ideal environment for the collection. It is the last two elements that are perhaps the most different from the management of traditional museum collections, for the conservation and ideal housing protocols for digital objects requires new and different sets of tools that are the opposite of the tangible object-based world museum professionals are accustomed to. Nonetheless, the work of digital preservation can be relatable to the ethical mission and day-to-day work of traditional museum collection management.

Although this chapter focuses on how digital preservation has developed over time, with great focus on the technical requirements, a major component to any successful digital preservation program is the institutionalization of digital preservation on a institution-wide basis, or perhaps on a field-wide basis. Management of digital objects requires funding and personnel to perform the technical requirements of any digital repository. These important factors can only be provided for the long-term through institutional commitment, and also by way of collaboration. The following chapters will discuss some of the key components that can aid in the institutionalization of digital preservation for the museum field: collaboration for support, and digital preservation policy.

Chapter 4: Memory Institutions and Collaborative Solutions for Digital Preservation

The literature review in the previous chapter was derived from the library, archive, and information science fields. While very technical and seemingly devoid of literature from the museum field, the cited resources are extremely useful tools that should be shared across the library, archive, *and* museum fields. There simply has not been much academic work published from within the museum field on the issue of long-term digital preservation. As of the publication of this thesis, relevant research has been funded and harnessed predominantly through large library and academic research initiatives. When applying digital preservation practices to the museum profession, the dominance of library and archive science raises the need to address the changing relationship between what the collectively deemed 'memory institutions,' 'social memory institutions,' or 'cultural heritage institutions.'

As such, this chapter will discuss the points of convergence between the distinct disciplines of libraries, archives, and museums to support the development and sustainability of digital preservation efforts in the museum field. This chapter will be divided into three sections: section one sets the stage for the changing landscape of the museum field in relation to other cultural heritage institutions; the second will define 'memory institutions' and discuss why libraries, archives, and museums are popularly categorized as such today; and the third will highlight why collaboration across memory institutions is a best practice to address long-term digital stewardship of cultural heritage.

Setting the Stage

The professional disciplines of library, archive, and museum work can be considered to consist of three distinctive 'cultures' that occupy different places within our social and informational space (Novia 2012, 2; Trant, 2009, 370). Although these institutions first emerged during the 18th century and were conceptually closer as "public institutions," they have evolved through different mandates, collection types, and professional cultures (Duff 2013). They may share common functions (such as collection, conservation, research, and public service), but the differences in professional practices, training, and organizational methods has traditionally

distinguished each field in modern society (Duff 2013). These distinctions have been recently questioned by a number of authors who believe that the commonalities among these institutions are more meaningful than their differences. For example, former director of the Institute for Museum and Library Services (IMLS) Robert S. Martin has published that libraries, archives, and museums "share a common institutional ancestry" that has always been rooted in a common cultural endeavor, and that these disciplinary separations are a relatively recent phenomena (Martin 2007, 81).

Also indicative of the reconvergence of these disciplines, the 2009 International Council on Monuments and Sites (ICOMOS) symposium, proposed a new acronym meant to acknowledge the reemerged commonality of LAMMS - libraries, archives, museums, monuments and sites, which has since been distilled down to just LAMs (libraries, archives, and museums) (Gwinn 2009, 1). One of the reasons why LAM disciplines have been re-identified can be attributed to the 'digital age' and the resultant paradigm shift in the cultural sector, in which the potential vision for 'ubiquitous knowledge' was fully realized (Kirchhoff et al. 2008).

Since the implementation of Web 2.0 and the semantic web in the early 2000s, the world has entered an age of transparent and rapid access to information, often referred to as the "Information Society" (MacDonald and Alsford 2010, 72-3). This major paradigm shift in the way our society interacts with technology and acquires knowledge has had a major impact on cultural heritage institutions such as libraries, archives, and museums (Iljon 1999, 23). Certain common digital technology tools, such as the Internet, have enabled new levels of user access that impose similar expectations and approaches to cultural materials and information, which memory institutions cannot ignore. We are now in an era of access. This paradigm shift in the world has broad implications for the LAM community, including the openness and availability of cultural materials, and whether they can be accessed, or should *become* accessible (Iljon 1999, 23). The access mandate of our digital age has been equally experienced among libraries, archives, and museums. LAMs all share the same duty in the digital age to be proactive in disclosing their holdings to the digitally-engaged public. Importantly, the duty and mandate have been discussed primarily within the ethical codes of all the memory institutions, which has had

an impact on the way their missions and roles have been redefined in recent years (Iljon 1999, 23).

Libraries, archives, and museums are still three distinguishable entities; nonetheless, the boundaries between them are slowly blurring due to the prevalence of access mandates, which involve integrated technology systems. During the last 15 years, considerable information and dialogue on the blurring boundaries between LAMs has emerged through conferences and journal publishings. One of the first was the Library Automation Group Conference in 2000 (*Archives, Libraries and Museums Convergence*), which looked in detail at cooperative digital projects across cultural heritage institutions (Higgins 2012, 3).

An especially important study about the merging practices between LAMS was the Online Computer Library Coalition research project called *Beyond the Silos of the LAMS: Collaboration Among Libraries, Archives and Museums*. This project was carried out between 2008-2010 to examine collaboration in the context of LAMs that have common organizational governance, and who were already committed to working together (OCLC 2011; Higgins 2012, 4). The project held five mediated workshops to identify a shared vision of "seamless collections access and community engagement." These workshops highlighted how collaborative activity has been project-based, and mainly focused on shared creation and *storage of digital materials*, and on search tools for their discovery (OCLC 2011). This OCLC study encapsulated the trend towards an ever-increasing acquisition of born-digital materials, which is causing traditional boundaries between memory institutions to blend (OCLC 2011; Higgins 2012, 4).

Another landmark report created in 2008 was funded by the International Federation of Library Associations and Institutions (IFLA) which included studies across the U.S, Canada, the U.K., Russia, Denmark, Norway, Sweden, Germany, Italy, Spain, South Africa, Australia, and New Zealand (Yarrow, Clubb, and Draper 2008). The research brought international attention to the dialogue on LAM best practices, and created a guide to successful collaboration, discussions on the benefits and risks of collaboration, and a list of sources to consult (Yarrow, Clubb, and Draper 2008). The most important collaborative activity that was identified takes place in event programming, integrated facilities, and digital resource creation (Higgins 2012, 4).

Since the onset of the many studies and conferences, 'memory institutions' and collaboration between them, has been encouraged at the highest professional level. The IFLA, ICA (International Council on Archives), and ICOM (International Council Of Museums) along with the more recently formed ICOMOS and the Coordinating Council of Audiovisual Archives Associations (CCAAA), have formed the International NGO Working Group on Convergence in 2008 (Higgins 2012, 3). This group is now known as the Libraries, Archives, Museums, Monuments and Sites (LAMMS) Coordinating Council, which works to find solutions on copyright and other legal matters, as well as work with the World Intellectual Property Organization (WIPO), and lobby for measures to ensure the safety of cultural heritage within the Blue Shield and UNESCO. In addition, the LAMMS Coordinating Council works on the development and standardization of global digital libraries (Higgins 2012, 3).

From all these significant studies about cross-disciplinary efforts, most if not all conclude that successful collaboration depends, ultimately, on the ability of the participants to identify both their commonalities as well as substantive differences in services and practices in order to build partnerships that recognize and respect these factors. It is also shown in the research cited above that the impact of the "information society" and the rapid evolution of technology are profound agents because 1) the behind the scenes technical challenges make LAMs more alike than different, and 2) this makes collaboration more likely and necessary. While libraries, archives, and museums may remain physically independent of each other, the ubiquity of digital technology in memory institutions is profoundly shifting the boundaries and changing definitions around "sharing" professional practices. To further this concept of 'ubiquitous technology,' in many ways the convergence discussed by the greater memory institution field has emerged from commonalities in strategic and also technical aspects that LAMs share (Iljon 1999); in particular, the software, hardware, formats, and digital objects across LAMs often closely resemble each other, whereas traditionally their physical holdings have been defined quite separately. Additionally, our custodial habits around new media and technology for long-term preservation will continually merge, and as a result the digital preservation practice of one institution is becoming relevant to other kinds of institutions (Harvey 2014, 3). As succinctly said by Sarah Higgins for the 20th Anniversary of UNESCO *Memory of*

the World Programme and conference *Memory of the World in the Digital Age: Digitization and Preservation* in 2012:

> *"...it is in the area of catalogue federation and digital content creation and management that collaborative projects have started to lead to shared services. As professional best practice develops and projects mature, LAMs are starting to converge over the need to provide for the long-term access, use and reuse of their digital materials, both digitized and born-digital, through the new disciple of digital curation"*(Higgins 2012, 2).

While it is commonplace to express that the ubiquity of digital technologies is changing our societies and our ways of comprehending things, it is somewhat less commonplace to actually create and experience new best practices that break from the comfort of habit (Iljon 1999, 23). Libraries, archives, and museums will need to actively break the old habit of viewing each other as separate disciplines, and accept that they are more similar than different. Because of a long history of separation, the increased permeability of previously rigid boundaries will likely progress slowly. Hopefully, there is will continue to be an upward trend as a result of growing openness and cross-fertilization of practices among cultural heritage fields. Collaboration is a hallmark of the growing digital preservation environment and is an essential characteristic of any digital management program.

Defining Memory Institutions:

As the cultural heritage field is shifting today, a new set of definitions and vocabulary is developing in tandem. Terms like 'open access,' 'open source,' 'digital asset management,' 'digital curation,' etc. from the technology industry have become integrated within the humanities. The term 'social memory institutions,' while not necessarily a new term, is certainly used often today as a way to collectively group libraries, archives, and museums as having common missions and ethical codes.

In 2000, from the early stages of the current digital age, Lorcan Dempsey published a paper that first popularized the term *memory institutions* as we know it today:

> *"Archives, libraries and museums are memory institutions: they organise the European cultural and intellectual record. Their collections contain the memory of peoples, communities, institutions and individuals, the scientific and cultural heritage, and the products throughout time of our*

> *imagination, craft and learning. They join us to our ancestors and are our legacy to future generations...Memory institutions contribute directly and indirectly to prosperity through support for learning, commerce, tourism, and personal fulfilment... They are social assembly places, physical knowledge exchanges, whose use and civic presence acknowledge their social significance, and the public value accorded to them."* (Dempsey 2000).

It would seem from this definition that 'memory institutions' are essentially synonymous with 'cultural heritage institutions'; however, 'memory institution' seems to emphasize the responsibility of not only holding objects, but also capturing the greater manifestation of human memory.

A more recent iteration of the term *memory institution* was published in a 2009 issue of *Museum Management and Curatorship* magazine by digital culture strategy consultant Jennifer Trant:

> *"The memory institutions has captured the imagination of policymakers as a powerful metaphor for the social role of libraries, archives, and museums. Charged with giving access to and shaping shared cultural heritage, memory institutions are sometimes characterized as storehouses, reservoirs to be tapped for many different purposes, from education to entertainment..."*(Trant 2009, 1).

These two definitions of social memory institutions establish that such entities are core elements of the civic fabric that is woven to encapsulate the long-term public identity of communities, cities, and nations. Importantly, 'memory institution' has also been identified by policymakers to be an important term used to unite the missions of LAMs.

The field of 'social memory,' from which we borrow to create the term 'memory institution,' actually emerged in the 1920s and gained momentum in the 1970s (Ippolito and Rinehart 2014, 14). Social memory is how and what societies remember - the long term memory of our civilization. It is the vehicle by which civilizations "carry forward their social traditions, commercial arrangements, and political operations from moment to moment, year to year (and if they are lucky) century to century. It allows a civilization to persist beyond the lifetime of one individual or generation"(Ippolito and Rinehart 2014, 14). Although this definition is vague, and while there is not a consensus about the boundaries of social memory, a consensus does exist that most social memory is sustained and transmitted through the conscious efforts of institutions

like libraries, archives, and museums, hence the term 'memory institution'(Ippolito and Rinehart 2015, 15).

Within the field, social memory is typically divided into two umbrella categories: formal and informal. Formal social memory is considered "canonical," and this is the form that is most associated with the materials stewarded by libraries, archives, and museums. This concept carries over from LAM's historical significance as society's "cabinets of wonder" (Wunderkammer), or our collective memory banks, material encyclopedias, and now in the digital age, databases of civilization. In contrast, informal social memory is characterized by folklore and is typically distributed in popular forms of remembering that are harder to tangibly define when compared to formal social memory (Ippolito and Rinehart 2014, 15).

The limitations and challenges for convergence between libraries, archives, and museums has previously been couched in terms of perceived differences in how libraries, archivists, and museum professionals view their collections, their users, and their missions (Martin 2007, 81). Libraries, for example, traditionally hold collections of mass-produced, textual and published materials; these encyclopedic collections provide access to the world's knowledge for a broad, often general public, and the audience and clientele of any particular library is diverse (Trant 2009, 2). Today it is also common for libraries to have special materials such as manuscripts, maps, and pictures. Audio, video, photographic, and of course digital materials are also commonly collected in libraries today (Harvey and Mahard 2014, 66). Libraries have developed sophisticated systems for assisting users to retrieve specific resources that correspond to their needs through automated catalogues and federated searches. The role of the librarian is to facilitate the discovery phase of the research process, but the actual research and learning on the user side takes place in an unmediated manner (Trant 2009, 2).

On the other hand, archives seek to preserve corporate and individual memory, usually in the form of original (primary) resources such as administrative records and cultural, historical, and personal records. The materials collected in archives are largely unique, and although typically paper based, archives can also contain objects like clothes, jewelry, badges, etc. (Harvey and Mahard 2014, 66). Since archival materials provide primary evidence of historical occurrences or transactions, their collections do not circulate, and therefore, the user experience

is generally mediated by archivists. (However, the influx of digitization initiatives has also added the universal platform of the Web to the archive field, which has changed the mode of access to many archival materials.) Archives organize their collection based on the principle of provenance -- maintaining *fonds* separately in the order given by their creators-- which has guided the development of finding aids that are hierarchically structured (Trant 2009, 3).

Finally, museums collect materials, typically tangible objects, but increasingly also audio, visual, and digital objects that are selected according to their aesthetic, historical, or educational value (Harvey and Mahard 2014, 66). The museum profession places heavy emphasis on keeping collections in excellent shape and ensuring the originality of objects. Because the objects are high-quality and often rare, the handling of museum collections is generally restricted to museum staff or credentialed researchers. In contrast to libraries, members of the public can only see objects when the museum exhibits them, and the public also cannot handle objects (Harvey and Mahard, 2014, 66). Therefore, the museum experience is a highly mediated one; unique artifacts are presented and assembled according to a curatorial directive and usually part of a supported argument, narrative, or theme (Trant 2009, 3). Visitors are guided through collections with didactic educational materials that provide context and meaning, oftentimes as the interpretation of the curator. Museums often single out specific works for special attention, and visitors are alerted to why a masterwork is important. The province of the museum professional is mainly within collection documentation, which is recorded for future professional and scholarly use. Large museums often have their own libraries to support the research of its staff and archives to document institutional history (Trant 2009, 3).

In sum, libraries, archives, and museums have established different communities of users who expect divergent services through these distinct modes of collecting, organizing information, and professional mediation. While perhaps overly generalized, these models exemplify how each institution's definition of access assumes a particular kind of use, and therefore different assumptions about patrons' needs and their preferred methods of interaction with the organization and the systems that support them (Trant 2009, 4). While the differing nature of collections in museums, archives, and libraries has contributed to diverging professional practices, as discussed in the previous section of this chapter, these boundaries are

being challenged by the changing demographic of collections and user platforms in the cultural heritage sector.

The reidentification of LAMs as social memory vehicles has contributed to the slow disintegration of their separate identities. Leadership in the LAM fields can help to establish a necessary level of trust that cultivates a spirit of mutual understanding and respect; leadership will be needed to enable library, archive, and museum's future ability to meet the high-level preservation mandates imposed upon them post-Web 2.0. Today, the international bodies representing the three professions (IFLA, ICA, and ICOM) each publish a code of ethics to guide its respective field with recommended values, principles, and activities. These codes reveal how institutional and professional practices remain divergent in ethical emphases; however, despite the sustained differences, the three Codes also reveal the cross-sectoral mission to care for collections and to provide access (HIggins 2012, 6). The universal role of LAMs as "sustained institutions to collect, organize, preserve, and provide access to knowledge-bearing objects" highlights the potential for coevolution of their expertise, methodologies, and tools for organizing and interpreting knowledge as collective social memory institutions (Higgins 2012, 9).

Technology presents a perfect platform for a reunification of LAM professional practices. According to Kirchhoff, Schweibenz, and Sieglerschmidt, it no longer matters where one finds their information, as long as they find it. This is all because the digital realm levels the playing fields for data access; in other words, it is no longer relevant whether the original materials are in a library or a museum or an archive because finding information on the Web is the result of a uniform set of digitization and preservation activities that feeds into the open access to data. The new form of digital heritage sets the stage for the new so called 'memory institution' (Kirchhoff 2009,, 252).

Collaboration is Necessary to Ensure Preservation

The purpose of recognizing how cultural heritage institutions have been redefined within the social memory field is to acknowledge not only the trend towards 'convergence' in practices and collections, but also to acknowledge that *collaboration* between libraries, archives, and

museums will be a critical factor for whether the greater museum field can achieve digital preservation to the level of a 'Trusted Digital Repository,' which arguably, is the ideal level of preservation for medium to long-term stewardship. The already large information world grows larger with the rapidity of material published digitally and on the Web. Since the library and archive field has invested much time and money to create useful tools, resources, and tutorials for digital preservation, it is logical that the museum field share and use those resources. Awareness of and need for digital preservation in museums will only increase as society continues to use technology as a tool and as a mode of self-expression. Although the separate missions of libraries, archives, and museums have grown more similar, and a paradigm shift towards collaboration has risen within the last 15 years, there is still an education gap between the museum field and the library/information science fields. Whereas it is generally required that all librarians and archivists have a graduate degree in order to work in the field, the same has not always been true for museums (Novia 2012. 5). The preceding formalization of library and information science professional degrees to that of Museum Studies is perhaps one reason for the gap; librarians have been engaging in scholarly dialogue about their profession for a long time. However, hopefully as the social memory institutions grow closer in digital collecting practices, so will the strategy of learning across the board (Trant 2009, 12). Jennifer Trant suggests in her article published in a 2009 issue of *Museum Management and Curatorship* magazine:

> *"Both Museum Studies and Information Science have a strong tradition of linking theory and practice, and of placing students in the field to apply their knowledge. When designing new curricular content, this strategy of learning and doing could be emphasized through teaching methods that cross institutional boundaries and draw upon strengths of each traditional specialization. It also meshes well with the need for lifelong learning in a technological environment of continuous change"* (Trant 2009, 12).

Emerging and future professionalism aside, if the museum field seeks to improve its stewardship of digital materials, collaboration will be required (Higgins 2012, 6). As proven by the redefining of libraries, archives, and museums as 'LAMs,' and as 'memory institutions,' no matter what the cultural differences, these institutions ultimately share the common goal of collecting and preserving our cultural heritage (Novia 2012, 5). For example, the Open Archival Information System Reference Model (OAIS) was created as a general, common framework for all digital preservation applications. It provides a platform-independent model for the information architecture and the organizational requirements of long-term care of digital materials. The basic nature of the OAIS Reference Model has spurred a large amount of collaborative activity between vendors, information scientists, IT professionals, big data research centers, *and* cultural heritage institutions to formulate tangible solutions and technical environments based upon a universal vocabulary and data model from the OAIS (Higgins 2012, 17). Collaborative materials created out of the OAIS model include storage solutions, LOCKSS systems, and metadata standards for creating information packages and describing preservation activities.

The collaboration happening around the use of the OAIS Reference Model follows the model proposed by the authors of the 2008 OCLC LAM collaboration study, *Beyond Silos of the LAMs: Collaboration Among Libraries, Archives, and Museums* (OCLC 2008). Building from issues raised at the RLG Forum in 2005, Zorich, Waibel, and Erway created a framework for collaboration based on a continuum model which identifies steps and activities that accumulate trust, investment, risks, and also benefits that accrue as an institution moves forward on the continuum (OCLC 2008). This continuum has been cited in many scholarly papers that discuss the issues of changing users, increased dependence on technology, open access to collections, digitization, etc. within memory institutions (Duff 2013; Novia 2012; Higgins 2012). For example, Sarah Higgins' paper for UNESCO in 2012 cites Zorich, Waibel, and Erway's continuum model as a best practice for libraries, archives, and museums who wish to incorporate active digital curation into their digital stewardship (Higgins 2012, 4). Her paper proposes that *digital curation* is the main "change agent' that will bring the inevitable full convergence between cultural professions, as they move through the digital content and management continuum. This 'inevitable convergence' of LAMs is a concept bolstered by Zorich, Waibel and

Erway's continuum model. Their view is that the organizational commonalities of LAMs, along with certain "change agents" that raise awareness of collaborative potential, will lead to collaborative efforts that act as catalysts for long-term projects and therefore convergence (Zorich et al 2008).

The steps in the continuum are conversance, contact, cooperation, coordination, collaboration, and finally convergence (Zorich et al 2008). The first step, conversance, is an activity that builds understanding of the professional landscape by keeping abreast of the most current developments in LAMs through channels such as media, newsletters, RSS feeds, social media, professional conferences, etc. (Higgins 2012, 3). Once members of institutions become more conversant in the possibilities afforded by collaboration, they may proceed to the next step, contact, which then leads to progress along the continuum (Higgins 2012, 3). Benefits only accrue as one moves along the continuum, and as trust and investment develops between LAMs, so does the vision for a committed, shared future (Zorich et al 2008).

Regardless of how collaboration happens (following the continuum model, or not), like Sarah Higgins suggests in her paper, recent studies reveal that the main "change agent" that inspires collaborative activity revolves around preservation and digital projects (Novia 2012, 3; Higgins 2012; Yarrow, Clubb, and Draper 2008). Collaborative practices between libraries, archives, and museums will enable memory institutions to close the knowledge gap between them, as well as encourage professionals to find solutions for complicated digital infrastructures amongst a wider network. Outside of technical methodologies, libraries, archives, and museums can share on the digital front to make up for the lack of resources, such as funding and space, required for digital mandates (Novia 2012, 8). Collaboration makes it possible for institutions to take advantage of professional customs and expertise from across a far-reaching group of practitioners. Because of the iterative, active nature of digital curation, Higgins see this as the change agent that will move LAMs along the collaboration continuum towards convergence. The technical challenges and investment required for long-term digital preservation (curation included) means that cultural heritage organizations are forced to pool experience and expertise to develop best practices, training, tools, and shared services (Higgins 2012, 17).

Examples of Collaborative Digital Preservation

Noteworthy examples of collaborative digital preservation efforts include the LOCKSS system, namely the MetaArchive, NINCH, the Colorado Digitization Project, and MOAC. The MetaArchive was formed in 2004 by six libraries in the southeastern United States to develop a digital preservation solution for their special collection materials. The outcome of this effort is a community-owned, community-led consortia comprised of libraries, archives, and other digital memory institutions. The MetaArchive works cooperatively with the Library of Congress through the NDIIPP Program to achieve a secure and cost-effective repository that provides long-term care of digital materials, not by outsourcing, but through active participation of their own content (MetaArchive 2014). Their methodology is a LOCKSS software developed by Stanford University that allows members to embed the technical redundancy infrastructure within actual memory institutions instead of through an outside vendor. Essentially each institution in the Cooperative runs a server linked securely to the network, and as each member readies content for ingest, the content source is visited by seven of the network's servers, from which a replicated and preserved copy is made (MetaArchive 2014b). The seven servers aid with regular checksums and curation to aid in detecting preservation issues, and also ensures versioning, or making sure a copy can be recovered (MetaArchive 2014b). The MetaArchive Cooperative was founded to encourage archives, libraries, and museums to build their own preservation infrastructures and expertise without outsourcing such a core service to vendors. To keep digital preservation affordable to many institutions, they offer three levels of membership ranging from $6,000-8,000/per year on a 2 Terrabyte example (MetaArchive 2014c). The MetaArchive also provides useful tools (such as its own TRAC audit tool) and a network of people to call upon for troubleshooting. While this cooperative seems like an innovative solution that keeps digital preservation secure within the memory institution network, and retain complete ownership/control over its digital assets, as of 2015 there were no museums listed as members of the MetaArchive (MetaArchive 2014d).

The next example of successful collaborative efforts is The National Initiative for Networked Cultural Heritage, or NINCH. NINCH is a US digital networking/digital cultural heritage endeavor that aligns many American organizations (many of which *are* museums) and

aims to provide leadership in the digital world to build a framework for collaboration (NINCH 2003b). The two major projects that have come from NINCH are an international database of digital humanities projects, and a document titled, "Guide to Good Practice in the Digital Representation and Management of Cultural Heritage Materials" (Yarrow, Clubb, and Draper 2008). The latter of these important projects is an excellent example of how collaboration across memory institutions can produce useful guidelines that serve as educational and policy-building resources (Yarrow, Clubb, and Draper 2008). The NINCH guide defines issues of digital preservation including migration, using non-proprietary digital formats, and metadata capture as long-term management strategies. The main museum participant in the creation of the NINCH guide was the Berkeley Art Museum/Pacific Film Archive, but the guide does also mention the noteworthy Colorado Digitization Project (CDP) as an excellent example of how various partners can combine talents and resources for digitized cultural heritage management (NINCH 2003, 165).

Formed in 1999, the Colorado Digitization Project brings together a variety of institutions from within Colorado with the overall aim of producing a digital resource that encapsulates material from collections of museums, libraries, and archives. Collaborations such as the CDP enable large-scale digital programs to realize its full potential of digital management, and long-term access. To coordinate between all the participating institutions, a CDP staff member acted as the project manager (NINCH 2003, 165). Skills and resources were shared through collaborative training sessions and labs throughout Colorado. These training resources provided a critical link between all the participating institutions because they enabled smaller institutions to learn how to use the technical equipment in the larger institutions, and enabled participation. Ultimately, the collaboration allowed large public libraries that house state-of-the-art equipment and small local museums with no equipment at all to work together in creating a unified, standardized, and high quality digital resource. The management of this consortium was controlled from a central point which allowed for the goals and deliverables to be agreed upon across the board. These goals and deliverables that all the institutions in the CDP adhered to include types of metadata, controlled vocabularies, file formats, standards for interface design, and guidelines for quality control (NINCH 2003, 165). While the CDP does not exclusively

focus on the preservation side of digital cultural heritage, it is a prime example of a successful LAM collaboration project. Unfortunately, the CDP is no longer active; by 2007, the CDP merged into the Bibliographical Center for Research (BCR), and then, by 2010, the BCR established a partnership with LYRASIS, a non-profit member organization that manages digital content on a subscription-based membership. However, the existence and success of such initiatives provide a beacon of hope for future collaborative preservation opportunities.

Another collaborative effort that comes directly from the museum world is the Museums and Online Archive of California (MOAC) from Oakland, California. This project was headed by Richard Rinehart and collaborated between 13 partners including the Berkeley Art Museum (at which Rinehart worked), the Japanese American National Museum, the Oakland Museum, the San Francisco Museum of Modern Art, and the UCLA and UC Berkeley museums, libraries, and galleries (Rinehart 2003). Funded by an IMLS grant in 1999, MOAC sought to enhance interoperability, integration, and seamless access to digital library and museum resources in order to lower the cost of participation for museums and libraries wishing to collaborate (Rinehart 2003). In order to raise the ability for museums and libraries to share digital content, a need for easily attained technical and descriptive metadata standards was necessary. Rinehart developed a 'community toolbox,' or a practical software tool, using FileMaker Pro (which is used frequently in the cultural sector) that allowed museums and libraries to easily produce standards-based metadata for content sharing, called the Digital Asset Management Database (DAMD) (Rinehart 2003). This tool provides basic digital asset management by easily transforming collections information into a variety of standards-base XML formats, such as METS and OAI. The Digital Asset Management Database is open-source and free to cultural organizations. Once downloaded, the DAMD has a specially designed export/transform function that allows organizations to customize the tools for themselves (Rinehart 2003). While this project is not focused acutely on digital preservation activities, it is an important initiative towards the convergence of LAM missions and long-term access to digital materials.

The collaborative efforts of groups like NINCH, MOAC, and the MetaArchive are encouraging examples of how libraries, archives, and museums can work together to lessen the burden of digital preservation. While there is still a larger gap between participation in the

museum field and the library/archive fields, it is possible for these memory institutions to share strategies to manage digital media, digitized collections, and born-digital materials as a complementary part of a unified resource. Dempsey aptly predicted early on that libraries, archives, and museums will address the issues of the digital age within their own curatorial traditions and organizational context; however, they can collectively develop strategies for the initial investment and managed intervention that is required of long-term digital stewardship (Dempsey 2000). Together, LAMs can ensure that 'born-digital' documents and artifacts become integrated into the cultural record through various levels of digital preservation activity that will help to keep them accessible, and to become a permanent part of the cultural memory of future generations.

Chapter 5: Digital Preservation Policy: The New Collection Management Policy?

As implied by the chapter title, the need to draft and implement a digital preservation policy is of equal importance to that of collection management policy for a museum. Considering the consistent parallel between the care of traditional museum collections and the care of digital collections highlighted in this thesis, creating similar parallels with high-level policy will be presented as a strategy to encourage the museum field to exercise its responsibility and duty of care for born-digital collections. In the 20th century, the professionalization of museums directed the field to implement collection management policy that addresses the ethical handling, accession, deaccession, storage, and conservation of the materials stewarded in the name of the public trust (Malaro and DeAngelis 2012). Now still in the early part of the 21st century, a similar call for policy is again needed, but this time in regard to the responsible stewardship of digital collections and assets. This chapter will discuss the importance of policy in the museum field, and especially for the implementation of digital preservation. Policy implementation will be placed within the context of institutional readiness assessment. The last section of this chapter will provide useful resources for modeling a digital preservation policy.

Currently, most museum collection management policies do not address the issues of digital preservation or digital stewardship. In 2010, a European Union-funded survey resulted in a white paper written by the organization Planets (Preservation and long-term access through networked services) that evaluated the current status of digital preservation readiness across over 200 collecting institutions. The survey found that only 3% of survey participants were museums, none of whom had digital preservation policies in place. Comparatively, 65% of national archives and 55% of libraries who participated did have digital preservation policies by 2010 (Sinclair 2010). The statistics confirm that libraries and archives are the leaders in digital preservation best practices (Planets, 2010), including policy. Nevertheless, as digital assets in museums increase, and practices across LAMs converge, the creation of digital preservation policy in museums will become increasingly important.

The Important Role of Policy in the Museum Field

In general, memory institutions ensure the sustainability of their collections and maintain a high level of public trust upon the realization as trustworthy repositories, which requires an organizational environment that protects the physical and moral integrity of its collections (Higgins 2012). That is the physical and intellectual security of collections, through effective collections management underpinned by policies that address the institution's commitment to cultural stewardship.

As noted earlier in the chapter, advocacy for policy in museums has mainly been in the form of the collection management policy. Two major figures for best practices in the museum field are scholars Marie Malaro and Ildiko DeAngelis, whose book *A Legal Primer on Managing Museum Collections* promotes prudent best practices from an ethical and legal standpoint, including the importance of policy for asserting a museum's good work outwardly to the public, as well as inwardly to its personnel (Malaro and DeAngelis 2012, 6). These scholars define a collection management policy as "a detailed written statement that explains why a museum is in operation and how it goes about its business. The policy articulates the museum's professional standards regarding objects left in its care and serves as a guide for the staff and as a source of information for the public"(Malaro and DeAngelis 2012, 46). Because museums are commonly structured as nonprofit organizations, they are set up as trusts for cultural objects and records held in the name of the people, and thus awarded with certain financial and tax privileges (Malaro and DeAngelis 2012, 6). This concept of acting as a 'public trust' is a significant reason why policies are necessary and recommended documents/practices in the museum field in order to have written proof of the museum's prudence. Malaro and DeAngelis affirm the trustee relationship between museum institutions and the public: "In its pure form, a trust relationship imposes a high degree of responsibility on the trustee. The trustee is charged with affirmative duties to protect, preserve, and increase the trust assets" (Malaro and DeAngelis 2012, 6). In this quote, Malaro and DeAngelis outline the main "duty of care" entrusted to the governing body (and arguably also the staff) of a museum. A collection management policy is therefore an effective leadership exercise and tool that outlines an institution's "duty of care," by covering its legal and ethical practices. Should the public ever question a museum's ability to care for its

objects and records, publicly available policies will be especially important to defend and protect the institution.

Ultimately collection management policies are preventative measures. Their adoption and implementation provides clear direction and prevents poor decisions that are not only hard to reconcile, but affect the public's opinion of the museum (Malaro and DeAngelis 2012, 46). Policies are considered board-approved documents that should take into account the mission and goals of the overall institution, so often they contain general rather than specific statements that together functions as a *guiding* document. This is not to be confused with more practical-level plans, which document directly actionable protocols and are often collection-specific (Corrado 2014, 22). Policies are voted on and/or approved at a high level, and thus act as a guiding document for topics that require governance approval such as staffing and funding, but also directs staff in their overall responsibilities, roles, and collection directives. Policies may be general, but they will often guide the building of plans, which are the important road maps that take into account practical implementation (Corrado 2014, 22). Since upper management is involved in the passing of policies, the creation of such can serve as the required stamp of approval needed to facilitate and acquire institutional commitment and resources for the long term. In addition the very exercise of creating (and then later reviewing) the terms of a collection management policy "provides a worthwhile educational opportunity for museum officers and staff. All who participate in writing and revising a collection management policy cannot help but emerge with a better appreciation of their respective roles and a firmer grasp of important basic principles" (Malaro and DeAngelis 2012, 46). The rules around differentiating policies from plans is not always rigid, and when necessary plans can also be directly based on mission, goals, or objectives of an institution, bypassing a formal written policy (Corrado 2014, 22).

The professional guidelines for museum policy goes beyond the literature, and can also be found within the leading professional associations such as the American Association of Museums (AAM) and the International Council on Museums (ICOM). The AAM promotes the practice of accreditation on a Continuum of Excellence, which is a pathway of programs that recognizes and promotes the museum field's commitment to standards, professionalism, best

practices, and helps nurture a culture of excellence (AAM 2015). One of the programs along the Continuum of Excellence is the Museum Assessment Program, or MAP, which is a one-year process of self-assessment and consultative peer review that analyzes a museum's strengths, weaknesses and provides a roadmap for improving operations and meeting standards (AAM 2015). The MAP program allows a museum to choose a type of assessment, one of choices of which focuses on collections stewardship (AAM 2015b). As part of this assessment, MAP helps the museum improve the following: its ability to raise funds to support collections, improve collections stewardship, prioritize long-term collection management issues, and most importantly develop, review, and/or revise collections *policies* (AAM 2015b). Policies are highlighted as one of the necessary documents a museum needs in order to qualify for the MAP program, and also as a standard needed to move onto the next step of the Continuum of Excellence, *Core Documents Verification.* The AAM MAP program is one example of how professional leaders in the museum field prioritize policies as one of the main functions towards professional accreditation.

Another such example can be found in The International Council of Museums' (ICOM) published *Code of Ethics for Museums*, which is revised regularly, the most recent iteration was published in 2013 (ICOM 2013). The Code reflects principles generally accepted by the international museum community and is considered a minimum standard for all museums as a series of guidelines for desirable professional practice (ICOM 2013, IV). The first two parts of the Code of Ethics encourage the enabling of policy in museums: "museums preserve, interpret and promote the natural and cultural inheritance of humanity"; and "museums that maintain collections hold them in trust for the benefit of society and its development." The ICOM Code of Ethics champions the most succinct description for why policies are important documents to reflect a museum's upholding of standards that it is worth directly quoting:

> *"Museums are responsible for the tangible and intangible natural and cultural heritage. Governing bodies and those concerned with the strategic directio nand oversight of museums have a primary responsibility to protect and promote this heritage as well as the human, physical and financial resources made available for that purpose. Museums have the duty to acquire, preserve and promote their collections as a contribution to safeguarding the natural, cultural and scientific heritage. Their collections are a significant public inheritance, have a special position in law and are protected by international legislation. Inherent in this public*

> *trust is the notion of stewardship that includes rightful ownership, permanence, documentation, accessibility and responsible disposal."* (ICOM 2013, 1-3).

Interestingly, the ICOM Code of Ethics specifies both the "tangible" and "intangible" types of cultural heritage. Although there is no specific language in the Code on the stewardship of *digital* assets (objects, documentation, and records), it can be deduced that as intangible materials, digital assets fall within ICOM's Code of Ethics. This thesis has taken great pains to assert that more and more of museum assets are falling within this "intangible" category. While one could argue that the Code could do better by the museum field by directly using the terms "digital stewardship" or "digital preservation", the recognition of the museum's role and responsibility towards "intangible" materials is still clearly evidenced here.

Considering the well-established advocacy for collection management policies within the museum field, it can be argued that digital assets ought to be included within the same considerations made for the already de facto standards of collection stewardship. While the library field typically differentiates policies (one for material and another for immaterial assets), the ICOM Code of Ethics demonstrates another school of thought -- that collection management policy and digital preservation policy could be combined into one. If the museum field has demonstrated its commitment to professionalism and ethical handling of its assets, the future of collection management policies will need to also include digital collections, documentation, and records because of the increased normalization of digital technology in our world today. The momentum towards such a change in museum policies perhaps begins with the recognition of "ownership" of digital assets on the same level as normal museum collections and documentation (Kenney and McGovern 2003).

The Importance of the Digital Preservation Policy

The 2010 Planets survey results white paper, "The Digital Divide", ably demonstrates the impact of policy for the implementation of productive digital preservation. The Planets survey states in plain language the need for policy in order to guide the future directives of a growing landscape of digital materials in cultural institutions.The survey of over 200 EU institutions found that the volume of digital content that organizations expect to archive will

increase 25-fold between 2010 and 2020 (Planets, 2010). At the time of the survey, only 27% of the participating organizations felt they had complete control over the file formats that they will accept and store in their digital archives, because of a lack of standards and policy available (Sinclair 2010). American initiatives, such as the Northeast Documentation Conservation Center, Institute for Museum and Library Services, the Library of Congress, the American Institute for Conservation, and the Center for Research Libraries recognize and echo the same need for digital preservation policy as the Planets findings.

More key language and points about the importance of digital preservation policy can be found in the Planets white paper: "A policy is a vital first step towards tackling digital preservation challenges. Articulating a policy helps to build a business case, which may lead to obtaining a budget and implementing a solution" (Sinclair 2010, 3). The focus of this statement on the foundations of funding for digital preservation is an interesting case to make for the creation of policy that is different from the normal focus of the museum field on ethical and legal handling of its assets like that asserted by Malaro, DeAngelis, ICOM, and the AAM. The Planets survey and white paper stated that organizations with a digital preservation policy are more likely to include digital preservation in their operational, business, and financial planning (Sinclair 2010, 9). In addition, they are three times more likely to secure a budget for digital preservation, four times more likely to invest in a solution in the immediate future, and three times more likely to have a long-term solution in place (Sinclair 2010, 9). Also in the survey's findings, institutions *without* a digital preservation policy are four times more likely to have no experience or be unaware of the challenges presented by digital preservation, three times more likely to have no plans for long-term management of digital materials, and more than twice as likely to put off investing in a digital preservation solution for more than two years (Sinclair 2010, 9). This last point poses a particular challenge since many digital assets can become obsolete in as quickly as 2-5 years.

Digital preservation policy plays an important role for the OAIS Reference Model, which inspired the concept of the Trusted Digital Repository (TDR) and the Trusted Repository Audit and Checklist (TRAC). For an institution to be OAIS compliant, it is essential to have documented policies and procedures for preservation (CCSDS 2012; Corrado 2014, 50). While

the OAIS Reference Model maintains a general perspective and does not specify what these policies should look like, the OAIS considers having strong policies in place a way to prevent errors and to add to the trustworthiness of a repository (Corrado, 2014: 50). Institutions that wish to qualify as a Trusted Digital Repository, and to pass the Trusted Repository Audit and Checklist (TRAC) must have a policy in place:

> *"Whether archival storage is centralized or distributed, it relies on a robust and well-documented policy for storage and maintenance and for the expected level of service...The policy must include systems for routine integrity checking of the bytestream, once it has been established within the storage facility, redundancy of data storage, and for disaster preparedness, response, and recovery"* (RLG-OCLC 2002, 26).

Since TRAC is based upon OAIS compliance, the standards required between them are naturally similar in nature. While the OAIS defines the specific steps, standards, and requirements for a digital archive, TRAC additionally includes considerations for managerial, organizational, and administrative standards required to establish trust. One of the main conclusions drawn from the 2011 TRAC Magenta Book is that a *trusted* digital repository is more than just an organization responsible for storing and managing digital files, but it must also pledge to provide reliable, long-term access to managed digital resources from its designated community, now, and into the future. (CCSDS 2011, 37). The point made here, regarding established *trust* by the TRAC standard, is parallel to the same standards already asserted within the museum field by Marie Malaro, ICOM, and the AAM. Furthermore, similar to Malaro and DeAngelis' characterization of the collection management policy, according to TRAC, having a digital preservation policy proves an institution's adherence to the necessary level of legitimizing trust and standards. In the same fashion that Malaro encourages museums of all sizes to have written collection policies, the OCLC-RLG paper on attributes of a Trusted Digital Repository also leaves little room for excuses:

> *"In the past, some organizations may have relied on vague or even unwritten policy for the management of traditional collections. However, to ensure effective and efficient mechanisms for long-term preservation of and continuing access to its digital contents, a repository requires well-documented and widely adopted policies – and well-documented procedures...a policy for the preservation of digital files needs to sit comfortably within or alongside policies for non digital content"* (RLG-OCLC 2002, 28).

Museums (or any type of memory institution) with a serious desire to address digital preservation now or in the near future should aspire to develop a digital preservation policy as soon as possible. Not only do such policies provide a basis for digital archive requirements and a solid intellectual foundation for practical solutions, it also forms an important step in securing organizational buy-in to the principles and practice.

Although creating a full-formed policy is not always the first step an institution takes when launching a digital preservation system, it is a way for the staff, and hopefully eventually upper management, to organize the overall mission, goals, scope, staff roles, and basic procedures. This may help better define how the staff can tackle digital preservation, making it a less intimidating process and to also document its official initiation (Kenny and McGovern 2003). Regardless of whether policy is something made during the onset of a digital preservation system, or after more trust around the system has been established, policy is still a significant requirement when becoming a mature digital repository. This fact is highlighted by the *Five Organizational Stages of Digital Preservation* written by Anne R. Kenney and Nancy Y. McGovern and discussed below.

The authors of this paper propose that developing a comprehensive and effective digital preservation program does not necessarily have to do with duration (Kenney and McGovern 2003). In fact, to them, digital preservation policy is one of the last parts considered. The first steps concern the acquisition of digital materials, and acknowledging the ownership of said digital materials, which will naturally lead to discussion of maintaining those assets. Digital preservation policy is often the capstone rather than the cornerstone of such efforts (Kenney and McGovern 2003). Therefore, the stages (as a form of self-assessment) suggest benchmarks for measuring development, where policy falls into the more mature stages of readiness. The five stages of organizational response to digital preservation are:

1. **Acknowledge:** understanding that digital preservation is a local concern
2. **Act**: initiating digital preservation projects
3. **Consolidate**: segueing from projects to programs
4. **Institutionalize**: incorporating the larger environment
5. **Externalize**: embracing inter-institutional collaboration and dependency (Kenney and McGovern 2003).

Policy is considered within each of the stages. While policy may be implicit by stage 2, policy does not truly take form until stage 3, in which the organization "makes explicit its commitment to digital preservation by developing basic, essential policies and by understanding the value of policies as part of the solution"(Kenny and McGovern 2003). It is not until stage 4 that the authors envision the integration of creating a TRAC report or mapping to the OAIS Reference Model. The authors' premise for assessing digital preservation in these stages is that far too many institutions are not far along enough in these stages to build a sustainable digital preservation program. Technology is not the greatest inhibitor, but organizational readiness is. Therefore, the *Five Organizational Stages of Digital Preservation* can serve the purpose of outlining when policy should be created, the importance of policy for moving forward in institutional readiness, and the power of policy for formulating a mature digital preservation system.

The *Five Stages* is a useful metric to use for evaluating where an institution lies within the maturity process, especially when it comes to considerations of funding. The *Five Stages* realistically considers the challenges of securing funding, which seems to hit maturity only by stage 4; institutions can rest easy in knowing that until it has moved along the spectrum of readiness and towards the ideal mapping of TRAC and OAIS, funding does not necessarily need to be fully figured out. In fact, institutions may linger in stage 3 for some time until critical mass or funding builds and the organization feels pressure to move onto the next stage (Kenney and McGovern 2003). Interestingly, the authors correlate the creation of official policy in stage 4, with the securing of institutional funds as well as well as an institution's ability to move beyond rudimentary/basic digital preservation tactics. Funding is of course another area where policy can be vital.

In 2004, Tim Au Yeung wrote a paper commissioned by the Canadian Heritage Information Network called *Digital Preservation: Best Practices for Museums* (Yeung 2004). Although created over ten years ago, this paper is one of the only of its kind that considers the digital preservation needs of the museum community specifically. One of the most significant conclusions in the paper concluded that one of the biggest constraints of digital preservation to the museum field is funding. While there are many cost models, open-source, and

community-collaborative methods of digital preservation available for reference and consideration, funding is a huge topic onto itself and is beyond the scope of both Yeung's paper, and of this thesis. Although not able to extrapolate on how to best secure funding, Yeung concluded from a survey of literature that the most important recommendation he could offer to the museum field is to create a digital preservation policy. Reaffirming the conclusions noted earlier from the 2010 Planets survey, policy is a key step and document for securing institutional commitment, and therefore increasing the chances of achieving minimal levels of funding or staff members for a digital preservation project.

To conclude, while every institution will have its own constraints, policy, at whatever organizational stage, as either a formal or informal document, is important for creating an effective and legitimate digital preservation system. To reiterate a major point from throughout this thesis, museums are institutions that uphold the public trust as ethical stewards of cultural heritage, and maintaining digital assets for long-term viability will become increasingly important. In 2011, the Canadian Heritage Information Network conducted a survey on Canadian museums and digital preservation preparedness that found that 37% of respondents had experienced the loss of digital data (CHIN 2013b). Hopefully more museums will not fall into this category before realizing that digital preservation is a serious need in their institutions. Kenney and McGovern's *Five Stages* paper highlights that unfortunately more often than not, it takes the tragic loss of data in institutions before it is realized their mistake was not planning ahead for long-term management and preservation of their data (Kenney and McGovern 2003). Museums of course are best advised to not risk their own collections or assets getting to such a critical point. Following the words of Marie Malaro, the easiest way to prevent errors is to begin action through creating policy that directs the staff and the greater museum with standards and best practices (Malaro and DeAngelis 2012, 46).

Resources for Creating a Digital Preservation Policy

This section of the chapter will provide some guiding resources that offer best practices, frameworks, and checklists for creating a digital preservation policy. As of 2015, there are very

few U.S. museums who actually have a digital preservation policy, especially one that is published for public viewing. A study conducted by Madeline Sheldon of the NDIIPP of the Library of Congress in 2013 found that since 2008, libraries and archives consistently remain at the forefront of digital preservation policy and best practices, while museums have consistently remained a distant third place (Sheldon 2013). Sheldon found only two museums who had a published digital preservation policy, The National Museum of Australia and the Rhizome ArtBase for The New Museum (Sheldon, 2013, 7-9). While in 2015 it is still true that museums take a back seat in digital preservation policy, research conducted for this thesis has found that the Museum of Modern Art New York has a newly developed digital preservation policy (Fino-Radin 2015), as well as the San Diego Air and Space Museum (Renga and Riney 2012), the Computer History Museum (Kott and Jabloner 2012), the Tate Modern (Tate 2013), and the Smithsonian Institution (Smithsonian 2011; Smithsonian 2012). Many of these noted policies are presented as digital repository plans, or digitization, digital asset management, or digital initiative policy rather than distinguished as a 'digital preservation' policy. The above names represents a small fraction of the undoubtedly many more museums who are safekeeping digital materials, yet no more policies are to be found online. Unfortunately this also means that there are far too few existing museum digital preservation policies that other museums can use as a model to create future policies. There are many useful resources however that offer advice about formulating a digital preservation policy.

Most of the resources available were created with research universities and libraries in mind. The Canadian Heritage Information Network (CHIN) is the exception, and has provided digital preservation tools and tutorials specifically for museums. Although this resource comes from Canada, it is equally relevant to U.S. museums. As noted in the above sections, the Canadian Heritage Information Network conducted a survey in 2011 to identify digital preservation issues facing museums. In response to the survey, CHIN has released the *Digital Preservation Toolkit*, which is a suite of documents that outline concrete steps to identify digital material found in one's museum, the potential risk and impact of lost materials, and how to begin the development of preservation policies, plans, and procedures (CHIN 2013c). The Toolkit addresses the context of digital objects commonly found in museums including: administrative

materials (office records), records of a museum's physical holdings (collection management records), and resources that are born digital (digital video, photographs, sound recordings, etc). The Toolkit includes a Digital Preservation Inventory Template, Digital Preservation Decision Trees, best practices for creators and preservers (from InterPARES), Digital Preservation Plan Framework, and importantly a Digital Preservation Policy Framework Guideline (CHIN 2013c).

The Canadian Heritage Information Network recommends that museums first use their inventory tool to take stock of what media and files must be preserved, and then establish the first rendition of a policy using their guiding document (CHIN 2013c). With the goal to become a museum community standard, CHIN outlines a digital preservation policy framework that:

- *"Addressed the seven attributes of a Trusted Digital Repository*
- *Presents the high-level perspective of an organization's digital preservation program*
- *Reflects current not future capabilities of the digital preservation program*
- *Provides links to documents containing more detailed and frequently-updated documents, e.g. lower level policies and procedures*
- *Points to the digital preservation plan for near-term priorities and timeframes*
- *Documents the policy approval and maintenance process"* (McGovern 2013).

The suggested framework is divided into seven sections, one for each attribute of a Trusted Digital Repository: OAIS compliance, administrative responsibility, organizational viability, financial sustainability, technological and procedural accountability, system security, and procedural accountability. The policy framework created by CHIN for the museum context is an important resource for museums of any size and from any region. Much of the components featured in this document resemble the recommendations offered in additional policy-making resources outlined below.

Another useful model for a digital preservation policy can be found on the Inter-University Consortium for Political and Social Research's (ICPSR) website (McGovern 2007). The document created for the ICPSR was drafted by Nancy McGovern in 2007 and outlines the ideal components of a digital preservation policy. See the Works Cited section of this thesis for the website link to McGovern's policy outline (McGovern 2007). The Electronic Resource Preservation and Access Network (ERPA) created their Digital Preservation Policy Tool in 2003; although over ten years old, the usefulness of this tool is still relevant today

because the basic requirements and scope of a digital preservation policy has not changed much over time (ERPA 2003). ERPA's policy tool is continually referenced as a relevant resource in many of the digital preservation policy resources outlined in this thesis.

In 2008, the Joint Information Systems Consortium (JISC) and Charles Beargie produced a Digital Preservation Policy Study that focuses on the policies in major universities (Beagrie et al 2008b). While the scope of digital preservation in the university context is not as relatable to the museum context, this study is a useful document nonetheless for any person working to create a digital preservation policy for their institution because of the models, clauses, an implementation recommendations provided in the study and the final report (Beagrie et al 2008b; Beagrie et al 2008).

Drawn from COPTR (Community Owned digital Preservation Tool Registry), one can access the "Catalogue of Digital Preservation Policy Elements" maintained by SCAPE (SCAlable Preservation Environments)(SCAPE, 2014). This catalogue was created between 2011- 2014 as a final report after reviewing a number of digital preservation policies, it includes a policy framework, a policy template, and guidance for ten digital preservation policy elements (SCAPE 2014). In addition to the Catalogue, SCAPE's wiki also has a webpage that lists many Published Preservation Policies, which reflects the many real-life documents used to create the Catalogue (SCAPE 2015). This concise list of digital preservation policies is actively used as models for other digital preservation policies around the world. The only museum digital preservation policy listed on this site is that of the National Museum of Australia (SCAPE 2015). In addition, SCAPE cites the recent research conducted by the Library of Congress NDIIPP by Madeline Sheldon, also noted earlier in this text. Her informal report on the Library of Congress' blog, *The Signal*, is a brief recap of her research and functions as a very accessible resource, especially to those just starting to outline their digital preservation policy (Sheldon 2013). Although an informal online publishing platform, *The Signal* blog has many useful entries on policy including one written in 2011 by Bill LeFurgy called "Facing Off with Digital Preservation Policy" (LeFurgy 2011). This blog post evaluated 13 policies and created a taxonomy for pertinent sections found within policies. This informal study of various digital preservation policy provides a useful look at the common elements of a policy based upon a

crosswalk of 15 categories common between all the evaluated policies already used in the field (LeFurgy 2011).

Furthermore, the second phase of the InterPARES project (International Research on Permanent Authentic Records in Electronic Systems) produced many digital preservation guidelines including "A Framework of Principles for the Development of Policies, Strategies and Standards for the Long-Term Preservation of Digital Records" (Duranti et al 2008). This document is heavier on the theoretical principles rather than practical implementation. For a more practical guideline, the InterPARES also published a paper in accordance with the International Council on Archives in 2012 titled "Digital Records Pathways: Topics in Digital Preservation, Module 2 Developing Policies and Procedures for Digital Preservation" that provides a detailed template and workflow for policy development and review (ICA and InterPares 2012).

The Northeast Documentation Conservation Center (NEDCC) and the MetaArchive also provide publicly available templates for creating a digital preservation policy (NEDCC 2008; MetaArchive 2010) . In 2007, the NEDCC, in conjunction with PALINET, SOLINET, Amigos Library Services, and the OCLC Western Service Center, created a National Endowment for the Humanities two-day workshop that resulted in the creation of "Digital Stewardship Questionnaire" and a "Digital Preservation Policy Template" (NEDCC 2008). Their template leads users through linear steps for creating a basic skeleton preservation policy (NEDC 2008). Similarly, the MetaArchive created its policy template from a digital preservation planning workshop in 2010 (MetaArchive 2010).

Most, if not all, of these useful guides, templates, and frameworks for creating a digital preservation policy strive to comply with the OAIS Reference Model as the foundation for any kind of digital archive or preservation program. In addition many of them refer to the Trusted Digital Repository (TDR) guidelines created by the OCLC/RLG in 2008, and often the Trusted Digital Repository Audit and Checklist (TRAC). Overall, when creating either a skeleton, or a mature digital preservation policy there are a variety of tools available that are almost exclusively online, as free and open-source materials. The foundations established for

policy-making by these many resources will be useful references for any museum wishing to start or review their own digital preservation policy.

Conclusion

Although policy is important to museums, and research has demonstrated its role in launching a successful digital preservation program, few museums have digital preservation policies in place. The 2011 CHIN survey on museum preparedness for digital preservation asked respondents whether they had a digital preservation policy, strategy or plan, or guidelines at their institution. The response showed a large number of resounding "NO's" to this question (CHIN 2013b). The data's demonstration of a significant absence of digital preservation policy in museums is concerning, especially now that museums are faced with a growing responsibility to steward digital materials. The development of guiding policies has proven to be a necessary prerequisite for the implementation of active digital preservation programs that the museum field needs (CHIN 2013b). Notably no international or national standard has been established for museum-specific needs. The absence of a professional standard will lead museums to proceed with digital preservation with great caution and apprehension. Although good guidelines for non-museum specific policy are available, there is a lack of widespread adoption because of the lack of education and exposure to such frameworks and models.

Hopefully future research, education, and recommendations from the museum field, such as that made in this thesis, will help foster the practice of creating digital preservation policies, whether informal or formal. Such policies will facilitate the development of digital preservation activities within museums. Such implementation is possible, and will be exemplified through three case studies of museums enacting digital preservation initiatives at various levels of program maturity within the subsequent chapters.

Chapter 6: Methodology

In this chapter, the research methods used in this thesis will be outlined. First, a brief overview of the topic selection and overall research questions will be described. Second, the literature review selection and review process about key research in the field of digital preservation will be discussed. Finally, the selection process for the chosen three case studies will be described, as well as an outline of the interview questions used to conduct original research with the content experts of each case study.

Research Question

As outlined in the Introduction chapter, the question guiding this thesis is the following: how are U.S. museums handling the long-term accessibility and preservation of their many digital assets? Furthermore, are U.S. museums well-equipped to be prudent stewards of digital cultural heritage records, objects, and data by way of digital preservation plans and policies? These questions are significant because the museum field is dedicated to the ethical responsibility to care and share all aspects of cultural heritage under its stewardship, as well as a responsibility towards due diligence of managing collections records and research related to its holdings. In addition, as efforts to integrate into the digital age take place, many museums are investing in large digitization projects for their collections and archives, much of which are irreplaceable. How are museums managing its digital records for long-term sustainability? Because of the fast rate of digital software and hardware degradation and obsolescence, fear that many museums will come to lose vital data will become more and more of a reality the longer museums wait to implement digital preservation plans, strategies, and policies.

Museums must acknowledge their responsibility as stewards of digital material and as trusted cultural institutions, and then act within their best capacity to ensure the viability of valuable digital materials. A review of literature found little to be published on digital preservation from the museum field, with most information instead derived from the library and information science fields. The research foundations of this thesis was formulated on the themes of threats to digital assets; digital preservation definitions, history, standards, and practical strategies; convergence between libraries, archives, and museums; and digital preservation

policy. Three case study institutions with large digital collections and/or established digital preservation practices were selected based upon their prevalence within digital preservation literature, and current dialogue within digital preservation conferences. These three case study chapters highlight these institutions' exceptional expertise, and the emerging status and needs of digital preservation within the museum community. The process for selecting the literature review and case studies will be outlined in the following sections of this chapter.

An Overview of the Literature Review

The Literature Review of this thesis is a practical encapsulation of digital preservation concerns, strategy, and resources, and was designed to be a useful resource in of itself to any member of the museum community. In light of the fact that the topic of digital preservation is vast, the literature review is organized into several broad themes. The first chapter of the literature review focuses on the realistic threats to digital materials, including an overview of the qualities that specifically make digital materials difficult to maintain for long-term viability. The second chapter discusses how digital preservation is defined, the history of digital preservation, and the significant standards, strategies and workflows that make up an encapsulation of 'digital preservation 101.' The third chapter of the literature review provides an overview of the collaborative role of the museum within the digital age, and as a converging member of the collective genre of 'memory institutions.' The fourth chapter of the literature review establishes the importance of policy in the professionalism of museum work, as well as the specific importance of digital preservation policy in the promulgation of a successful digital preservation program. The literature included in these four chapters includes information from information science publications, library and research institute surveys and research, as well as professional standards such as those published by the American Alliance of Museums, the Society of American Archivists, the Library of Congress, the Canadian Heritage Information Network, UNESCO, Research Libraries Group, the Online Computer Library Center, and the International Council on Museums. Academic journals, articles, books, essays, blog posts, and key websites were also consulted.

The Case Study Selection Process

Case studies of three institutions managing large digital collections and utilizing digital preservation strategies were conducted in this thesis. The case study process included building of selection criteria, the selection of ten institutions, contact with institutions, content expert interviews, and analysis. The initial selection process for case study institutions emphasized the accessibility of a digital preservation policy; however, since digital preservation is currently an emerging topic in the museum field, it was found that this criteria was unrealistic. Instead, consideration was paid to institutions that manage large digital collections and have been mentioned in various literature to be pursuing digital preservation activities. Potential case study institutions were evaluated using a variety of sources such as: peer-reviewed journals, museum websites, blog entries, conference presentations, conference attendance, 990 tax forms, and institutional publications. The review of these sources (most of which are accessible on the Internet) took place between October, 2014 and March, 2015. This is of note because of the rate of change within the digital preservation field as well as the internet; webpages, websites, and content from many of the resources and museum websites in this thesis may or may not be represented in current or future iterations of said websites.

From a long list of potential institutions, further evaluation was narrowed down to a list of ten. This list was then evaluated according to operating budget size as a way to further categorize the choices into small, medium, and large institutions. The initial plan for case study selection was to choose an example of a museum digital preservation program from institutions with small, medium, and large budgets. However, further assessment elucidated that so few (none to be found online) small-budget museums were using digital preservation standards. Instead of using operating budget as a criterion, the list of ten case study options were categorized by the *maturity* of the organization's digital preservation program: from emerging to mature, which is similar to the maturity models crafted by Anne Kenney and Nancy McGovern, and which were outlined in Chapter 5.

Final selection of case studies was arrived at by taking into consideration the institutions that were most often discussed in current digital preservation dialogue. The final three case

study institutions coincidentally are all organizations that focus on art collections, although the focus on art was not intentional since this thesis strives to apply a holistic approach to digital preservation strategy to *all* kinds of digital material, even outside of art collections. It was also a coincidence that the three case studies are all major internationally recognized museums. Born-digital art, and major art digitization projects happen to be the current source of digital preservation discussion and needs when involving museums specifically. In addition, since digital preservation is still new within the museum field, it requires a fair amount of money and staff investment, which is why larger museums are tackling digital preservation issues first, setting the example for the rest of the field. The three case studies were therefore selected based on their knowledge of digital asset management, knowledge of digital preservation, and the work they are conducting to promote digital preservation activity within the greater field. Interviews with content experts were conducted as part of the case studies, and standardized contact scripts and interview questions were used for all three institutions.

Case Study Selection

The first step in selecting which institutions were a best fit for case studies was to conduct a brief internet survey of museums with digital preservation policies. This survey found that most museums with accessible digital preservation policies are outside the U.S. Because of this discovery, the internet survey was manifested into a spreadsheet of museums mentioned in library and digital preservation literature, in particular, the existence of blog posts of *recent* digital preservation highlighted relevant activity within those museums. The museums listed in this spreadsheet were then ranked, based upon their prevalence in the digital preservation literature used within the literature review of this thesis, as well as by the collection type (art, history, natural history, science, etc). Much of this data was derived from a 2014 survey made by the Museum Archives Section Standards and Best Practices Working Group of the Society of American Archivists, in which the topic of museum archives and electronic records was investigated among many museum and museum-libraries across the U.S. (SAA, 2014). Ultimately, although the initial intention of this thesis to provide examples of digital preservation from museums of a variety of sizes and budgets, it was determined that the institutions employing best practices and exemplary digital preservation systems was currently mostly within

major U.S. museums. As an emerging field, best practices in digital preservation for the museum context are still being developed and have yet to become widespread across the field.

In the end, three institutions were selected because of their importance to the work of digital preservation, availability/access to content experts, as well as a general assessment of the museum's overall efforts in managing digital preservation. The three museums selected were:

- The Metropolitan Museum of Art
- The San Francisco Museum of Modern Art (SFMOMA)
- The Museum of Modern Art in New York, NY (MOMA)

A set of interview questions was developed to examine the systems, strategies, and policies that guided each museum's digital preservation initiatives. A contact script was sent to each museum in April, 2015. Since the primary need for digital preservation in museums has been focused on collections management, especially that of time-based media, either digital asset managers, registrars, archivists, digital repository managers, information managers, or collection managers were considered for contact. Interviews were conducted in-person at all three chosen museums. The interview at the San Francisco Museum of Modern Art was conducted on May 6th, 2015 with Layna White, Head of Information and Access. The interview at the Metropolitan Museum of Art was conducted on May 18th, 2015 with Jenny Choi, Digital Asset Manager, along with email follow up with Dan Lipcan, Digital Initiatives and Metadata Librarian at the Thomas J. Watson Library of The Met. The interview at the Museum of Modern Art was conducted on May 19th, 2015 with Ben Fino-Radin, Digital Repository Manager.

The questions asked of each content expert were developed to examine the systems, processes, staff roles, preservation strategies, rationales, successes, and future plans for digital preservation within the museum. The same questions were asked at each institution with little variation to the order or wording. The SFMOMA and MOMA both requested the questions in advance to better prepare their answers, and the questions were sent by email a few days before the interview. The questions reflected various themes: digital preservation technology systems, digital preservation standards, administrative planning and policy, and history of the implementation of digital preservation systems.

Interview Questions

The themes presented within the interview questions were intended to ascertain the practical implementation of digital preservation within the museum context, as well as unearth the standards factored into the planning and policy of said digital preservation efforts. The questions were intended to be neutral in terms of the type of collection being preserved. In addition, the questions were inspired by the recent survey conducted by the Museum Archives Section Standards and Best Practices Working Group of the Society of American Archivists (SAA 2014). Many of the questions asked within this survey highlighted key themes in digital preservation planning. In light of the goal of this thesis to address a holistic approach to digital preservation (including both collections and records/archives), the questions asked in this survey addressed general standards, digital asset management, technology systems, planning, policy, and future plans.

The first three questions of the interview asked about how digital preservation came to be a raised issue in the museum, and then as a working project within the museum, including questions about how many staff members worked on the project and what resources were used to launch the digital preservation efforts. The next five questions of the interview asked about the more technical aspects of a digital preservation system, including inquiries about OAIS compliance, ISO 16363 (TRAC), digital preservation software, archival file formats used, selection process for preservation, metadata schemas, data migration practices, normalization of formats, ingest, digital storage, and backup systems.

The next four questions were concerned with inter-museum collaboration such as the relationship of a Digital Asset Management System to digital preservation, communication amongst the greater museum staff regarding the acquisition and maintenance of digital preservation, the scope of additional electronic records saved (email, website, blogs, social media, etc.), and access to preserved digital materials or records. The last three questions of the interview were more disparate in theme, the first being a question about how the museum funds its digital preservation efforts, the second question about whether the museum has a digital preservation policy or statement, and the third question about the future plans of the museum regarding digital preservation.

The fifteen questions asked in each interview were intended to gather a general picture of the digital preservation and management practices within each institution. The actual questions can be found in the appendices of this thesis. Furthermore, the case study interviews were designed to elicit data regarding the directions museums are moving towards in regards to digital archives, digital stewardship, and ultimately future digital preservation. The answers provided in each interview reveal the underlying planning, needs, and philosophical underpinnings of each institution that paints both an interesting and galvanizing picture for the future of digital preservation for the museum field.

Each case study is individually presented in the following three chapters of this thesis. The case study chapters will discuss: a brief background of the museum, its relationship to digital technology, the museum's current digital preservation practices and future plans, the museum's story about how they got to where they are today, and an analysis of the practices and rationale of the museum compared to the literature, standards, and strategies discussed in the literature review.

Chapter 7: The Metropolitan Museum of Art

Introduction to The Met

Founded in 1870, the Metropolitan Museum of Art (MMA or "the Met"), located in New York City is one of the oldest American museums. By the 20th Century, the Met had become one of the world's greatest art centers, boasting an encyclopedic collection (Met 2015). Today their permanent collection includes more than 2 million works of art, spanning 5,000 years of world history and culture. These vast holdings are managed between 19 curatorial departments, each one responsible for a comprehensive and specialized genre, whether American art, European art, Ancient Egyptian, Islamic art, Asian art, photography, costume, or decorative arts (Artstor 2015).

From the museum's original charter, it was charged with the purpose of "establishing and maintaining in [New York] a Museum and library of art, of encouraging and developing the study of fine arts, and the application of arts to manufacture and practical life, of advancing the general knowledge of kindred subjects, and, to that end, of furnishing popular instruction"(Met 2015b). This very statement of purpose had guided the Met for over 140 years. On January 13, 2015, the Trustees of The Met continued the spirit of the original charter and supplemented it with the following mission statement: "The Metropolitan Museum of Art collects, studies, conserves, and presents significant works of art across all times and cultures in order to connect people to creativity, knowledge, and ideas" (Met 2015b).

The Met's museum library and archive were both authorized by the original 1870 charter and formally established in 1880 (Fleming and Lipcan 2012). In 1965, the museum library was moved into its current building next to the museum today and renamed the Thomas J. Watson Library, the founder of IBM and a Museum trustee. Today it is one of the world's great collections of art historical research materials with over 900,000 volumes (Thomas J. Watson Library 2015). The Thomas J. Watson Library is the center for research and archives relating to The Met's art collections, and its mission is to support the research activities of the Museum staff as well as serve the international community of scholars. The foundation of the Lita

Annenberg Hazen and Joseph H. Hazen Center for Electronic Resources in 1997 positioned the library as a leader in collecting and managing online resources (Met 2015c).

Since its creation, the objective of the Metropolitan Museum of Art's Archive was to collect, organize, and preserve in perpetuity the corporate records and official correspondence of the Museum (Met 2015d). Until the 1960s, the Archives primarily served as a resource for the Museum's secretary, officers, and trustees but has since expanded its collection scope to serve the needs of the whole Museum and of the general public. Today the Museum Archive holdings include Board of Trustees records, legal documents, Museum publications, office files of selected Museum staff, architectural drawings, press clippings, and Museum-related ephemera (Met 2015d).

Relationship to Digital Technology

With a trifecta of important departments including the museum's collections, library, and archive, the Metropolitan Museum of Art is committed to the stewardship of a wide range of materials. This stewardship has expanded to include a multitude of digital assets in addition to its traditional physical holdings. In our current 'information society' that depends on technology, the Met's distinctive mission to provide experience and knowledge to the public has directed the museum towards innovative digital initiatives. As for any museum in the 21st Century, the internet provides the new frontier for public engagement, sharing of collections, and access to knowledge. As stated by Thomas Campbell, the Executive Director of the Metropolitan Museum of Art, to The New York Times in 2014: "Impacting all of us is technology. We've made a huge investment in transitioning from being an analog museum to a digital museum and there are great opportunities in that to see the collections on the whole, to deliver the information to our audiences in new ways" (Pogrebin 2014).

The Met has embraced its website to become an extension of the museum itself. The Met has over 1,000,000 digitized works of art, over 400,000 of which are online (Choi 2015). The ultimate goal, as mandated by the Museum Director, is to try to put as many photographs of

the collection as possible onto the Web, and to provide encyclopedic access to its holdings - achieving a new level of museum transparency, and public access unheard of before the paradigm shift of Web 2.0 (Choi 2015). In addition, the Met contributes to the digital image library Artstor with over 9,000 images represented, 7,800 of which are available as high-resolution downloads for academic publishing (Artstor 2015).

One of The Met's significant online collection resources is The Collection Online, a comprehensive image catalog of over 400,000 artworks, searchable to users by artist, genre, date, location, culture, or by curatorial department (Met 2015e). The data available on this feature includes curatorial research, exhibition history, provenance, publications related to the work, and more. All of these valuable resources are available for free to the user, including the publications most of which have been digitized and are downloadable directly from the web (Met 2015e). In addition The Met participates in various social media outlets, such as its group pool on Flickr in which visitors can post their own photos taken while touring the galleries. The Met's Flickr group has grown to over 2,400 members and over 22,000 photos (Wall 2015).

Other initiatives include One Met. Many Worlds, an online interactive interface provided in 11 different languages; Viewpoints: Body Language, an online learning tool that includes audio and video of experts discussing how body language is communicated through art; Connections, recordings of The Met's curators speaking about their personal connections to art that is both personal and academic; and the Heilbrunn Timeline of Art History, which in an invaluable research and visual tool that presents the Met's collection via a chronological, geographical, and thematic exploration including 300 timelines, 930 essays, and close to 7,000 objects culminating in a robust index of global art history (Met 2015f). The robust number of digital resources available on The Met's website indicates an enormous digitizing initiative, whether of artworks themselves, or of academic resources such as essays, catalog entries, publications, etc. In addition, The Met's website features a number of interactive web applications that involve both audio and video.

The Museum's digital holdings also extends to the library and archive. The Thomas J. Watson Library has its own digitization initiative with the primary mission to "expand access to the Library's rare and unique materials by developing, supporting, and promoting a distinctive

digital collection of these items" (Met Library, digitization initiative). Part of the goal of the library's digitization project is to preserve many of the original printed materials that are rapidly deteriorating from heavy use and acidic paper. As of 2012, the Thomas J. Watson Library has digitized more than 3,000 items both independently and in collaboration with the Museum's curatorial departments, as well as other art museum libraries and galleries (Fleming and Lipcan 2012). This digitization initiative is also extended to the Museum Archives to identify and include additional Museum publications not held by the Thomas J. Watson library. With preservation of valuable information in mind, the ultimate goal of the Library's digitization is to compile the "digital library record for early Metropolitan Museum of Art publications" and resources (Fleming and Lipcan 2012).

The Met is an important case study for digital preservation because of the institution's commitment to using digital technology as a tool for public access, preservation, and long-term sustainability as a memory institution. The Metropolitan Museum of Art's relationship to digital technology is as fast growing and vast as its encyclopedic collections. The Museum has invested in digital technology to position itself as a pioneer and significant contribution to our socio-cultural record. The sheer vastness of the Met's museum, library, and archive collections requires an immense amount of money and staff effort to digitize and provide access to its materials.

In addition, The Met uses digital technology to aid in the preservation of analog assets and records. While The Met has a clear directive about using technology as a tool for preservation of analog collections, much less information is available concerning the actual preservation of the existing digital objects themselves. The next question, and the next step relating to The Met's relationship to digital technology, is to inquire how these vast holdings of digital images, audio, video, social media, publications, and web applications are to be managed and cared for so that they are viable resources for future scholarship and public engagement. This topic will be discussed in the subsequent sections of this chapter. Most of the data discussed in the below section were derived from interviews with Jenny Choi, Digital Asset Manager, and Dan Lipcan, Metadata and Digital Initiatives Librarian (Choi 2015; Lipcan 2015).

Additional information was derived from an recent interview conducted by the Society of American Archivists with Jim Moske, Archivist for the Met (Bowling 2014).

Status of Digital Preservation at The Met

Many institutions will find different pathways towards digital preservation, whether that is by lobbying for a full contract with an OAIS-compliant system, or building ownership of digital assets over time and employing basic preservation tactics until something more robust is needed. The Metropolitan Museum of Art follows the latter pathway. This section will outline how The Metropolitan Museum of Art's approach to workflows and systems that bolsters its current stewardship of digital materials.

The Metropolitan Museum of Art has a large staff of over 2,000 employees. The staff is divided by the three sectors, Museum, Library, and Archive, with a variety of teams within each. Although currently no staff members distinctly work on digital preservation, designated staff manage digital objects for long-term use. The Digital Asset Management Team at the Met consists of 6 staff members. There are also 8 photographers and 4 catalogers who contribute to metadata control, technical standards, display standards, and quality control (Choi 2015). At the Thomas J. Watson Library, there are 2 staff members who help manage their (Lipcan 2015). As of late 2014, the museum Archive has 3.5 permanent staff members and 3 temporary/grant-funded staff members. Within all these teams, there is currently no one titled as a 'digital archivist' (Bowling 2014). However, the Information and Technology department supports the whole museum in the installation and maintenance of database software, collection management software, and content management software (Bowling 2014). These technology systems will be outlined below.

Technology Systems

Access to collections, and therefore, access to digital assets, continues to be one of the Met's primary concerns. According to the library staff, preservation has always been an active concern on the mind of the museum's units, although it has not yet acted upon it (Lipcan 2015). Regarding e-records, there is no institution-wide electronic records program, although batches of

born-digital material from selected departments and sources around the Museum are collected by the Archive (Bowling 2014). For the greater museum, the Met's digital preservation efforts are currently embedded in their IT and Digital Asset Management teams. Implementing a Digital Asset Management System (DAMS) for the museum, library, and archive is still a newer initiative and has been a work in progress since 2007 (Choi 2015). The decision to invest in DAM software was a much more natural progression from the Museum's digitization efforts than jumping into thinking specifically about digital preservation.

Digital Asset Management software enables management tasks and decisions surrounding the ingestion, annotation, cataloguing, storage, retrieval and distribution of digital assets across many platforms, web interfaces, and for different user-types. Digital asset management refers to the protocol for downloading, naming, backing up, rating, grouping, archiving, optimizing, maintaining, thinning, and exporting files. The wide variety of management achieved with this type of software enables the storage and retrieval of assets that the Met immediately needs to carry on its digital initiatives. A DAMS is not technically a digital preservation system because it does not necessarily run format-checks or checksums, nor maintain the high-level requirements for ingest (SIP) or storage (AIP) modeled by the Open Archival Information System (OAIS) using preservation-specific metadata schemas. The focus on a DAMS is typically around user interface and storage for *retrieval*, rather than for long-term repository. Nonetheless, the use of a robust digital asset management system can be a starting point for thinking about the long-term stewardship of digital materials.

By 2007, the world's mode and expectation for receiving information was changing, as so much data was going digital (Choi 2015). To stay relevant to the needs of its public, the Met had an fast-growing influx of digital images that could have been at risk of being hard to retrieve, lost, misused, or corrupted without a way to manage and maintain them in a server (Choi 2015). The Met first recognized the need to preserve its growing digital assets once it began to create a massive amount of digitized museum collections mostly in the form of photography, which was used by a multitude of departments such as curatorial and marketing for creating exhibit materials, catalogs, flyers, online materials, etc. (Choi 2015). New online initiatives to make the collections more publically available also initiated more digitization of the Met's encyclopedic collections.

The initial digital asset management software implemented at the Met was Media Bin, which was the same software used by the San Francisco Museum of Modern Art at the time (Choi 2015). Media Bin was able to be linked to the Museum's collection catalog, TMS (The Museum System) which is another important relational database that managed additional digital assets, such as exhibit media (Choi 2015). The Met has been using TMS for over 20 years; during this time, the Museum's integration of this software, together with the DAMS software, allowed the Museum to control access, metadata, technical standards, display standards, copyright, and other quality control. Although the Met does not use any digital preservation-specific software, the above listed functions achieved between their collection catalog and Media Bin enacted a basic level of preservation that enabled access and long-term tracking of digital media.

Recently, the Met's increased need to manage video and audio files has proven to be problematic through Media Bin, and it was determined that a new DAMS would be needed. After publishing a Request For Proposal, the Met chose to invest in NetExposure, in Spring, 2015. This software is the same digital asset management system currently used at SFMOMA and the Museum of Modern Art, both of which also decided to cease using Media Bin because of its inability to handle rich media files (Choi 2015). At the time of the case study interviews in May, 2015, the Met had not yet migrated to using NetX. The Met hoped that this new software system would give them some of the same functionality as The Museum of Modern Art, who had been using NETX since 2005. MOMA has used NETX to streamline content between its legacy collection management system, TMS, to the museum's website, as well as to staff (NetX 2015). This new software system was chosen by all these museums because of its ability to manage more complicated digital assets such as video files. Ultimately, using a DAMS allows the Met to maintain an organized digital record of its collections, along with all the data associated with them (photos, video, research, paperwork), supporting the Met's standing as one of the world's most reputable education and research centers for global art and culture.

The library and archive units of the Met use separate technology systems from the Digital Asset Management team. The Museum Library uses a collection management system

called CONTENTdm, which is produced and managed by the Online Computer Library Center (OCLC) (Bowling 2014; OCLC 2015). CONTENTdm enables libraries to store, manage, and deliver more content to the Web (OCLC, 2015). Some of the features of this collection management software include abilities to customize a digital collections website, the ability to upload metadata of collections to WorldCat (an internationally used database of library materials), and the ability to store any kind of document, image, video, or audio files (OCLC 2015). Similar to NETX used by the Museum staff, CONTENTdm has the end goal of maximizing end-user discovery, access, and display of materials rather than long-term storage, or standards of a trusted digital repository.

The Met's Archive unit uses its own technology tool called Archivist's Toolkit (Bowling 2014). Archivist's Toolkit is an open source archival data management system that supports the archival processing and production of access abilities, promotes data standardization, and efficiency (Archivists' Toolkit 2009). The Archive will also occasionally store digitized materials in the Thomas J. Watson Library's CONTENTdm system including some digital surrogates of audio/visual materials (Bowling 2014). Although only the unit at the Met that uses Archivist's Toolkit is the Archive, the staff is working towards more cross-departmental records management and sharing. For example, PDF preservation using MediaBin or NETX has been an ongoing discussion between the Digital Media and Archive teams (Bowling 2014). There is currently no active network connection between the DAMS (NETX), Archivist's Toolkit, or CONTENTdm, although the respective departments that manage these systems will occasionally deposit material amongst each other (Lipcan 2015; Bowling 2014).

Many of the digital assets from the Museum and Library were contributed to the major web-archiving nonprofit, The Internet Archive. The Met has contributed over 140,000 of its digital images to the Internet Archive with over 5,000,000 visits by the public as of October 2015 (Internet Archive 2014). In addition, there are over 2,800 digitized texts from the Thomas J. Watson Library contributed to the Internet Archive (Internet Archive 2015). By contributing these materials, the Met participates in the Internet Archive's mission to provide permanent access to cultural and historical collections that exist on the internet and in digital format, thus

participating in one of the largest digital preservation/web-archiving collaborations and projects (Internet Archive 2015b). Allowing the Internet Archive to host many of the scanned publications and images from the Met's digitization efforts means that the Internet Archive will sustain the collections in perpetuity for the Museum.

For internal storage, most if not all of the Met's digital materials, regardless of whether they came from the museum, library, or archive units, are stored across an internal network server (Bowling 2014). This network server functions as the museum's storage for material ingested into the various digital asset management softwares utilized within the different units. The network server is routinely backed up by the IT department (Bowling 2014). The use of digital asset management software across different units of the Met exemplifies the museum's approach to managing digital assets with a stronger focus on retrieval and access, and less on long-term viability or active preservation activities (digital curation), although the shared server does link the whole museum for access and storage needs. The level of preservation activity that is being used at the Met, while not purposefully following industry standards like OAIS and TRAC, works for the institution's needs right now. These activities will be outlined below.

Preservation Protocols

The Metropolitan Museum of Art has a series of protocols involving the creation, management, and storage of digital surrogates, digital files, and other digital materials that may be ingested into the digital asset management system and internal server (Choi 2015). Having such protocols enables better retrieval and access, but also bit-level preservation.

Digitization as preservation was the first phase for the Met's major digital initiatives. The *selection* process for the creation of digital surrogates was mainly based on the sensitivity of storage media or the fragility of objects. For example, slides were one of the first collections to be digitized because of the fast deterioration rate of slide media (Choi 2015). Much motivation for creating digital surrogates of their collection also came from internal and external requests for high-quality images, such as from curators, publishers, or researchers. Since Media Bin had been linked to the collection catalogue database (TMS), digitization was also prioritized for analog images or paper files relating to collections which could then uploaded to bolster

corresponding catalog records (Choi 2015). Outside of the Digital Media department and within the Met's Library, selection criteria is typically based on curator's recommendations, rarity/uniqueness, condition, research value, and intellectual property concerns.

According to the Digital Asset Management team, a distinctive *selection criteria* for digitization and deposit into the server does not perhaps truly exist since any and all collections are desired to be digitized, made accessible online, and then stored for accessibility. *All* collections that are digitized are mandated by the Museum Director to be put on the Met's website. The Director's mandate for online access greatly explains why implementation of digital asset management software has been a higher priority than establishing a standardized long-term repository (Choi 2015). With a need to quickly keep up with the progression of public interaction online, the Met's most recent priorities have been around control and access to digital surrogates.

While digitization has been on the minds of staff at The Met, they have not acted upon it much yet (Lipcan 2015). Regardless, there are ways in which the institution effectively controls the viability of its digital materials, such as *format standardization.* All digitized slides, transparencies, and photos are maintained in TIFF format, which is the standard recommended by the Library of Congress for long-term archiving (Choi 2015). For any new digital photographs of museum objects, the raw file is maintained as a master copy, and another copy is migrated to TIFF format to be manipulated for use copies (Choi 2015). These use copies are controlled by the DAMS, which allows for the download of images to JPEG formats. Furthermore, file formats are also controlled for publications and records, which are all maintained in PDF format, another standard recommended by the Library of Congress (Choi 2015).

The Museum Archive's workflow for the submission of files is a good example of their use of format standardization and normalization. The Museum Archive typically receives museum records to be deposited into the Archive as an email attachment; which is downloaded from institutional intranet and saved onto the server as a PDF, WAV, or TIFF format. If any materials are submitted on a Compact Disc, those materials are also normalized to PDF files stored on the internal server (Bowling 2014). Although the Museum-side has put film, VHS, and

audio recordings on the backend of priorities for now, the Library and Archive does consistently maintain audio recordings in WAV format for preservation copies (Lipcan 2015; Choi 2015).

In addition to controlling what digital file formats are maintained for all items deposited to digital asset management software and into the server, the Met's team uses a standard *backup routine and magnetic tape storage* to aid in basic bit-level digital preservation. To connect the various departments and sections of The Met, all digital assets are ultimately stored on an internal network server after being ingested. The data on the internal network storage is automatically backed up nightly onto magnetic tape storage (Choi 2015). Magnetic tape storage is an ideal backup and storage media because it allows for lossless data compression, is less expensive compared to disk and cloud-storage, and is reliable for retrieval. The tapes are maintained for two months before being recycled and rotated for new back ups. For emergency and disaster preparedness the tapes of the backup copies are stored off-site (Choi 2015). The shared network server is accessible to any museum staff that has been granted access rights, but the backup copies are only accessible by the IT department for security reasons (Choi 2015; Lipcan 2015). Additional backup systems are maintained for individual staff email accounts on a cloud-server and the museum's website linked to the network share, and so is also backed up routinely by the IT department (Choi 2015; Lipcan 2015).

This routine back-up system and use of magnetic tape storage is an effective way to achieve long-term bit-level preservation of digital assets. Because a backup copy is an exact replica of the original, the museum can preserve the bitstreams for their image, audio, and video files for eternity as long as they continue their routine protocols. While this strategy is sufficient for bit-level preservation, it only remains effective assuming that future software and hardware will run the chosen preservation formats used by the museum. Format-level preservation is the next level of digital preservation not fully realized within their backup and storage protocols.

Another way that the Met controls the viability of its digital assets is through implementing *consistent metadata*. Digital asset management systems can aid in controlling metadata entry for all kinds of deposited assets. To enable this ability, the Digital Asset Management team at the Met must set the boundaries, requirements, and fields for metadata for the various types of files and the various needs across the Museum's departments. The broad

applicability of metadata can be challenging because some content creators will be able to fill-in certain metadata fields more than others, resulting in incomplete metadata in some cases. The Digital Asset Management team at the Met works with a variety of departments to teach them how to use the DAMS, and the metadata requirements for submission and cataloging (Choi 2015). Sometimes metadata entry by more than one person may be required to complete a catalog entry. If the metadata recorded between the collection catalog (TMS) and the DAMS is consistent, data and access functions can crosswalk more easily between them. In addition, consistent metadata allows for more control and consistency with information uploaded to the museum's website (Choi 2015).

Furthermore, consistent metadata more easily facilitates data migration that occurs as their software may be updated over time; consistent metadata allows for the crosswalk of data fields from one software to another to be far more lossless. For example, as the Met transitions from Media Bin to NETX in 2015, the Digital Asset Management Team has created metadata protocols and a migration plan to make sure that the data from one system transfers to the next without any data corruption along the way (Choi 2015). As a hypothetical example, if Media Bin has a metadata field for Creator, but NETX is designed to use an analogous field called Artist Name, when transferring the data from one to the next, the systems will not know to save the data recorded under Creator and place it into Artist Name unless a migration protocol is written, or the fields are changed to be the same. Nor would the two systems know that entries that read "V. Van Gogh" is the same person as "Vincent Van Gogh," resulting in further migration issues. Making sure that the way metadata is entered into a database or DAMS is consistent is extremely important for long term retrieval and management.

The staff who works with TMS and NETX have developed their own cataloging standards that are tailored to their needs; they does not use any specific metadata schema standard, but borrows various elements from many schemas including the controlled vocabularies published by the Getty (Choi 2015). The Museum Library uses Qualified Dublin Core for its digital collections, which is a standard widely used across libraries worldwide (Lipcan 2015). Being proactive in maintaining consistent metadata that can be shared and made intelligible to many users is very important to the Met's practice in safekeeping digital materials.

Institutional Management and Future Plans

The Metropolitan Museum of Art's digital preservation plans are still emerging as their digital asset needs continue to grow and evolve. One such future plan for the Met is to create a digital asset management policy, which will include topics around the digital preservation and stewardship of its digital collections (Choi 2015). Creating an institutional policy from the Museum's Board of Trustees is a long and involved process because such policies not only need to be created from the governing body, but institutional policy is a sensitive document that reflects the ethical handling of its assets, and therefore, its creation must be carefully thought out. The Met otherwise has a variety of departmental procedures; the creation of digital asset management procedures are already in place for the DAMS, and further documentation is in place for the migration of their old DAM (Media Bin) to their new DAM (NETX) (Choi 2015). The Museum's Library also relies on the Digital Media team to implement preservation policies for content in the DAMS, so this future responsibility will likely be led by that those staff persons (Lipcan 2015).

The Digital Media team in charge of the DAMS is lucky enough to be supported as part of the general operations budget of the museum. The Met's full dedication to open access to its collections inherently led to the necessary support (aka funding) for the management of all digital surrogates. Policy has therefore not been needed in order to help create a business argument for financial support; instead, a future policy for the Met will function as a necessary document to help guide the projects and roles of the staff involved, and perhaps to exemplify to the public the Museum's digital stewardship practices.

Other future plans will involve time-based artworks. The Metropolitan Museum of Art has plans to continue collecting contemporary art in addition to traditional genres (Choi 2015). Since digital art is on the horizon for the Met, collaborations with other museums, notably MOMA New York, as a resource for how to care and preserve born-digital artwork, have only just begun. This genre of art collection is uncharted territory for the Met, however the growth of such art collections could be a possible major motivation for implementing a true digital preservation system one day (Choi 2015).

Social media outlets are another area that the Met is beginning to prioritize for archiving. While social media is considered an element of popular culture, the posts and thoughts recorded on these forums reflect a relevant log of how the public interacts with the Museum, as well as a record of current events and updates from the Museum. As of mid-2015, the Met was beginning to archive photos made by Met Museum staff on their Instagram account and has a running spreadsheet of the museum's Twitter "tweets" (Choi 2015). A record of the Museum's Facebook is currently not archived in any way; however, the museum staff are starting to recognize the importance of doing so in the future, because there is content posted there that may be important for the Museum's institutional history (Choi 2015).

The Met's Library staff has had digital preservation within their future goals for some time; one goal in the immediate future is to move their archival files into the Museum DAMS so that they will be more effectively managed and preserved on an institutional level than just remaining on an internal network shared server (Lipcan 2015). One of the Library staff members has worked directly with the Digital Media department to submit some archival scans to the DAMS repository (Lipcan 2015). This cross-departmental process has only been done on a pilot level for a small number of assets because of the heavy workload on the Digital Media team. For now, the library's digital assets are typically a lower priority, unless they pertain to high priority projects, such as exhibitions (Lipcan 2015).

In conclusion, the Metropolitan Museum of Art's future plans for digital preservation are to one day increase institutional management through policy-creation and integration of library digital surrogates with the Museum's active DAMS workflow. Time-based media and born-digital artwork is also on the horizon for future digital preservation needs and concerns at the Met. Social media continues to be a growing asset that the Museum wishes to selectively archive one day. On an ongoing basis, the Met will continue to increase its holding of digital assets and it will continue to make those collections accessible to the public as an enrichment to the internet as a new medium for accessing our cultural record.

Analysis

This analysis will discuss the digital preservation efforts at The Metropolitan Museum of Art in the context of the best practices and topics presented in the literature review. Taking into consideration the Met's major digital initiatives that contribute to the art historical research community, as well as provide public access to information, the Met is a true pursuant of the digital age and of the 'information society.' Even as a large, internationally recognized museum and research center, the Met continues to find its way for digital preservation needs, much like most museums nation-wide. The key themes to be discussed about the Met's digital preservation case study are: defining digital preservation; distinctive practices between the library, archive and museum units; and bit-level preservation.

Defining Digital Preservation as a Common Challenge

Defining digital preservation is challenging in the museum field because of the many misconceptions about the very term 'digital preservation'. One of the most common misconceptions is that 'digital preservation' is meant to mean digitization. While The Metropolitan Museum of Art itself has not succumbed to this misconception, the Met still serves as an excellent example of a museum whose relationship to digital preservation started with digitization, followed by access, leading them to their current status as a major steward of digital assets with emerging digital preservation practices underway.

The major focus around 'going digital' for the general museum field has revolved around public engagement, internet tools, social media, and of course the digitization of museum collections to create surrogate cultural records. It is worth restating that a high-resolution photograph or digital scan captures a frozen snapshot of a museum object, and therefore preserves a visual account of that object. However, to believe that these digitized surrogates are in of themselves forever stable, is a false concept. Digitization is not necessarily preservation (Rinehart, Prud'homme, and Huot 2014, 29). Hence the very foundation to this thesis is born: digital assets also need their own level of care, maintenance, and active curation in order to withstand the fast-evolving technological world.

Digitization as *one* form of preservation is certainly a valid concept; however, cultural heritage institutions need to be careful not to believe that digitization is the end point for digital

preservation. The Met has made strides in avoiding this misconception by making sure that the many digital assets it creates are managed professionally through submission criteria, specific metadata standards, and robust digital asset management software. These factors, as outlined earlier in this chapter, ensure the ongoing access that the Museum needs and would be commended by digital preservation professionals, such as the Library of Congress National Digital Stewardship Alliance. The Met does not misconstrue the need for digital preservation, however the institution has not yet fully developed its goals or plans for digital preservation on the level of long-term stewardship, such as that required of a Trusted Digital Repository (ISO 16363). It can be speculated that this is largely due to some misconception over the concept of digital preservation within the Met, and may also be due to more energy being put on the digitization and access of collections, rather than of building a trustworthy digital repository.

Distinctive Practices Between Library, Archive, and Museum Units:

Another interesting point of analysis is the presence of three separate systems contributing to the Metropolitan Museum of Art's digital resources: its museum, library, and archive. Each unit has its own DAM software, creates its own digital assets, and manages them with particular standards separately. Currently the three units collaborate only on a pilot level, however, increased continuity between the three units to streamline repository processes is on the horizon. The current separation between the three units, but the desire to diminish its silos, is indicative of the emerging collaboration opportunities at the Museum. As discussed in Chapter 4, much like how libraries, archives, and museums have traditionally operated separately, the Met's own internal units followed a similar pattern. However, the boundaries between the units are becoming blurred because each unit's digital initiatives are starting to look and feel similar, ultimately achieving the same goals and using the same kind of materials: *digital* resources. The Metropolitan Museum of Art encapsulates its very own insulated cultural memory community, and as such, it follows the trend towards library, archive, and museum collaboration, such as that supported by the International Council on Museums (ICOM) and UNESCO.

Bit-Level Digital Preservation

As an institution with emerging digital preservation plans, The Metropolitan Museum of Art has matured over time to understand that managing digital assets is a necessary

responsibility that goes hand in hand with creating and stewarding a large number of digital assets. This perspective of ownership of digital assets is vastly important towards the realization of digital preservation, as outlined by the requirements of a Trusted Digital Repository by the Research Libraries Group (RLG) and the Online Computer Library Center (OCLC). As such, the level of digital preservation achieved at the Met is mostly at the *bit-level.* The Met's digital asset management system software (DAMS) and backup/storage system for its digital assets maintains an exact copy of the digital bits, and will do so for eternity as long as the Museum follows its same backup protocols. Bit-level preservation most plainly ensures continued access. However, another common misconception of digital preservation is to think that if an item is accessible, then it is fully preserved (Rinehart, Prud'homme, and Huot 2014, 29-30).

Access is in fact not always required for an object to be considered digitally preserved; access is more of a desired component or outcome. As a point of contrast, full digital preservation includes bit-level preservation, *as well as* services intended to ensure that the information content of the files will remain usable into the indefinite future, as defined by the Library of Congress in Chapter 2. Digital preservation has much more to do with the long-term storage requirements and periodical refreshing or migration of the bitstream, and less to do with access. The access provided by the Met's DAMS may be a higher priority for now, and indeed provide more immediate satisfaction than true digital preservation. However, without additional preservation strategies within their current system, access to the valued born-digital assets created by the Met will not be reliable over time. The DAMS, backup, and storage strategies at the Met are effective for their needs currently; however, the risk exists that if the institution continues to sidestep the issue of longer-term preservation in favor of providing access to materials, that at some point this approach may eventually fail, leaving the museum with a preservation crisis.

While the Metropolitan Museum of Art has considered preservation issues, the institution simply has not yet matured their digital initiatives to include that step. This situation is perhaps the most relatable to most large, medium, or small sized museums in the United States that struggle with their role as memory institutions in the digital world, and which have aggressive access mandates from upper management on their hands. In light of this fact, the Met

as a case study for an emerging digital preservation initiative serves as an inspiring and practical learning opportunity for other museums to see how even large institutions struggle with defining and implementing digital preservation.

Within the context of Nancy McGovern's *Five Organizational Stages of Digital Preservation*, the Metropolitan Museum of Art would be considered to fall within Stage Two or Stage Three, where the institution is accumulating commitment to digital preservation by setting up technical requirements that apply to each digital project on its hands, but digital content is still dispersed across multiple locations. As specified for Stage Three, the Met, at a minimum, practices some assessment of the basic technology investment, and focuses on creating safe spaces for its digital resources (Kenney and McGovern 2003). As also outlined in the *Five Stages*, a digital preservation policy is not expected to be formed until Stage Four, and as an institution defined within Stage Two or Three, the Met has yet to form its own digital preservation policy. Evaluating the Met within the context of the misconceptions of digital preservation, as well as within the *Five Stages* provides context for the expected qualities of an emerging digital preservation initiative.

As a case study, The Metropolitan Museum of Art's practices highlight what a museum can do to safeguard digital resources, even without a mature digital preservation system. Bit-level preservation, while not necessarily fulfilling all necessary standards, is nonetheless an effective process. Digitization products (images, video, etc,) must meet current standards and guidelines in order to achieve suitable quality for long-term preservation. The Metropolitan Museum of Art imposes quality control by normalizing all digital resources to stable file formats, regardless of what format they were originally submitted. By having all their digital resources consistently in standard "archival" file formats, there is far less disparity among assets; as a consequence, these formats are less likely to become obsolete as quickly as other proprietary formats. If migration to a new system or a new format is needed one day, at least the Met's team only has to orchestrate the migration of a limited number of formats.

The Museum's digital asset management system, NETX, also aids in basic repository needs and has an easy user-interface that enables more efficient distribution and access to digital resources. For example, upon ingest, NETX will run a MD5 checksum to make sure there

are not duplicates within the system to reduce storage waste (NETX 2013). Whether this checksum feature in the software runs *routine* checksums over the stored digital assets is undetermined. If it does, the system would be performing a key digital preservation activity of 'fixity checking' to determine if there is any kind of bit rot, metadata changes, or other tampering that could compromise the long-term viability of a digital resource. If it does not, the checksum feature regardless ensures that the digital assets deposited have consistent fixity checks, and virus scans upon ingest.

When reflecting back on the Library of Congress' National Digital Stewardship Alliance "Levels of Digital Preservation" as mentioned in Chapter 3, the Metropolitan Museum of Art would be commended for achieving many of the qualifications for Level 1 and/or Level 2. These strengths from Level 1 include:

- Two complete copies that are not collocated
- For data on heterogenous media, get the content off the medium and into storage system
- Check file fixity upon ingest
- Identify who has authority to read, write, move, and delete individual files
- Ensure backup and non-collocation of inventory
- Encourage use of a limited set of of known open formats or codecs

From Level 2, the Metropolitan Museum of Art fulfills the following NDSA recommendations:

- At least one copy in a different geographic location
- Document storage systems and storage media and what you need to use them
- Inventory file formats in use
- Store descriptive and administrative metadata

The accomplishment of establishing many digital preservation recommendations is notable. Significantly, The Metropolitan Museum of Art is building its assets, as well as building upwards its management and stewardship activities for digital resources. As an institution that strives to follow the best professional standards in all aspects of its work, there is no doubt that the Metropolitan Museum of Art will continue to climb the ladder towards the final level of the NDSA's standard for digital preservation recommendations.

On the most basic level of digital preservation, without attention to bit-level preservation, there will be no digital assets to display or use in the long-term. Therefore, the Met's proactive diligence to using bit-level preservation strategies, including backup copies,

metadata control, and off-site magnetic tape storage, addresses its strength in accomplishing a basic level of digital preservation. Regardless, it is still important to consider the recommendations from the Library of Congress and other digital preservation resources: "bit preservation does not address the long-term needs for appropriate software to display and use the 'photographs and descriptions that will aid users' understanding of when, where, and how the photographs were taken and, at an even more complex level, the subjects of the photographs and their context within larger events" (Anderson 2011). Bit preservation is, however, the building block for a more complete set of practices and processes to ensure the survival of digital assets over time. As digital asset managers and stewards, bit-level preservation remains a practical step to keep digital content viable now.

The Metropolitan Museum of Art sets the example for how museums should progress towards digital preservation. It is important to recognize that digital preservation can be accomplished within stages, and from a practical perspective. From the onset of its major digitization and access initiatives, the Met has practiced due diligence to manage and organize its digital resources to ensure ongoing access and in many ways bit-level preservation. As needs change, the Museum is on the pathway to adopt more policies, which will lead to not only more streamlined processes between its museum, library, and archive units, but will also lead to future digital preservation enhancements.

Chapter 8: San Francisco Museum of Modern Art

Introduction to SFMOMA

A true contemporary art pinnacle for California, the San Francisco Museum of Modern Art (SFMOMA) has been dedicated to collecting art that exemplifies important challenging and contemporary practices within the art historical canon since its founding in 1935. As the first modern art museum on the West Coast, SFMOMA has a reputation for being one of the first museums to recognize photography as a legitimate art form, as well as embracing fresh ways of seeing and thinking about the art world by exhibiting and collecting a variety of both modern masters, and younger, less-established artists (SFMOMA 2015; SFMOMA 2015b). The SFMOMA boasts a collection of about 30,000 works including photography, painting, sculpture, architecture and design, and media arts (SFMOMA 2015c). The museum's interest in collecting works of art that challenge and express the way that we think and experience the world today has contributed to SFMOMA's forward-thinking with regards to modern technology, whether that be in the form of collections, education, or digitization.

Located in the heart of the San Francisco metropolis, the SFMOMA is surrounded locally by the fast-paced world of technology. The reputation for nearby Silicon Valley as the country's hub for technological innovation has slowly crept north within the last five years as companies like Facebook, Twitter, Google, and Apple have made their presence and influence felt within the city of San Francisco. The changes occurring within the skyline and demographic makeup of San Francisco will certainly take effect on the changing art scene. SFMOMA is located within proximity to other major new media-focused art entities such as the Bay Area Video Art Coalition, Pacific Film Archive at Berkeley Art Museum, and the Kramlich Collection, which is the largest digital art collection in the United States. SFMOMA's location and local culture positions it well to be a leader within the changing face of art collecting and the dialogues surrounding the world's increased dependency on technology.

This chapter will discuss SFMOMA as a case study within three sections: 1) the Museum's relationship to digital technology; 2) the status of digital preservation; and 3) an

analysis of the museum in the context of digital preservation best practices. SFMOMA's relationship to digital technology will be greatly focused on two avenues, the New Art Trust and Matters in Media Art, which are important contemporary art dialogues that relate to digital preservation. The status of digital preservation at SFMOMA is divided into two parts, the digital art vault and the digital asset management system.

Relationship to Digital Technology: Matters in Media Art and Digital Assets

This section will outline the relevancy of Matters in Media Art to digital preservation in museums, and a major connection for SFMOMA to digital technology. Additionally, the section will outline the major digital assets created at SFMOMA that are also valuable materials with ongoing preservation and access concerns. In lieu of avoiding redundancy, since both SFMOMA and the Museum of Modern Art are involved in Matters in Media Art, the information below will provide context for Matters in Media Art that is relevant for both this case study, and Chapter 9: The Museum of Modern Art.

The New Art Trust and Matters in Media Art

In 1997, Pam and Dick Kramlich founded the New Art Trust (NAT), a non-profit consortium for the advancement, collecting, preserving, exhibiting and understanding of of time-based media scholarship, for works such as video, film, audio, and computer-based installations (ArtDaily 2008). This international research collaboration involves four institutions that are leaders in time-based media art: the San Francisco Museum of Modern Art, the Tate Modern London, The Museum of Modern Art New York, and the Bay Area Video Coalition, San Francisco (Art Daily 2008). The major initiative that has evolved from the NAT is Matters in Media Art, "an ongoing project that aims to develop guidelines for the care and preservation of time-based media works..."(SFMOMA 2015d).

Begun in 2003, Matters in Media Art constitutes a multiphase project whose aim is to produce best practices and guidelines based on the collaboration of curators, conservators, technical managers, and registrars; the results are published digitally on the Tate Modern's website (Tate 2015). The first two phases of the project from 2005-2008 focused on collaborating on the process of acquisitions and loans of time-based media. These documents

discuss best practices for cataloging and minimum metadata capture, pre and post-acquisition questionnaires for artists, and properly documenting installation requirements in order to better understand and prepare for long-term preservation/exhibition (Tate 2015b). The process of loaning time-based media artworks is not one that had any kind of standard within the museum field before the Matters in Media Art consortium published its findings. For example, Matters in Media Art put forth recommendations and templates for interviewing artists before and after accession in order to gather the necessary information about the an object *before* preservation issues even arise; this is a very different process than what is used for traditional collections in which such questions would not be asked until preservation complications are recognized (Harvey and Mahard 2014, 8). Additionally, the guidelines that Matters in Media Art published regarding loans covered condition reporting, facility reports, loan agreements, and budget expectations; these are all familiar topics within the museum field, but when considering ephemeral artworks, had been very unfamiliar territory (Tate 2015c).

Since the third phase in 2008, Matters in Media Art has used many of the technical protocols that museum registrars and collection managers are familiar with, but addressed within the special category of digital media, which lies somewhat outside the normal doctrine of thought for collection management in the museum profession. With these best practices published, the third phase of the project now aims to "expand content to keep pace with changing demands, not only to reflect new media formats that artists are using today, but also to extend this model for exchange…[between] our connections to other networks of allied research and practice" (Tate 2015 b). This fostered sense of exchange and inter-museum collaboration brings to mind the importance of collaboration between LAMs discussed in Chapter 4.

Although not explicitly described this way in their publishings, Matters in Media Art promotes a type of preservation technique known as *encapsulation*, or the practice of maintaining digital objects by linking all the necessary files and content that contains information required for the object to be deciphered, understood, or accessed (SAA 2015b). In essence, encapsulation is about storing technical, descriptive, and preservation metadata with an object (such as by using logical structures like "containers," "bags," or "wrappers") so that it may be stored for long periods of time (National Library of Australia 2001). The process of using an

Information Package (like an SIP or AIP) as defined by the OAIS Reference Model is also a form of encapsulation (Paradigm 2008).

By gathering vital technical and display information about an art object through artist interviews and by following other metadata guidelines such as those from Matters in Media Art, this metadata grouping process lessens the likelihood that any critical components needed to render a digital object will be lost. In addition, encapsulation is considered a key element of emulation (MIT Libraries 2012d), which as described in Chapter 3 of this thesis, is a major digital preservation technique that involves using emulator software to render an obsolete format on new technology. An example to consider might be that of a video game that has become an accessioned object in a museum. Assuming that the museum has used best practices such as those made by Matters in Media Art, and has encapsulated a broad range of metadata, the museum should then be able to use those "bagged" files as instructions for running an emulator such as a Universal Virtual Computer (UVC) to render the video game in the future. A UVC is a computer program that is independent of any existing hardware or software that can simulate the basic architecture of every computer, which allows users to create and save digital files using any application of their choice. Exhibiting and maintaining the video game in the future would only require a single emulation layer --that between the UVC and the contemporary computer being used (Tristram 2002). This example exemplifies how encapsulation and emulation are often dependent upon each other to be successful.

Matters in Media Art does an excellent job in explaining encapsulation guidelines for the museum collection context. However, it is important to understand that on its own, encapsulation cannot preserve digital records; it is only a method that prescribes how digital objects will be reconstructed in the future or how accessibility should be preserved (Boudrez 2005, 5). What it does is ensure that the metadata about the digital object's original relationships is packaged with it, and so aids in the future employment of both preservation strategies of migration or emulation (Boudrez 2005, 5). The encapsulation methods described by Matters in Media Art is an effective way to track such relationships with special consideration for artwork and artist intent.

Matters in Media Art also contributes to the digital preservation dialogue by addressing basic bitstream and format preservation within the published guidelines for the 'post-acquisition

stage'. These guidelines recommend the development of a conservation plan which ought to consider: "installation equipment (maintenance requirements and equipment replacement), media migration cycle, storage specifications, future conservation strategies and costs"(Tate 2015b). Based on the conservation tactics labeled here, Matters in Media Art recommends a combination of digital preservation strategies presented in Chapter 3: the "computer museum" strategy and migration. The combination of these strategies would be ideal for the museum context, although preserving original technology media is inevitably a short-term solution, and migration is not always suitable for original artworks when historical context may be compromised within the migration process. .

Matters in Media Art, including SFMOMA's direct participation, is one way in which the conversation about digital preservation within the museum context has started. There are similar initiatives in the field that involve digital preservation and museum inter-dialogue, such as the Variable Media Initiative at the Guggenheim Museum, and international projects like DOCAM (Documentation and Conservation of the Media Arts) Research Alliance from Canada. Even more art collection-focused initiatives include: the Independent Media Arts Preservation (IMAP), the Electronic Arts Intermix (EAI), the Smithsonian's Time Based Media Art Initiative, the Association of Moving Image Archivists (AMIA), the European Commission on Preservation and Access, and the Electronic Media Group (EMG) of the American Institute for Conservation of Historic and Artistic Works (AIC) (Tate 2015d). Since there are an immense number of projects and work revolving around the care and conservation of digital artwork, the genre of contemporary art collections dominates the conversation about digital preservation within museums.

SFMOMA's Digital Assets

Preservation of accessioned collections is inevitably mission-critical for any museum. However museums steward more than just digital artworks; and such materials such as databases, digitization projects, emails, archival records, institutional photographs, etc. are also worthy of some kind of digital preservation cycle.

Out of all of SFMOMA's digital collections, it's institutional digital asset holdings are the most vast, and mostly made in-house to serve the Museum's public mission online. The Museum is involved in a variety of projects that engage digital platforms, which has resulted in an accumulation of assets that need management for long-term use. Such projects include: SFMOMA's Storyboard, "a digital hub for texts and video, dialogue, and a constellation of outside links offering windows onto the worlds of SFMOMA artists and artworks"; Google Art Project, an online compilation of high-resolution images and virtual gallery tours from a broad range of art institutions; Explore Modern Art, an IMLS-grant funded project that is an online learning environment that integrates interactive multimedia programs, collections information, and calendar of public programs/events; Steve: The Art Museum Social Tagging Project, in which the public can create labels that describe each museum image they view (similar to the social tagging system used on Flickr) that will generate a user-based taxonomy to help close the gap that exists between the way that art is described by museums, and the way in which it is understood by the public; among other collection digitization and archive projects (SFMOMA 2015e).

The many digital projects managed at SFMOMA involve layers of special APIs and certainly a large amount of visuals that were created by SFMOMA's photography team. As an extension of the museum, the online access to these educational projects and online collections are key and vital elements of how SFMOMA reaches its contemporary audience. Therefore, the many photographs, data, records, and online projects can also be considered mission-critical for SFMOMA's presence as a cultural institution within the digital age.

The ways in which SFMOMA addressed the long-term preservation and access to its digital artworks and important digital assets will be addressed in the section below. Most of the data presented was derived from a 2015 interview with Layna White, the Head of Information and Access at SFMOMA, as well as from recent interviews of Mark Heller from the Smithsonian Time-Based Media Initiative, and of Marla Misunas from the Library of Congress Digital Preservation blog, *The Signal*.

The Status of Digital Preservation at SFMoMA

While SFMoMA is a key player in the collections-focused initiatives sponsored by Matters in Media Art, their digital preservation efforts are even more wide-ranging. Like most other museums in the field still, the SFMoMA does not have a mature digital preservation system yet, but they do employ many established digital preservation tactics for bitstream preservation, digital asset management, and archiving digital art. There are two main aspects to the digital preservation efforts at the SFMoMA: their digital art vault (also referenced as the digital art server), and their digital asset management system (DAMS) (White 2015). While the digital art server supports the long-term care, access, and display of digital or new media artworks, the digital asset management system supports the Museum's actions and thinking around easy, reliable access to assets related to artworks and relevant for their public mission (Manus 2014b).

The Digital Art Server

Currently, SFMOMA considers itself further along with its project to preserve digital artworks than the digital asset management side of their efforts (White 2015). Overall there still is not a substantial "one size fits all" solution for SFMOMA, and this is especially true for their art server, or art vault. Since the time-based media artworks each have their own individual preservation needs, SFMOMA's art server functions more like a stable preservation vault; although adjusting this to a more active preservation system (or digital repository) is a next step for the museum (White 2015).

The development of SFMOMA's art server started in 2008 when the Museum acquired two works that were commissioned in 2000 for an online exhibition called e.space2 (Sanchez and Smith, 2013, 1). One of the works was a multimedia/website work called *Predictive Engineering II* by Julia Scher, and the other was *Agent Ruby* by Lynn Hershman Leeson, which was a Java program that presented an artificial-intelligent website based on a character Tilda Swinton played in Hershman's 2002 movie, Technolust (Sanchez and Smith 2013, 1). *Agent Ruby* and *Predictive Engineering II* had both been live, running, and accessible online artworks

since 2000, and remained active even after it was officially acquired by SFMOMA (Murray 2014). These two pieces led Curator of Media Arts, Rudolph Frieling, and Director of Conservation and Collections, Jill Sterrett, to consult the executive director of the Bay Area Video Coalition (BAVC) for collection management advice now that they were collecting software (Sanchez and Smith 2013, 2). BAVC's preservation department had already been helping SFMOMA on and off with digitizing many kinds of old video formats such as 1" open reel, U-matic, Betacam, and laser disc (Sanchez and Smith 2013, 2). This relationship led SFMOMA to partner with Mark Heller as a consultant in 2009, who was the digital media specialist of BAVC at the time (Sanchez and Smith 2013,2). The art server that resulted from the conservation needs of SFMOMA's first software-based artworks was created as a collaboration between Mark Heller, conservation fellow Martina Haidvogl, and SFMOMA's Team Media (White 2015).

Team Media is a group at the Museum that meets once a month and includes staff members from conservation, curatorial, exhibitions, and registration (White 2015). Team Media is the key interdepartmental group that helps care for the artwork, its files, the digital art server, and SFMOMA's stewardship practices (Manus 2014b). Once a fully-realized project, the art server has resulted in various in-house standards that deal with preservation of digital materials. Many of these were called upon from other institutions (like those involved with Matters in Media Art), advice from experts such as Karen van Malssen of AVPS, and working with software vendors themselves (White 2015). The in-house standards created around the art server will be discussed below.

The process of digital preservation itself for the art server is based upon the need of the specific artwork (White 2015). SFMOMA does not need to use a particular selection/appraisal process for their digital art server because *all* artworks of this category are immediately prioritized to be ingested into the art server once it has been approved to be in an archival format from the artist (White 2015). While ideally the object would already be in a format considered archival once acquired by the Museum, such is not always possible. This can be problematic since the Museum does not "normalize" artworks that are in the art server in order to avoid compromising the work's artistic integrity (White 2015). Normalization is somewhat necessary

to prevent accumulation of proprietary formats in the vault, and for the Museum to control its own preservation standards upon ingest. So, if the Museum is unable to consult the artist (in the case of deceased artists) on these format requirements before acquisition, a whole set of special attention and consideration must be applied to that work. However, when SFMOMA receives certain types of media, such as tapes which are considered common formats, certain normalization standards can be applied such as digitizing to 10-bit uncompressed video in a QuickTime wrapper (White 2015). The ingest of such digital materials has been more spelled out for SFMOMA than the process for software artworks (Sanchez and Smith 2013, 3). SFMOMA rightfully identifies a distinction between preservation approaches for video works and software-based art (Murray 2014). Whereas video objects are considered unchangeable after creation - they are more fixed - software-based artworks on the other hand have many moving components, including parts that are at risk of obsolescence and technical vulnerability within the fast-paced world of software updates and changing platforms (Murray 2014).

This next section uses a practical example from SFMOMA's experiences to explain how variability of software led to new best practices in the field. The two artworks that motivated the development of SFMOMA's art server, *Ruby* and *Predictive Engineering II*, are both software works, but are otherwise nothing alike (Sanchez and Smith 2013, 3). Each work contains a number of components that need different considerations for preservation, so they cannot be batch processed like more normalized materials (such as video). As described by Mark Heller:

> "Ruby *had a Java program that was a natural language interpreter, which communicated via a Web server to a Flash multimedia interface. You would enter user input and it would be would be analyzed by a Java program, then it would scan a database that Lynn Hershman's programmers created to return an artificially intelligent response. It was a very exotic set of components. Julia Scher's piece was also a network of components; it was a little less complex than* Agent Ruby, *but it was still about 11 HTML pages, each containing a Macromedia Flash object. Each Flash object contained hundreds of animation layers—images and sounds, then ActionScript code to make them interactive"* (Sanchez and Smith 2013, 3).

As Heller described, these two software-based works are written in different source code languages and therefore have very different components. The technical complexity behind the

variety of components and behaviors of these kinds of artworks led to the conclusion that a new form of documentation was required for ingest into their preservation system (Sanchez and Smith 2013, 4). This new form of documentation has become an integrated part of SFMOMA's standards for the digital art server, which they call the "technical narrative." This "technical narrative" standard consists of four parts that integrate various forms of metadata capture:

1. *A very high-level summary of how the work operates as a whole.* This is meant to be a platform-neutral functional description of the work.
2. *An examination of the components, what they do, and their related functions.* This section dissects each component individually, such as the Java natural language processor in *Ruby*, or the Flash files in *Predictive Engineering II*. In addition, this section provides a high-level look at how these components work as a complete system.
3. *A detailed description of the artwork in its current state upon acquisition, including technical metadata like hardware, software, operating system, environment, languages, code, versions, etc.* This section is meant to acquire an understanding of how the technical requirements serve the operational requirements of the work.
4. *An analysis of the current technology, its longevity, and evaluation of obsolescence.* This section considers the long-term stability of the piece as it stands now, and possible preservation strategies outside of maintaining its technical requirements. "For example, both Ruby and Predictive Engineering have Flash components, and we know that Flash is going away. It doesn't play on Apple devices, and there's no longer support in Android. So we looked for alternatives (like HTML 5, for example) and a strategy to migrate those Flash components to them" (Sanchez and Smith, 2013, 3).

As an internal standard being used to prepare artworks for long-term storage, SFMOMA hopes that their template for the "technical narrative" can be shared and used with other organizations (Sanchez and Smith 2013, 4). For example, the Tate Modern is interested in the "technical narrative," and they are employing it on their software-based works. While this standard developed at SFMOMA is recognized like an official ISO standard, the fact that it was developed within the museum community has garnered interest in sharing best practices for

digital preservation. Although the OAIS standard was not used in designing the "technical narrative" when it was developed in 2009 (OAIS was not considered in depth at SFMOMA until about 2012), there are certainly parallels between the two (Sanchez and Smith 2013, 4). In the OAIS Reference Model, the concept of maintaining Representation Information is a major overlap. In speaking to this inadvertent relationship between SFMOMA's "technical narrative" to the OAIS Representation Information, Mark Heller describes:

> *"The general idea is that you have this digital object -- the bits -- and you need to represent it in the way it was originally intended. So what documentation do you need? In these cases we are discussing, how do you maintain the artistic integrity or the intent of the artist? When I think of the technical narrative and then read the OAIS model, I can see a relationship between what we have done and its concept of representation information"* (Sanchez and Smith 2013, 4).

As made evident here, SFMOMA has created its own best practice and then audited it against the de facto OAIS standard, which adds considerably to its efficacy.

Another way SFMOMA's in-house "technical narrative" has aligned with digital preservation standards is through its use within a lifecycle approach to help manage change. From the technical narrative, the managers of the digital server can identify specific components of a work and ask, where is this in its life cycle? If nearing obsolescence, what options do they have to upgrade it? (Sanchez and Smith 2013, 5). In the case of *Ruby*, the program was running on Java 1.4, which was quickly reaching the end of its life as of several years ago. The Museum was able to find the latest version of Java and recompile the code in Java 1.6. Before even doing this step, Mark Heller presented the migration plan to Lynn Hershman Leeson and reviewed the work's upgrade needs to make sure it will continue to run for the next five to ten years. By doing so, the Museum was able to make sure that any changes to the work stayed true to her original vision (Sanchez and Smith 2013, 5). This approach to anti-obsolescence follows closely with the guidelines from Matters in Media Art, but also with the digital curation approach.

A couple of other anti-obsolescence tools used by the managers of SFMOMA's digital server are file format databases such as FITS and PRONOM (Sanchez and Smith 2013, 6). After the technical narrative has been completed, artworks are run through the open-source Bag-It program from the Library of Congress to extract metadata and run a checksum in order to

ensure that the piece is fit for long-term storage. After ingest and backup onto magnetic LTO tape, further checksums are automated regularly in order detect if a file has been corrupted over time, and allow the Museum to react as soon as possible. If a file is corrupt, the backup from the tape storage will be used, after running another checksum, to replace the corrupted file.

Regarding other standards in use, stemming from the information science field, SFMOMA also incorporates standardized metadata schemas including Dublin Core, VRA Core, and CDWA Lite (White 2015). PREMIS is a standard that SFMOMA strives to follow, but they are currently not sure if they are capturing all the core fields to be considered aligned with the PREMIS standard; this will be an area of further evaluation in SFMOMA's future. The use of PREMIS will likely become more relevant to SFMOMA's needs as it ramps up its "level" of digital preservation (White 2015).

While currently SFMOMA is not following digital repository standards like ISO 163163 (TRAC) or something similar, they are working to launch a digital repository one day that is OAIS compliant (White 2015). As a next step, SFMOMA's contractors, Mark Heller and Martina Haidvogl, are currently evaluating if a preservation software, namely Artefactual's Archivematica, could work for SFMOMA. So while the art server exists, it seems that this is a temporary stage for SFMOMA's digital preservation program; the future of SFMOMA's digital preservation efforts lies in integrating an official system that follows ISO standards, and that encourages a more robust digital preservation policy that will perhaps incorporate the "technical narrative" standard already in use at the Museum (White 2015). As a closing thought, Mark Heller succinctly describes SFMOMA's current approach and situation regarding their emerging role in the world of time-based media collections:

> *"...I was asked, "What will you do when you have to deal with 1,000 works?" My answer was, "Well, we have about eight right now; so I don't think we're going to have to deal with 1,000 works any time soon." We're giving individual attention to all these works because things are just emerging. In a way we're lucky because the collection is quite small and we can pay a lot of attention to each work and define standards where they feel appropriate. So hopefully when 1,000 works in a collection is the norm, we will have some kind of structure. We're exploring, discovering, and defining that now"* (Sanchez and Smith 2013, 7).

SFMOMA may only have a handful of software/web-based works of art, but its entire collection of 250 pieces of time-based media includes a diverse array of formats from single and multi-channel video, slides, film, and digital photography (Murray 2014). With such a wide range of new media-type collections in its holdings, the Museum must be capable of dealing with a variety of complex technical requirements, such as display parameters, but also artists' intention when making preservation decisions. Ultimately the preservation activities around digital or new media art are tied to keeping the piece alive through *using* it, and for museums, this includes installing and exhibiting works over time and in different situations. The digital art server is therefore designed and managed with use at its core mission (Manus 2014b). SFMOMA's art server and involvement in Matters in Media Art has resulted in good work to launch standards that highly engages the artists throughout the process, as well as digital preservation standards, and therefore responsible stewardship of the collection.

The DAMS

Although SFMOMA may only have a limited number of time-based artworks, the amount of other types of digital assets in its holding is comparably vast. As said by SFMOMA's Collection Information Manager, Marla Misunas, in an 2014 interview with the Library of Congress digital preservation blog, The Signal: "In a way, we're building a library or directory of artwork that anyone can access"(Manus 2014). This section will discuss how SFMOMA uses its digital asset management system as it relates to its digital preservation efforts. Most of the information gathered in this case study was derived from an in-person interview with Layna White, Head of Information and Access at SFMOMA in May 2015.

With so many worthy projects within SFMOMA's scope, the museum materials contributed to all these avenues (photos of collections, multimedia features, exhibit media, educational materials, etc.) are being used by a plethora of departments and staff, many of whom are not often interacting with each other or using the materials for the same goals. Thus these digital assets require a high-level of management to avoid loss, or human tampering, maintain metadata and aesthetic standards, monitor format and display resolution standards, and streamline museum branding (White 2015). Enter the tools used by SFMOMA to achieve the

goal of well-managed digital assets: the Collection Management System (CMS) and the Digital Asset Management System (DAMS). These software tools allow for efficient sharing, and a directory-like flow of information for the museum staff, and ultimately to the open public-facing side of the museum via their many digital projects. In addition to these functions, the DAMS at SFMOMA is being used as a kind of control vault for storing digital assets (besides the artworks in the digital art server). The DAMS stores and controls standardization of digital materials until they are needed for access throughout the Museum's departments. While a digital asset management system is certainly not to be confused with a digital preservation system, SFMOMA is effectively using this new technology as a tool to promote certain digital preservation practices.

SFMOMA has been using a software called Embark as its collection management system (CMS), or database software, since 2003 (White 2015). While the Museum may not consider Embark to be the most perfect CMS for them, it was the best out-of-the-box option without needing any major reconfiguration for their purposes (White, 2015). Marla Misunas describes SFMOMA's uses for Embark:

> *"...to track, document and manage our collections and works loaned to us. Staff members around the museum contribute to documentation about our collections via the system, starting before objects come in or accession, through their "life" cycle at the museum...The database is our authoritative source for information used by our digital asset management system, our online collection and just about any project where object data appears"(Manus 2014).*

As implied by Misunas, the CMS works in tandem with an overarching *digital asset management system*, or DAMS. The first DAMS employed at SFMOMA was called Media Bin (White 2015).

Around 2004 SFMOMA began to rapidly acquire digital assets, especially as more photography was being produced in digital formats, as well as the accumulation of other kinds of digital materials. The Museum was prompted to look into the need for a digital asset management system, which is a type of software that was more widely used in the big business world at the time; this made finding the right system for the museum context a little more challenging (White 2015). After putting out a request for proposal, and working with vendors to

find a product that meet their needs, Media Bin was deployed in 2006. The Museum of Modern Art in New York and the Metropolitan Museum of Art, who also used Media Bin, were other great influencers for this decision (White 2015).

By 2010 however, SFMOMA came to realize that Media Bin was not robust enough for the museum's growing needs and expectations, especially in regards to dealing with video, a regularly accumulated asset, as SFMOMA continues its oral history project to record artist interviews for research and exhibit documentation (White 2015). In addition a huge number of relevant photos was continuing to be produced museum-wide, but the metadata capture was not

easily attainable within the individual staff departments because Media Bin was not particularly user-friendly. A new request for proposal was put in place, and SFMOMA looked into trying a system called Net Exposure (NETX), which was put into place January 2015 (White 2015). The Museum of Modern Art in New York was already using NETX by January 2015, and the Metropolitan Museum of Art was also switching to NETX as of April/May 2015. The commonality of this software between other leading institutions made the decision for SFMOMA easier, but also the prospect of future sharing and exchange of data more promising (White 2015).

Migration of digital material in Media Bin to the new system was very difficult. The descriptive metadata crosswalked to NETX easily, but the technical metadata failed to crosswalk between the two softwares smoothly. Head of Information and Access, Layna White, highlighted this data loss challenge to be a major foray for SFMOMA with digital preservation issues outside of its conservation practices (White 2015). Luckily the technical metadata remained in the old system, so SFMOMA resolved the issue by working closely with the vendor about the metadata crosswalk. To do this the Information and Access and IT departments at SFMOMA had to communicate very carefully with NETX to help them understand why the Museum valued this technical metadata, and why it is important for them to hold onto that information for long-term usability of the assets (White 2015). From the vendor's point of view, keeping that level of technical metadata is outside their concerns since most technology industries are used to thinking about data use in the 1-5 year span, not the long-term archival timelines desired by cultural memory institutions (White 2015).

The types of assets involved with this technical metadata migration debacle mostly included thousands of still images, some that relate to collections, others that relate to other departments such as exhibition, marketing, and education (White 2015). For example, dozens of images can be made to document views of an artwork as installed in a particular setting. This photo documentation can help the Museum to understand how visitors interacted with the work at the time and within the space; this type of documentation is especially relevant for time-based media works that have variable exhibition possibilities. The photo documentation is vital for the long-term planning and usability (therefore the preservation) of such works. These images, audio, and video files related to artworks and artists, as well as data about those files, are stored in the internal NETX system (Manus 2014b). As mentioned previously, video is a growing digital asset for SFMOMA and will be the museum's future foray into digital archiving (White 2015).

Managing the DAMS at SFMOMA is an extremely collaborative effort, so the number of producers, consumers, and administrators is vast (White 2015). The departments that deal most directly with NETX are the Visual Resources Department, Information and Access, Registration, Conservation, Curatorial, Web Team, Information and Technology, and Marketing (White 2015). The production and management practices of the DAMS are informed by community and industry guidelines within three categories: *metadata*, *formats*, and *storage* (Manus 2014b).

The first, metadata, is considered a quality control issue among the many users of the DAMS and its stored assets. As discussed in Chapter 3 of this thesis, the control of metadata capture is a vital element of digital preservation, especially when actively using a schema designed for long-term preservation purposes. At SFMOMA, the Information and Access team acts as quality control for SFMOMA's data by reviewing for standards that are maintained internally, but derived from the greater data management community. SFMOMA refers to controlled vocabularies and metadata schemas from the Library of Congress Name Authorities, the Union List of Artist Names, the Thesaurus of Geographic Names, Dublin Core, CDWA Lite, VRA Core, and PBCORE (for audio). This metadata quality control is processed before and sometimes after ingest into NETX (White 2015). Using various industry-wide standards that

are widely applied allows SFMOMA to seamlessly contribute to federated databases like Artstor or the Google Art Project (Manus 2014). The participation in such "databases" is somewhat of a digital preservation practice itself because the Museum is able to extend its holdings under other web servers outside the Museum; those managing Artstor and the Google Art Project save the high-quality images and related metadata contributed to them for as long as needed, and so promotes the preservation philosophies of a redundancy system.

The other practice used with by SFMOMA's teams for their digital asset management system is normalizing objects to industry *standard file formats*. This is a way in which SFMOMA monitors and maintains format preservation within their current system. As noted in Chapter 3 of this thesis, format preservation involves being aware of formats that are considered de facto archival standards by the digital preservation community, but also actively normalizing assets from proprietary formats, and regularly refreshing the media and checking for at-risk formats over time. SFMOMA uses format registries such as PRONOM and FITS to aid in their format lifecycle assessments (Sanchez and Smith 2013). Still images are normalized to TIFF or DNG files as master copies, and to JPEG format for use or distribution copies. Most files kept for conservation purposes are normalized to DNG files, including some exhibition and installation photos (White 2015). For photos of the collection, SFMOMA keeps the raw files. Video assets can be normalized to 10-bit uncompressed formats, or they will take the least compressed file format that the artist or producer can provide. The recommended file formats maintained in SFMOMA's stewardship is driven greatly by other practices in art institutions, like NY MOMA, and by recommendations by their vendor partners (White 2015).

SFMOMA's digital asset management system functions as a quality control unit, dissemination unit, but also as a storage unit for digital assets that come from all directions in the museum. In terms of *storage*, the DAMS is backed up nightly to magnetic tape storage, and their servers are mirrored regularly at another museum outside SFMOMA's region (Manus 2014). To safeguard their raw master image files, these assets are saved in a sequestered section of the server; while any corrected master files are stowed safely into the DAMS, from which derivatives can be made as needed (Manus 2014). The magnetic tape storage is mostly

maintained by SFMOMA's IT team, while the corrected master files and distribution copies are managed by the Information and Access, and the Visual Resources departments (White 2015).

While the assets stored in the DAMS come from a variety of sources and relate to a variety of purposes, from marketing photos to conservation documentation of artworks, NETX is currently unable to interface with SFMOMA's other digital management entities like Embark or the art server (White 2015). In SFMOMA's future plans, their conservation fellow Martina Haidvogl is interested in finding a program that would function as a "portal" into the art server's holdings and that connects the descriptive metadata between all of SFMOMA's systems (Embark, NETX, and the art server). Until such a program exists, the museum's various data sites will remain separate (White 2015).

A rising topic at SFMOMA is archiving email communication (often between artists and the museum) about digital artworks or other important topics that may be needed long-term (White 2015). There is a need to reconcile original emails as archive materials and relate them with the descriptive and technical metadata entries for artworks in the art server (White 2015). Currently email is not considered an archive-worthy asset at SFMOMA, although Layna White strongly recognizes the need to track such communication. White worries that email memory is being underestimated; once the Museum starts to lose, misplace, or empty its email buildup, there is potential data, correspondence, and/or important artist dialogue that could be lost. Such institutional knowledge recorded in emails can be pertinent to the system, as well as to the collections themselves (White 2015). Email preservation in of itself has garnered much discussion in the digital preservation field, but there are many issues surrounding its execution including legal parameters, search capture of important email content, technologies, storage, formats, etc (Prom 2011). However, SFMOMA has the support of such digital preservation issues from its tenured staff and will likely revisit this issue when their new DAMS system is more mature and they are well poised for the next project.

Arguably the end goal of any kind digital preservation system is the *continued access* to the digital materials in questions (Chapter 3). On this note, one of the major outcomes of utilizing a DAMS in a museum setting is the increased level of transparency and available information provided to the museum field and to the general public. Via internet access, material

can be monitored by the DAMS on the backend, but published online for front-end users. SFMOMA provides access to its digital collections by way of various public portals on their website. Examples of these portals include their Open Public Access Catalog of its collections (called Artscope) and the online initiatives of Explore Modern Art (White 2015). SFMOMA has over 11,000 images on their website now, which amounts to over a third of its collections. The digitization of SFMOMA's collections and publishing online is an ongoing project.

Policy and Future Plans

With the many protocols, standards, and workflows happening between SFMOMA's various data storage and management venues, it would be ideal to have a policy of some kind to guide each of these initiatives. Because museum policies are such high-level mandates and digital preservation is not yet a common program in museums, such policies are far and few between. SFMOMA does not have a digital preservation policy currently, however they are still in the process of developing their OAIS compliant system for its digital artworks. Since the museum uses its DAMS as a storage and management tool for digital assets, they do have a digital asset management system policy.

The Digital Asset Management System Policy at SFMOMA functions mostly for internal staff and outlines for users of the DAMS the ideals for contributing and sharing digital assets. Because the policy is considered for internal use, it is not available for distribution to the public. The policy places much emphasis on the responsibility of the staff member to provide enough information about a digital asset in order to contribute to museum-wide clarity and discoverability of digital materials. SFMOMA's Digital Asset Management Policy looks to help staff members, and advocates for regular use of the software, as well as encourages a sense of generosity towards sharing assets and data in order to ultimately optimize the DAMS' capacity as a useful tool. This policy is focused on collaboration, and clarifying how the DAMS can help the museum as a whole; but it does not cover distinct preservation issues, except in reference to its statement on *trust*. The statement on trust, as the final section of the policy, defines the Museum's goal to sustain a trusted system that is underpinned by good practices in order to ensure the availability and integrity of assets over time (SFMOMA Digital Asset Management

System Policy 2014). Although the Digital Asset Management System Policy is not a digital preservation policy, SFMOMA does view the DAMS as a vehicle for digital preservation by way of using the software for establishing intellectual control and long-term trust over its digital assets.

For managing and guiding the work around artworks preserved within the art server, SFMOMA currently depends on its overarching collection management policy for high-level ideals around the acquisition and responsibility of stewarding the collection. For policies specifically regarding digital materials, SFMOMA looks to Matters in Media, other institutions namely the Museum of Modern Art in New York, the library field, as well as vendor consultants (White 2015). A copy of the Museum's collection management policy was not able to be provided for this research.

A large topic revolving around the importance of digital preservation policy is often a policy's purpose in implementing a secure source for funding. Because the creation of a digital preservation policy involves review from upper-management and often governing bodies, the implementation of policy can clarify a museum's commitment to digital stewardship, which often feeds into the financial needs to meet those stewardship goals. For SFMOMA, they are fortunate that all aspects of its digital preservation efforts, the art server and the DAMS, are currently accounted for in the Museum's overhead budget (White 2015). The internal group Team Media has been in place for 20 years, and the dedication and maturity of this group created a lot of trust within the greater museum. Team Media's existence and work was an very big leverage tool for institutional support for both the conservation of digital artworks and implementing a museum-wide digital asset management system (White 2015). When these initiatives were first being pitched, the biggest concern was around lobbying money for magnetic tape storage; however in the end the cost of storage was lumped into the internal funding used for the art server and the DAMS (White 2015).

For the future of digital preservation at SFMOMA, the biggest challenge facing the museum is that of staffing. They currently have Mark Heller and Martina Haidvogl managing the art server, but unfortunately neither of them are working full-time. The staff position of a time-based media conservator, is the next step for SFMOMA's future digital preservation needs.

Additional staff members that have some background understanding of digital preservation issues will also be key in the future (White 2015). Recently, SFMOMA was able to hire a digital asset manager, although it took a long time to lobby for this full time staff position (White 2015). The Museum is slowly moving forward with having the right team of personnel required to make its digital asset and digital art conservation initiatives successful, however there is concern that these projects will be held back if the Museum cannot give the current teams the manpower that it needs to move forward.

The amount of digital art that is stowed within the art server is rather small compared to SFMOMA's collections as a whole, and only amounts to about 5% of the collection (White 2015). This of course makes lobbying for money and full-time staff even more difficult. The efforts to implement a future digital preservation system, like Archivematica, is a hard case to make if only 5% of the collection is affected. However, the staff already involved with the art server see the benefit of planning for the future preservation needs of SFMOMA's digital collections (White 2015). In terms of the other side of the Museum's digital preservation work, making the case for purchasing the first, and then second, DAMS software was a hurdle, and only recently was SFMOMA able to hire a staff person to be in charge of managing that system. However, through strong collaboration between a variety of departments, SFMOMA was able to make a strong case to its governing body for the need of a system that will make access, retrievability, and storage of digital assets achievable (White 2015). SFMOMA's digital asset team holds hands with IT, the Web team, Visual Resources, and even with Marketing and Development to advocate the use and preservation ideals of the DAMS. SFMOMA's formation of strong collaboration across departments made the issues of digital asset viability a museum-wide initiative, which today serves them well in achieving control and preservation of its digital collections.

Analysis

The analysis of the San Francisco Museum of Modern Art as a digital preservation case study will be made within the context of motivations for digital preservation, maturity of program, and contributions to the museum field.

Motivations for Digital Preservation

The case for good practice in digital preservation will ultimately vary from institution to institution, however all museums share the same calling for responsible stewardship. This stewardship can be more easily advocated when considering the needs of rare, unique, collections versus that of documentation and institutional records of an institution, although both are important to continue the mission of a museum. Fortunately, SFMOMA addresses the needs for both types of digital assets, and the Museum's staff has achieved a lot in the balance of practicing due diligence to both sides of its digital collections.

The motivations for digital preservation at SFMOMA has been separate between the art collection and the records, mostly due to the specific needs of time-based artwork which is entirely more complicated and dynamic compared to the standardization and management requirements of the Information and Access team. Although separate, the SFMOMA has done good work in dedicating time and funding to support two programs that address the Museum's immediate needs. Nonetheless, the motivations for digital preservation at SFMOMA, as defined by the recommendations of the digital preservation field, are focused around the conservation of the time-based media art collection, while the digital preservation of the records management unit was motivated by the shorter-term goal of retrievability and use of those digital materials. The creation of SFMOMA's digital art vault was certainly envisioned around the requirements of long-term digital preservation, and the digital asset management system was contracted for management and use, with digital preservation as a positive side-effect that the Museum has been able to take advantage of to promote basic digital preservation for those materials.

Team Media and Matters in Media Art, as two collaborative groups for dialogue, certainly supply the SFMOMA team's education about the unique stewardship needs for time-based or digital artwork. These two groups (one internal and one external) are

'collections-focused' themselves, however as evidenced by their two robust technology systems, the staff at SFMOMA have regardless drawn the connection between the needs of its various assets, such as the needs of an original digitized film, to the needs of a recorded artist interview created by the Museum itself (i.e. between accessioned artworks that the museum is primarily responsible for, and institution-made digital assets that enrich the artistic record and research of the collections.) It is this connection, that all digital collections will share the same concerns for longevity and viability (albeit with varying levels of need), that can be recognized as a good practicing foundation for digital preservation at SFMOMA.

Digital Preservation Maturity

SFMOMA is an institution following the right path towards a 'mature' digital preservation, but it can be said that they are currently working with an 'established' digital preservation system, one that is still a work in progress and that evolves as more of SFMOMA's team collaborates with other professionals to fully realize what they can do to optimize and enhance their current systems. Since SFMOMA views its digital preservation efforts within two separate categories, the digital art vault and the DAMS, the level of digital preservation 'maturity' will be considered for each separately.

The digital art vault at SFMOMA is one of the few prevalent digital preservation systems in museums today, and as such is a best-practicing ideal for the broader museum field to follow. Their system is well-established by 2015, but is also still very much a work in progress. For this reason, SFMOMA's digital art vault is considered an 'established' digital preservation system with a bright future for more potential, or maturity, as the museum's dedication to learning and optimizing systems progresses.

SFMOMA's digital art vault, and essentially the steps to prepare works for long-term storage, were created with the Museum's needs in mind, and less focused on using digital preservation software tools and standards already in the field. For example, the Museum's 'technical narrative,' acts as a process for evaluating an object and collecting descriptive and technical metadata, and is a significant *internal* tool produced from evaluating the long-term technical needs of their artworks. Significantly, on its own, the 'technical narrative' outlines a

similar process of preparing an Archival Information Package from the digital archive requirements outlined by the OAIS model off the scientific big data and library fields. The 'technical narrative' is effective for the small collection of artworks in the digital art vault currently, but to ensure it will continue to work for the future scope of digital preservation at SFMOMA, their goal to map in-house standards to the outside standard of the OAIS model will allow the institution to ensure it is meeting the recommended object-level preservation requirements of a trustworthy digital repository that are greatly accepted in the broader digital data fields. Although SFMOMA did not initially use many standards for digital preservation, the Museum nonetheless exemplifies best practices in how such standards could be used after-the-fact for self-assessment and as tools for future improvements or modifications to a digital preservation system.

Other of the steps within SFMOMA's digital preservation system already use many laudable tools from the digital preservation community, such as Bag-It and the PRONOM format registry. The use of these tools exemplifies how SFMOMA has been able to apply digital preservation processes to the museum context. With the expertise and help from key staff personnel, SFMOMA has been able to learn and adopt new standards from outside the museum field, and use it within distinctive parts of its own digital preservation workflow. The connection between the digital preservation systems from allied fields, such as the library field, with the Museum is key to their success.

SFMOMA's digital art vault team is currently looking to enhance their current metadata protocols to the OAIS and to recommended metadata schemas, namely PREMIS and METS, in the near future. The Museum's journey to become a best-practicing digital repository is underway, and their natural pathway to best practices began first with meeting the museum's immediate repository needs, and then later mapping to standards from the digital preservation field to bolster and improve the basic architecture already put into place. Such a pathway is an ideal model because it has allowed SFMOMA to consider the protocols for digital preservation within its own capacity of time, staff, and funding first; and then to consider further enhancements by looking at standards from the digital preservation community that will elucidate how the museum can optimize or improve for the future.

Importantly, SFMOMA's use of running checksums for automated fixity checking, as well as their concept of "keeping the piece alive through *using* it" parallels concepts from digital curation, essentially of the necessary reiterative and cyclical preservation activities required for digital preservation (White 2015). Long-term stewardship of digital materials is not a static process, unlike the collection management practices that work for traditional museum collections. SFMOMA's recognition that accessing, and 'using' code-based artworks is the surest way of 'condition reporting' its status as a working, functioning artwork is a significant example of adjusting expectations for collection care in museums. SFMOMA has accepted the cyclical, and constant need for activity that is required to responsibly care for such digital collections.

Even as a new endeavor at SFMOMA, the museum has taken huge steps towards digital preservation. However, when evaluating the maturity of SFMOMA's program within the context of Nancy McGovern and Anne Kenney's *Five Organizational Stages of Digital Preservation*, SFMOMA sits somewhere between Stage Three or Four, in which the institution has established technical infrastructures but is still reaching for the additional institutional support to consolidate and institutionalize the current system for a broadening scope of digital preservation. Once the Museum has established long-term personnel in charge of the digital art vault, created more policies, and has considered additional systems that optimize their workflow (such as OAIS-compliant tools, such as Archivematica), the Museum will be well on its way to a mature system.

Within the context of the National Digital Stewardship Alliance's *Levels of Digital Preservation*, SFMOMA has achieved the basic technical requirements up to Level Three, including having one redundancy backup, monitoring fixity through regular checksums, storing standard descriptive and technical metadata, and addressing format obsolescence up front. Considering the future next steps in SFMOMA's plans, including establishing full-time staff for the digital art vault and mapping to digital preservation standards for self-assessment and improvement, the Museum will soon enough lead itself to last tier (Level Four) of NDSA's model.

The importance of the SFMOMA's digital asset management system to the implementation of basic digital preservation for institutional assets is also very prevalent to the Museum's maturity level. Although the technical architecture of the digital art vault aligns more closely with digital preservation standards like the Open Archival Information System (OAIS), the chosen digital asset management system, NETX, also aims to foster trustworthy long-term access for non-collections assets. The digital asset management system has been leveraged to fulfill many basic digital preservation activities; although the primary goal of a DAMS is streamlining and managing content for users, the quality of safe-keeping assets while they are not in use is also present in the way SFMOMA uses its DAMS.

The digital asset management system at SFMOMA achieves bit-level preservation through their backup system, maintenance of 'master copies,' and protocols to combat obsolescence. To eliminate the problematic potential of having too many proprietary formats in their system, the Museum has established standardization of the formats that can be accepted into the DAMS. Standardized format protocols and metadata capture has also aided the museum in avoiding migration problems when possible, which almost resulted in data loss during their initial migration from Media Bin to NETX. SFMOMA's team took great care, attention-to-detail, and due diligence to ensure the migration ultimately proceeded smoothly. SFMOMA's digital asset management system is an important way that the Museum addresses digital preservation, however it is also important to note that a DAMS is still not a true digital preservation system, and is mostly sufficient for short-medium term maintenance, but does not consider long-term preservation needs like that of its digital art vault.

SFMOMA overall, approaches digital preservation from a practical perspective, even between the dichotomy of the preservation of artworks versus institutional assets. The fact that SFMOMA has systems in place for both is far more advanced than most museums today, and as such is a noteworthy model. SFMOMA's journey to establishing digital preservation activities, and systems is important for the museum field to contextualize the pathway and needs of digital preservation in the museum context. Future ways that SFMOMA could improve its systems include implementing digital preservation policies for its art collection, as well as addressing digital preservation more fully in its current digital asset management policy. Such policies will

not only be an exercise for the museum to consider the holistic management and technical details of their systems, but it will better define for the staff and the whole Museum the needs and commitment for ongoing digital stewardship.

Contributions to the Museum Field

The story that SFMOMA has to share with the museum field about its journey in achieving established levels of digital preservation came from collaboration and openness to learn and advocate for digital preservation museum-wide. SFMOMA is an ideal representation of a museum that has garnered the initial work for achieving some level of digital preservation, and as a case study, exemplifies how the museum field can expect to envision the considerations for digital preservation in museums, the technology systems to consider, and the standards from the digital preservation field to apply to the museum context.

Outside of SFMOMA's good work, another contribution to the field is encapsulated in SFMOMA's culture for collaboration and sharing. So much of the work and success of SFMOMA's digital preservation initiatives stemmed from collaboration, whether that was with other institutions, vendors, or with artists themselves. The Museum's Team Media, as well as key participation in Matters in Media Art, highlight SFMOMA's collaborative nature, and how such groups successfully contributed to advocating the needs of the Museum's digital collections in order to jumpstart digital preservation initiatives. It can be concluded that SFMOMA recognizes that preserving ephemeral artworks, and institutional assets, for the long-term cannot be done alone, and requires support institution-wide, as well as from partners outside of the Museum. Collaboration at SFMOMA has provided the institution with the support and infrastructure it will need to enable a greater digital preservation mandate in the future.

Conclusion

SFMOMA's work in digital preservation is an idealized and practical example for the museum field. They are recognized by leaders in digital preservation, namely The Library of Congress, as one of the few museums working towards integrating digital archiving and

preservation into their scope (Murray 2014). They have also reached out to other museums and partnered with them to gather a strong team that can openly discuss challenges in the preservation of digital-based museum assets. There is much work ahead of them, but SFMOMA sets a leading example of how digital preservation can be achieved for the museum field. As for the application to other smaller or mid-sized museums, much of the techniques tested and used by SFMOMA --bit level preservation, format preservation, and metadata encapsulation -- do not require large amounts of funding, but are nonetheless important steps one can take to better prepare digital collections for long-term access and preservation. SFMOMA has gathered a team that has necessary skills in metadata cataloging, digital asset management, and contractors who specialize in digital media or time-based media conservation. SFMOMA's open attitude towards collaboration, both internally and externally amongst the greater museum field, is perhaps one of its most important attributes. The desire for collaboration will be vital for the survival of the Museum's digital preservation initiatives as it continues to grow in within the digital age.

Chapter 9: The Museum of Modern Art

Introduction to The Museum of Modern Art

The Museum of Modern Art located in New York City was established in 1929, the same year it acquired its very first acquisition of eight artworks. From its beginnings, the MoMA has been dedicated to being the foremost museum and educational center for modern art in the world. Today it's evolving permanent collection consists of more than 200,000 paintings, sculptures, drawings, prints, photographs, architectural models and drawings, and design objects spanning the last 150 years. In addition, MoMA stewards about 25,000 films and 4,000,000 film stills. This vast art collection is also bolstered by a library and archive that supports research and scholarship related to modern and contemporary art (MoMA 2015). The units of the MoMA's collection, library, and archive are of particular interest when considering the needs and narrative for digital preservation within the museum field as exemplified by MoMA's own journey.

To support the many exhibition, stewardship, and educational mandates of the Museum, there are seven curatorial departments: architecture and design, drawings and prints, film, performance and media art, painting and sculpture, and photography (MoMA 2011). Sharing a similar span of genres, the Museum's Library contains a noncirculating collection of over 300,000 books, artist books, exhibition catalogs, and periodicals that document emerging art history from 1880 to the present (MoMA 2011). The Museum Archives was established in 1989 to "collect, organize, preserve, and make accessible documentation concerning the Museum's art-historical and cultural role in the 20th and 21st centuries" (MoMA 2015b). The archive unit holds approximately 2,500 linear feet of historical documentation; and a photographic archive of tens of thousands of photographs, including images documenting past exhibits and the Museum's building over the years (MoMA 2011). In addition, MoMA has secured itself as a leader in film art by establishing the Celeste Bartos Film Preservation Center back in 1935. Sustaining a study center, repository, and circulating Film Library, the Celeste Bartos Film Preservation Center is home to one of the world's most important collections of film art. (MoMA 2015c).

The Museum of Modern Art serves a diverse audience of local, national, and international communities for which it strives to promote a deeper understanding, enjoyment, and appreciation for modern and contemporary art. Central to the MoMA's mission is to be a place that "ignites minds, and provides inspiration" as a venue that "is dedicated to the conversation between the past and present, the established, and the experimental"(MoMA 2015d). Through leadership of its Trustees and staff, MoMA realizes its mission by "establishing, preserving, and documenting a collection of the highest order that reflects the vitality, complexity, and unfolding patterns of modern and contemporary art; by presenting exhibitions and educational programs of unparalleled significance; by sustaining a library, archives, and conservation laboratory that are recognized as international centers of research; and by supporting scholarship and publications of preeminent intellectual merit" (MoMA 2015d).

As made evident by its mission, the Museum of Modern Art puts great emphasis on its prudent work in contemporary art collections, including conservation, research, documentation, and exhibition. These four activities are also central to the motivation to incorporate digital preservation strategies within the Museum. As of 2014, the MoMA was the first museum to create a standards-based digital preservation repository specifically for museum collections (Fino-Radin, Van Malssen, and Gillean 2014). In as such, the journey that MoMA pioneered to achieve its leadership in digital preservation within the museum profession, makes it an ideal case study for digital preservation planning and policy within the museum context.

MoMA's Relationship to Digital Technology

The Museum of Modern Art's commitment to digital assets is made evident by the sheer volume of museum collections that are housed on digital media carriers, are digitized, or are born-digital. These collections are spread between the museum, library, and even archive units of the institution; the Museum's overall relationship to digital technology will be discussed within the context of the Museum's art collection and involvement in Matters in Media Art; the Collections Online webpage; and the Museum Library and website relationship to the New York Art Consortium's web-archiving program.

With 25% of all MoMA artworks existing in some form of audio-visual (AV) format (around 35,000 pieces), the Museum's top priority as high-level, trusted stewards of contemporary art is to ensure that these artworks are preserved, discoverable, and accessible for long-term preservation (Arkivum 2015). Some of these artworks include the 500 hours of Andy Warhol's 16mm films, all of which will need to be digitized and then stored for long-term archiving, but also remain easily accessible if needed for exhibition (Arkivum 2015; Fino-Radin 2015). The dire need for solutions for conservation and preservation of the many audio-visual, software, and other time-based media artworks in the MoMA's holdings led it to become a participant in the Matters in Media Art consortium as noted in Chapter 7, in partnership between the Tate Modern in London and the San Francisco Museum of Modern Art. As members of the New Art Trust, MoMA was naturally incorporated into the meetings of Matters in Media Art in 2003 to work towards solutions in collection management and registration methods for time-based media art. Since the second phase of Matters in Media Art ended in 2008, MoMA has continued its commitment to the topic of long term preservation of ephemeral and digital artworks. For more information on Matters in Media Art, see Chapter 7: San Francisco Museum of Modern Art.

Details on the major digital preservation initiative occurring at the Museum of Modern Art will be discussed in the subsequent section in reference to their state-of-the-art digital repository (DRMC). MoMA has fostered a relationship to digital technology in other noteworthy ways including the many digital assets produced for the Museum's website, as well as those created by the Museum library and archive units. The Collections Online web page on the Museum's website features about 60,000 artworks from nearly 10,000 artists (MoMA 2015e). The many digitized images created of its collection contribute to the MoMA's mission and commitment to helping the public understand, enjoy, and use their collection by making it more accessible on the Web (MoMA 2015f).

The MoMA's Library manages a large number of records for its books, periodicals, exhibit catalogues, special collections, and electronic resources using DADABASE, an online catalog that supports the holdings of the Library and other MoMA study centers (MoMA 2015g). It lists scholarly materials located in MoMA's library in various mediums, however electronic or

digitized resources are accessible in DADABASE itself (MoMA 2015h). Records on primary source collections are reserved for the Museum Archives in a separate database. DADABASE feeds into a greater online catalog of art museum library holdings called Arcade, which is managed by the New York Art Resources Consortium (NYARC) (Moma 2015g). NYARC is a multi-museum library consortium working towards its own digital preservation efforts through a collaboration of its three partners the Museum of Modern Art Library, the Brooklyn Museum Library, and the Frick Collection Library.

Starting in 2006, with funding from The Andrew W. Mellon Foundation, NYARC was formed as a collaborative effort to enhance accessibility to art historical resources across a number of New York museum-libraries. As a consortium, NYARC aimed to create more cost-efficient and sustainable scholarly research programs, while also improving access and discovery of an ever-expanding number of resources through technology. In collaboration with the Online Computer Library Center (OCLC), NYARC created a 2008 and 2009 report on art museum-library access collaboration including additional institutions such as the Thomas J. Watson Library at the Metropolitan Museum of Art and Columbia University (Lavoie and Waibel 2008). In particular, one of the important ways MoMA's participation in NYARC deepens the Museum's relationship to digital technology is through NYARC's web-archiving project.

In Fall 2013, NYARC was awarded $340,000 from The Andrew W. Mellon Foundation to initiate a web archiving program to harvest the many online art historical resources made by New York Museums/Libraries. A study from 2012, "Reframing Collections for the Digital Age" demonstrated that the materials collected by the NYARC libraries/museums were increasingly migrating to online and digital versions, some exclusively available only on the web (NYARC 2015). The study concluded that there was an urgent need to document the web-based versions of valuable auction catalogues raisonnes, scholarly research projects, as well as artist, gallery and museum websites. Without such web archiving efforts, NYARC acknowledges a real and imminent danger of a "digital dark age" in the art historical record, very much analogous to the same 'digital dark age' feared by cultural memory institutions and the tech industry alike (NYARC 2015).

To support NYARC's web archiving efforts, a tool called ArchiveIt was deployed in 2013, a premier web-archiving service for collecting and assessing cultural heritage on the web. Archive-It was developed by the Internet Archive, and is a web page crawler that periodically archives certain web pages, stored as a WARC file. A primary and backup copy of archived web pages are stored at the Internet Archive data centers (Archive-it 2014). For MoMA, NYARC's subscription to Archive-It keeps versions of MoMA's websites including MoMA collection records (from Collections Online noted above), exhibition sites, MoMA PS1 site, Inside/Out blog, and the POST blog (Archive-It 2014b).

The Museum of Modern Art's collaboration with the New York Arts Resources Consortium (NYARC) led to a cutting-edge project that helps to protect data from the unstable, ethereal nature of the Internet. This digital preservation project, while separate from its in-house digital preservation efforts to be discussed below, is nonetheless an important factor in MoMA's leadership in the acknowledgement and practice in promoting digital preservation within the museum field. By collaborating with well-known digital preservation experts, such as the Internet Archive and NYARC, MoMA has found a way to sustainably archive some of its important digital assets -- in this case web-based resources. It is important to acknowledge the multi-faceted ways in which MoMA is engaged with digital technology, however the main driver for creating a trusted digital repository, and the first of its kind for the museum world, was by way of collection preservation.

Digital Preservation at MoMA

While there are many aspects of digital collections at the Museum of Modern Art between the museum, library, and archive units, the institution has devoted most of its innovation for digital preservation systems within the needs of its art collection. This section will report on the current status of digital preservation at MoMA within the context of its Digital Repository for Museum Collections (DRMC) Project. The content of this section was mostly drawn from a 2015 personal interview with Digital Repository Manager, Ben Fino-Radin, relevant 2015 posts from MoMA's Inside/Out blog, as well as interviews conducted by the Smithsonian Institution Time-Based and Digital Art Working Group in 2013.

The topic of digital preservation began at the Museum of Modern Art as a result of the first Matters in Media Art symposium in 2005. The conversations brought up through Matters in Media Art made evident that at the time art museums were ill-equipped to manage the longevity and accessibility of artworks stored on sensitive media carriers, or for collections that are born-digital. A conservation, and therefore digital preservation program would be needed as soon as possible - a process underway within many art museums today including those from Matters in Media Art (Fino-Radin 2015). The questions around time-based media and digital art stewardship led Jim Coddington, chief conservator, to realize MoMA needs to take action towards hiring a time-based media conservator to aid in resolving this new frontier of stewardship issues (Fino-Radin 2015). As a starting point Glenn Wharton was brought on in 2005 to perform a survey of the media collections at MoMA. By 2007 when the position of Media Conservator was created at the Museum, Wharton stayed on with this title to continue with the development of conservation for media artworks (Sanchez and Smith 2013).

Starting with MoMA's vast film collection, Wharton was able to determine shortly after the assessment that one of the first necessary steps would be digitization; many of the time-based media works at MoMA were still housed in their original media carriers many of which were not only at risk of technological obsolescence, but also slow deterioration of the media itself. Maintaining the artworks on the original media was no longer going to suffice. Due to MoMA's long history of acquiring video art, the collection contained a variety of formats, such as two-inch Quadraplex reels, hundreds of U-matic tapes, and thousands of VHS tapes among others (Oleksik 2015). Video tape was never designed to be a long-lasting medium; hence the entire collection was in danger of becoming unplayable because of degradation issues due to age and inherent fragility (Oleksik 2015). For example, a large concern with magnetic media is binder hydrolysis, or "sticky shed syndrome," in which the magnetic particles separate from the binder media over time. Along with a huge digitization project, intellectual control of each piece's format history would need to be documented; such information is often inseparable from the work's historical and artistic significance or meaning (Oleksik 2015).

In 2011, Wharton hired on Peter Oleksik as an Assistant Media Conservator to do much of the digitization of the video collection (Fino-Radin 2015). Previously, migrations to digital formats had largely been carried out on an ad hoc, exhibition-driven schedule. However with

Wharton and Oleksik's help, MoMA realized that a focused effort was now a top priority in order to care for the over 6,000 tapes in the collection (Oleksik 2015). Much of the digitization was done in-house, and some was also outsourced to specialized vendors when necessary (Fino-Radin 2015). In addition, metadata was imperative to the digitization process and the integrity of the artworks. The history of each digitized artwork was documented fastidiously based on documentation of the artist's practices, institutional knowledge, and general conservation knowledge about methods in which artists would work with video. In addition to metadata capture, special care was made to maintain the authenticity of each piece by analyzing its video signal after the data transfer, which ensured the digitized version remained as close as possible to the analog original and the intent of the artist (Oleksik 2015). As each analog video work in MoMA's collection was migrated to a digital video file, no compression was applied to allow for the maximum digital latitude in accurately representing the analog video signal (Oleksik 2015). The resulting digitized versions of the works now allow the Museum to extend the life of the works and to ensure that the artist's' work can continue to contribute to the art historical canon through accurate and faithful exhibition in the future. As said by Peter Oleksik, "Now that the material is in digital form, there is no risk of damaging tapes upon playback, no lengthy wait time after requesting material from off-site storage, and files can be easily transcoded to copies for viewing. This has allowed unprecedented access to this historically significant collection"(Oleksik 2015).

Of course, now that much of MoMA's media collection was digitized, methods for long-term storage and long-term viability were required for the resulting digital objects. Digitization saved many of these analog video artworks, but also marked the beginning of a new set of challenges and risks that are unique to digital objects. The unique nature, qualities, and risks to digital objects are outlined in Chapter 2 of this thesis. Starting in 2010, Glenn Wharton worked with a collaborative team made up of leadership from IT, conservation, collections, exhibitions, curatorial, and outside experts to carefully formulate MoMA's needs for stewardship of digital collections, and what functional requirements would meet those needs (Sanchez and Smith 2013b). With help from Karen van Malssen of AV Preserve, Wharton

authored a document fully articulating these needs (Sanchez and Smith 2013b). This event is perhaps the true starting point for the beginning of digital preservation at MoMA.

In 2013, Ben Fino-Radin, who had been working for the Rhizome ArtBase of the New Museum at the time, was approached by Glenn Wharton to assist with the development and management of MoMA's up and coming digital repository project. The vision was to create a true repository that would essentially be akin to traditional art storage, except specifically only for digital objects including software-based works, digitized films, video games, and other forms of digitized audio/visual art (Fino-Radin 2015). Through Wharton's analysis of MoMA's needs, it was determined that having the works off their media, digitized, and on centralized storage in two locations (which was MoMA's initial setup) was not quite enough to make sure the works were fully managed and that they were preserved properly (Fino-Radin 2015). From his experience with preserving digital artworks at Rhizome, Fino-Radin believed that most of MoMA's repository functional requirements could be accomplished with an open-source digital preservation tool called Archivematica. The development of MoMA's digital repository began with Archivematica in 2013, and since has deployed a full sweep of systems in production since Fall 2014 (Fino-Radin 2015). Today, in late 2015, MoMA's state-of-the-art digital vault includes three technical parts that can be broken down as such: the packager- ***Archivematica*** as the ingest pipeline; the warehouse ***-Arkivum*** for digital storage system; and the indexer -an MoMa-collaborated software tool called ***Binder*** made in collaboration with Artefactual Systems (Fino-Radin 2015; Fino-Radin 2015b). See the infographic below from a blog post written by Ben Fino-Radin for MoMA's Inside/Out blog:

Figure 4, from *Inside/Out* blog posted by Ben Fino-Radin April 14, 2015

An inherent challenge of digital materials is that all digital files are encoded and thus require special tools to interpret the code in order to be understood as something other than a series of 1's and 0's which cannot be understood by humans. It is worth restating the challenging nature of digital objects here because these are the same issues that were very present within the minds of MoMA's digital repository team. Similar to how a VHS tape is useless without a VCR, a digital file is useless without the right combination of software (and sometimes hardware) that understands how to render it, or tell the user about its contents. Just by looking at a file, for example a Quicktime .MOV file, one cannot just tell what kind of software is needed to view it. This is especially true when considering future generations; we cannot guarantee that they will understand how to render digital objects of our time without the proper roadmaps, technical history, nor without some kind of assurance process that makes sure the object's bits are not corrupted or changed in any way over time. Knowing that the specialized tools we rely on to interpret digital objects -- be it an operating system, software application, or something very specialized -- will not always be around, we may also not understand all the formats that we do today. Even if we manage to maintain a perfect copy of a digital object for 150 years, no one may be able to understand what that file is, let alone what to do with it.

Archivematica, as 'the packager', addresses this fundamental challenge upfront as digital objects are prepped for the repository. Developed by Artefactual Systems in 2009, Archivematica follows the Open Archival Information System (OAIS) model for metadata harvesting and ingest quality control to create preservation-ready, platform independent information packages. According to its website, Archivematica is a "web and standards-based application that allows institutions to preserve long-term access to trustworthy, authentic, reliable digital content" (Archivematica 2015). Archivematica includes a series of micro-services as an integrated suite of software tools that allows users to process digital objects from ingest to access in compliance with the ISO-OAIS functional model (Archivematica 2015). Some of the micro-services perform granular processing tasks such as virus checking, checksum verification, file format conversions, etc (Owens 2012). Other standards used in Archivematica includes METS, PREMIS, Dublin Core, the Library of Congress BagIt specification, among others to provide trustworthy, authentic, reliable, and

interoperable archival packages (AIPs) for storage into a repository setup. It also provides a web-based dashboard from which users can monitor and control the processing workflows (Archivematica 2015). In other words, MoMA can use Archivematica to analyze all the digital collection materials as they arrive to the repository, and use it to record the results in an obsolescence-proof text format that is packaged and stored with the information object itself. This makes up the Archival Information Package (Fino-Radin 2015b). The preservation roadmapping and preparation activity conducted by Archivematica makes it analogous to the concept of "the packager." See the infographic below from MoMA's blog, Inside/Out that exemplifies a simplified visual for how Archivematica works:

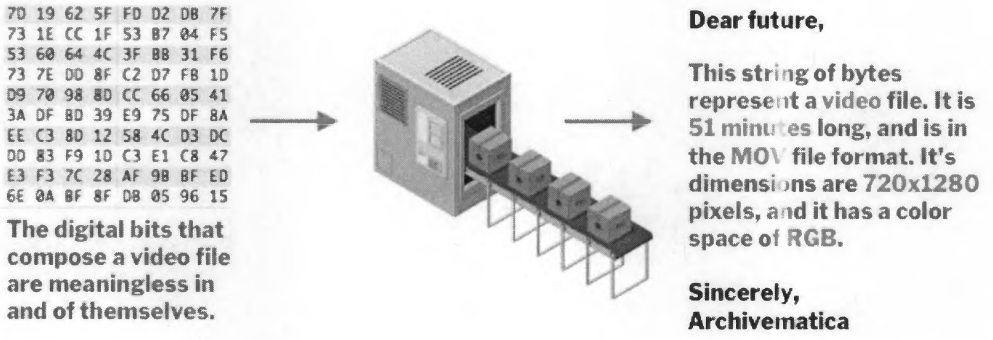


Figure 5, from *Inside/Out* blog posted by Ben Fino-Radin, April 14, 2015

In addition to mitigating the issue of ensuring that our successors will be able to understand how to render a digital object, Archivematica also addresses the problem of authenticity, which is critical when dealing with original art collections (Fino-Radin 2015). Because bit streams are very particularly ordered bits and bytes, if any one of those components is modified, maliciously or not, the digital object's original function or visual integrity could be compromised. Archivematica addresses this issue by passing each and every digital object through a cryptographic algorithm, or checksum. The subsequent checksum value for a digital file is recorded, and allows MoMA to run the object through this value over and over to make

sure none of the bitstream has been altered (Fino-Radin 2015). Having the ability to run integrity checks at any time is critical when dealing with an art medium that is inherently reproducible - integrity of 'the original' artwork is forever rethought when dealing with digital formats. In addition, MoMA uses a standalone disk imaging program called FTK Imager, from Forensics Toolkit by AccessData, which saves an image of a hard disk, calculates MD5 hash checksum values, and confirms integrity of the data before closing the files (Fino-Radin 2015). FTK Imager is popularly used by digital preservationists, and allows an image file to be saved in several formats. Such authenticity verification provided by Archivematica and FTK Imager allows the archival packages created at MoMA to not only preserve the digital collections, but also to hold the information the museum needs to confirm at anytime the authenticity of its digital collections (Fino-Radin 2015).

The next component of MoMA's digital repository is the storage system to which the archival packages are sent off to. The 'warehouse' infrastructure vendor used by MoMA is called ***Arkivum***, which is maintained by MoMA's IT department (Fino-Radin 2015b). Arkivum's storage system at MoMA was only recently deployed as of mid-2015; the first 'warehouse' component of MoMA's repository was purchased back in 2010 during the major film digitization initiative led by Wharton and Oleksik (Fino-Radin 2015). These subsequent sections about Arkivum are very technical, and while perhaps difficult to understand, nonetheless is a useful account of the methods deployed by one of the most advanced trusted digital repositories in a museum to date.

The digital storage infrastructure can be thought of as analogous to physical museum art-storage facilities, such as MoMA QNS, the Museum's offsite storage facility in Long Island City (Fino-Radin 2015b). The first digital storage used by MoMA since 2010 was a very large cluster of hard drives configured as a Redundant Array of Independent Disks, or RAID, that lives in a data center at the Museum. A duplicate of the entire cluster also lived off site at MoMA QNS (Fino-Radin, 2015). This set up served MoMA well for the last five years, but it was found that this type of disk-based storage became an untenable expense with very large amounts of data (Fino-Radin 2015b). Since MoMA's current digital collection is upwards 80 terabytes in size (80,000 gigabytes) and growing, this large amount of data is estimated to grow

over the next 10 years to approximately 1.2 petabytes (1.2 million gigabytes) as the Museum acquires more digital artworks. Knowing the rate of growth of MoMA's collection, it was decided that it would be irresponsibly expensive to continue to use this kind of spinning disk storage (Fino-Radin 2015b).

By mid-2015, MoMA was in the final stages of designing a completely new "warehouse" with a company called Arkivum (Fino-Radin 2015). This new system is a hierarchical storage that includes a small cluster of hard drives, but also adds the new element of data tapes (IBM LTO [Linear Open-Tape] magnetic tape) for primary long-term storage (Fino-Radin 2015). When the archival packages are first stored, they are placed on the cluster of disk storage (local hard drive cache), but are shortly thereafter copied to the data tape library. Hierarchical storage management is a data storage technique that automatically moves data between high-cost (high-speed hard disk drive arrays) and low-cost storage media (magnetic tape drives). While it would be ideal to have all data available on high-speed devices all the time, this is prohibitively expensive for many organizations. Instead, hierarchical storage systems, like that designed by Arkivum, will store the bulk of data on slower devices, then copy data to faster disk drives when needed. In effect, such a system turns the fast disk drives into caches for the slower mass storage devices (Dillon and Leonard 1998, 126-7). The hierarchical storage system maintained by Arkivum will allow MoMA to store the projected 1.2 million gigabytes of digital collections material redundantly in three locations: the Museum data center, the off site art storage facility in Long Island City, and the Celeste Bartos Film Preservation Center in Hamlin, Pennsylvania (Fino-Radin 2015). The added third copy places MoMA in compliance with the digital preservations standards recommended by the Library of Congress' National Digital Stewardship Alliance (NDSA 2015).

The first two parts of MoMA's digital repository, Archivematica and Arkivum, work together to facilitate digital preservation designed for the long-term. This part of the digital vault can be likened to carving information into stone in a universal language, then storing it an underground vault (Fino-Radin 2015b). However neither Archivematica or Arkivum facilitates day-to-day, active management, access, and big-picture analysis of the contents of the "warehouse" (Fino-Radin 2015). Museum professionals, especially those of collection

management, curatorial, and conservation units needs to understand the contents of the "warehouse" at all times in order to practice good stewardship.

Fino-Radin, along with the rest of the digital repository team, collaborated with experts from the private sector, libraries, archives, and other museums (including our other case study SFMoMA) to see what systems others may be using to address the need of a "portal" into the repository. The systems discovered in their findings did not completely meet the digital preservation requirements of media conservators working with museum collections; if MoMA was going to invest in a tool, they wanted it to work as optimally as possible and there was a gap between what they needed and what was available (Fino-Radin 2015). So they decided to build their own system, which has resulted in the third part of MoMA's repository software stack: ***Binder*** (Fino-Radin 2015).

Binder was developed by MoMA and Artefactual Systems in 2014 as a Web application designed to oversee and manage the active preservation of MoMA's digital collections (Artefactual Systems 2015). Binder is integrated with MoMA's custom branch of Archivematica and AtoM (an application for standards-based archival description and access), and is also directly linked to MoMA's collection records management system, The Museum System (TMS) (Artefactual Systems 2015). Essentially Binder provides a central user-interface through which users can access, view, and manage the rich technical metadata extracted from Archivematica for Archival Information Packages, but also see the full object record from TMS. Although the standardized metadata in AIPs formed by Archivematica are in a format that does not require special tools or techniques to understand it, the human-machine readable format makes it difficult to run quick and effective analyses about the entire collection (Fino-Radin 2015b). What good is the preservation of digital objects if the repository managers and other stakeholders cannot access the collection information easily? This is where Binder steps in as a key tool for indexing, analyzing, and seeing into the collection. It does so by managing and describing the relationships between the components of a collection object, its constituent digital objects, and the various external dependencies required to preserve and display the collection over the long-term (Artefactual Systems 2015). Before the archival information packages are

sent to the Arkivum storage system, Binder sifts through them, indexes their contents, and stores what it finds in a database that is designed to be good at large dataset queries (Fino-Radin 2015b). Therefore Binder allows the repository team, conservators, and curators to see the bigger picture of the collection within the repository at any given time.

In addition, Binder provides an outlet for MoMA's staff to understand the smaller details about the pieces stored within the digital repository (Fino-Radin 2015). Since Binder supports standards-based repository management, it provides a single place to view the four kinds of metadata related to objects: administrative, technical, descriptive, and preservation metadata (Artefactual Systems 2015). It also recognizes relationships between the various data of any given object that the user may not realize upfront. This in turn enables repository managers the key information and analyses they need to craft appropriate preservation policies and implement decisions for long-term care. A brief description of MoMA's digital preservation policy will be discussed below.

Binder's widget-based dashboard includes many useful features that helps MoMA's conservation and curatorial teams to understand the past technical life of an artwork, as well as the future technological variability potential of an artwork (Artefactual Systems 2015). The particular technological dependency that any digital artwork possesses is described and drawn on a visual graph-based context browser to aid in preserving the dependency itself. From Binder's user manual, one can achieve the following tasks, reports, and access when using Binder:

- "Import AIPs and reference copies of digital objects from Archivematica, and relate them to descriptive metadata imported from TMS or created in Binder.
- Gain at-a-glance collection-wide statistics about fixity, ingest, and use via the widget-based dashboard.
- Relate the components [like required software or hardware] of a work to derived AIPs and any supporting technologies required to preserve and display them in the future, using a node-based graphical user interface.
- View and download an AIP's digital objects and technical metadata.

- Sort search and browse results based on facets drawn from both descriptive and technical metadata, allowing for a high degree of precision and granularity - and then save your search parameters for future re-use.
- Run and manage fixity checks of preserved AIPs, and receive alerts if a fixity check fails.
- Track who downloads digital objects from your repository, and why.
- Compare the descriptive and technical metadata of up to 4 digital objects from an AIP side by side in a graphical user interface.
- Generate and save reports on ingest, fixity, usage, and more." (Artefactual Systems 2015).

Lastly, Binder has a 'digital object viewer' feature that allows the user to view an actual proxy version of a digital object in the repository (Fino-Radin 2015). This feature not only gives the user a way to see a glimpse of the actual object, it also includes technical metadata extracted from the METS file generated by Archivematica, a copy of the Dissemination Information Package (DIP) for easy visual reference, and the ability to download files directly from the Binder web-API (Artefactual Systems 2015).

After using Binder for some months at MoMA, both Artefactual and MoMA realized they can expand the utility of the project by open sourcing its code and making it available to other developers. MoMA's hope was that providing a free version of Binder could help a broad number of cultural heritage institutions to achieve their long-term preservation goals; they are excited to see the Binder project develop into its own full-fledged, production-ready, open-source application with its own vibrant museum community (Artefactual Systems 2015b). In late 2014 and early 2015, MoMA took initial steps to generalize and open-source the code for Binder. By May of 2015, MoMA officially released the free, open-source software code for Binder on GitHub, including documentation of its features, technologies, and API on ReadtheDocs, a website that hosts documentation for the open-source community (Fino-Radin 2015b). Now any cultural heritage institution who wants to achieve digital preservation, especially for the museum context, can download, adopt, modify, or redistribute Binder for their own use, and all for free (Fino-Radin 2015).

Standards Overview and Policy

It is apparent that the Museum of Modern Art's Media Art staff has worked very hard since 2010 to develop a system and workflow for digital preservation in the museum context, not

only for themselves but also for the greater museum community. The Museum can feel confident about the best practices it sets forth due to much research from allied fields in library and information science. The MoMA has used the Open Archival Information System (OAIS) by default of choosing Archivematica for their packaging pipeline software. Other standards borrowed from allied fields used in MoMA's digital preservation system are focused on metadata, TRAC/ISO 16363, and the NDSA Levels of Digital Preservation (Fino-Radin 2015), all of which are reflected within MoMA's Digital Preservation Policy.

For metadata of MoMA's digitized and born-digital artworks, MoMA hired Peggy Griesinger, who was a National Digital Stewardship Resident, to conduct a project that assesses the metadata standards MoMA needs to preserve the technical history of media works (Fino-Radin 2015). Griesinger's position was part of an Institute of Museum and Library Services (IMLS) grant-funded program that places recent graduates from library and archive programs in prestigious cultural heritage institutions across New York City to help those institutions find solutions for digital preservation problems (Griesinger 2015). Griesinger's project was to research standards for describing the digitization process history (e.g. how it became a digital file, what was its original format, what equipment was used to digitize it, etc.) (Fino-Radin, 2015). Before this metadata project, the information about an object's technical history was recorded by museum staff in unstructured text and/or on proprietary formats such as Microsoft Word and Adobe Acrobat, which is not sustainable for long-term archiving (Griesinger 2015). While these formats are common today, we must always remember that 20 years from now, Microsoft may not support a version of Word from 2015. Since her project ended in May 2015, Griesinger has developed an XML profile for MoMA that combines quite a few metadata standards since there was no one standard that met the Museum's needs. In the end the metadata standard constructed was a combination of a METS wrapper/file with elements of PREMIS, PBCORE, and REVTMD (a very little-known metadata standard that has been used previously to describe digitization history) (Fino-Radin 2015). XML was the chosen file format because it is a non-proprietary, text-based format that allows information to be encoded in a way that is both human and machine-readable. In addition, since MoMA now has Binder, it made sense to them to create specifications to build the functionality to create this XML-mapped

metadata standard in Binder, that way this particular set of metadata can be recorded and viewed directly in Binder. MoMA hopes to have the funds to integrate this metadata functionality in about a year (Fino-Radin 2015).

Museum collection management and conservation is very different from libraries or archives in that they do not dictate the formats they receive from artists; the format is received based on the artist's process and therefore the Museum does not want to change it if possible (Fino-Radin 2015). While it is most desirable to maintain the native masterformat, MoMA's Media team will sometimes be able to request a different format that is easier to manage but is analogous to the native format (Fino-Radin 2015). Standard formats used at MoMA include Quicktime .MOV for video files, TIFF for still images, .WAV for audio files, DPX for digitized film, and raw disk images (made by FTK Imager). Much consideration is currently being made in MoMA's digital repository team to use forensic disk images as well, simply because it has been reversed engineered, provides better metadata for preservation, and is heavily adopted (Fino-Radin 2015). Software art never has a master format, but MoMA does collect the source code from the artist which is included in their collection policy for digital artworks (Fino-Radin 2015).

The Museum of Modern Art's aim to have a repository that meets the collection's functional needs for long-term preservation has resulted in the byproduct of mapping many of its digital repository elements to the Trusted Digital Repository Audit Certification Checklist (TRAC, now an ISO standard [ISO 16363]). Meeting the ISO 16363 requirements as a standard, and therefore specifically for certification as a TDR (Trusted Digital Repository) is not MoMA's main goal because certification goes beyond the Museum's purposes; but they are using it as a guiding tool for self-assessment (Fino-Radin 2015). Using the ISO standard as a self-assessment tool allows the Museum to check if its practices parallel that of a Trusted Digital Repository, and therefore to verify the integrity, security, and longevity of MoMA's digital collections. MoMA specifically did an audit with guidance from AV Preserve in which they combined TRAC and the NDSA Levels of Digital Preservation, along with some of their own criteria, to audit MoMA's digital preservation and generate a report that they can present to stakeholders and report on their progress (Fino-Radin 2015).

The MoMA's approach to following standards only to the extent that they are relevant to the Museum's actual on-the-ground needs has been a consistent approach during the development of their digital repository. For example, one of the interview questions from this thesis research methods addressed the best practice of 'normalization' of file formats upon ingest; this is something Fino-Radin would not advise, not just in the museum context, but in general. Unlike libraries and archives that work with digital objects in mass digitization formats or of a more homogenous nature, museums must respect certain aspects of 'the original' which is part of the historical significance of each piece (Fino-Radin 2015). This perspective favors digital preservation strategies of metadata encapsulation and emulation as preferred solutions to combat format obsolescence over 'normalization' (Fino-Radin 2015). The strategy can even be likened to the more homogenous collections of libraries and archives, predicated on the idea that the original format of even mass-produced digital materials can reveal historical significance. The format used to write an original work of American literature in 2015 may provide important historical information about a writer's process of our present time. In order to not lose the contextual information of a digital object, thorough metadata tracking, and emulation whenever possible, best maintains the trusted integrity of the object. As a museum, and perhaps an odd-man out in the digital preservation world, MoMa adopts best practices from the digital preservation community, but only after thoroughly evaluating what standards the Museum truly needs; the Museum favors using and customizing standards and best practices to their specific needs. As concisely said by Fino-Radin in a 2013 interview with the Smithsonian Time-Based and Digital Art Working Group: "In practice, carefully informed action that is sensitive to the needs of your institution is more important than meeting a best practice" (Sanchez and Smith 2013b, 12).

The perspectives on adopting practices from outside fields and adapting them for the museum context is also reflected on MoMA's use of digital preservation policy. The digital repository team has created its own protocols, plan, and workflow for the object-level steps of using the digital repository, but the Museum has overall adopted its very own formalized digital preservation policy. The MoMA's digital preservation policy is not published openly for the public to view, however a copy was provided for this thesis research. The fact that MoMA even

has a high-level, institution-wide adopted policy is particularly special, especially considering only two years ago (2013), Madeline Sheldon's study from the Library of Congress confirmed that only two museums worldwide had published digital preservation policies.

Following in true fashion of other types of museum policy recommended by the American Association of Museums, MoMA's digital preservation policy is maintained as a high level formal document (Fino-Radin 2015). From the policy itself, the main purpose is outlined: "The purpose of this policy is to document the principles, standards, and practices that guide the care and preservation of the Museum of Modern Art's digital collections. This policy is not intended to be a handbook or operating manual, but provides a comprehensive framework for decision making and for the development of digital preservation procedures at MoMA"(Fino-Radin 2015c). For MoMA, policy best works as a blanket formalization that touches on the legal, ethical, and very basic preservation issues, leaving the more specific details for the digital repository managers to handle internally.

MoMA's digital preservation policy also outlines important vision-focused statements such as the Museum's mandate for digital preservation. As quoted from the policy itself: "The DRMC's mandate for the preservation of digital collections is drawn fundamentally from MoMA's Collection Management Policy, which documents the museum's overall commitment to the care of the collections through conservation, proper environmental conditions, security, and proper handling. The role of the DRMC is to enable the realization of this mandate for all *digital* collections materials" (Fino-Radin 2015c). This policy mandate is a very important statement on its own that contextualizes digital materials as worthy of the *same* ethical handling and collection management that is normally outlined in a museum collection management policy for traditional objets. The policy outlines the 'digital collections materials' to include not only the artworks themselves, but any supporting documentation from the past or future that will affect the future exhibition and conservation of the artwork (Fino-Radin 2015c). By aligning the handling of digital artworks and other digital collection materials with the same *duty of care* that Marie Malaro calls out as an ethical obligation made accountable through policy, MoMA exhibits a mature understanding of committed stewardship to digital objects. The deeper

implications of MoMA's digital preservation policy mandate encapsulates the points made in Chapter 5: Digital Preservation Policy, the New Collection Management Policy.

MoMA formulated its digital preservation policy not just from the ethical and legal topics of the institution's collection management policy, but also from a series of digital preservation policies from outside the Museum. Greatly inspired by the SCAPE (Scalable Preservation Environments) project's *Catalogue of Policy Elements* (2014) from Europe, MoMA's digital repository team was able to reference fourteen different digital preservation policies that helped the Museum to develop the correct content and depth for MoMA's own policy (Fino-Radin 2015). Some notable policies used as models for MoMA's policy are the National Museum of Australia's 2012 Digital Preservation and Digitization policy (which is the very first *museum* digital preservation policy), Cornell University Library's 2004 Digital Preservation Policy, and the UK National Archives Preservation Policy from 2009 (Fino-Radin 2015c). Other sections within MoMA's digital preservation policy include: guiding principles, staff roles and responsibilities, standards, selection and acquisition, overview of preservation strategies (such as storage, metadata, and bit preservation), security, disaster recovery, and access/use (Fino-Radin 2015c).

With its standards-based digital repository up and running and a digital preservation policy in place, MoMA's future in digital preservation is focused on optimization;. The DRMC team is looking to streamline the entire digital preservation process to make it faster and easier for the staff to process (Fino-Radin 2015). Regarding improved access and exhibition of its digital collection, the DRMC team is also looking to potentially integrate with the bwFLA Project. This project makes the software architecture for 'Emulation as a Service'- which aims to provide ready-made, easy-to-use emulation services that are scalable and affordable. The hope is that Binder could be integrated with bwFLA's Emulation-as-a-Service to simplify access to preserved digital assets by allowing end users to interact with the original environments running on different emulators from a server to a web browser (Fino-Radin 2015). The future of digital preservation at MoMA will only continue to grow and improve as the institution maintains their commitment to collaboration with the digital preservation community and through sharing its resources with the greater museum field.

Analysis

The analysis of the Museum of Modern Art as a digital preservation case study will be made within the context of digital preservation motivations, maturity of program, approach to standards, and contributions to the museum field.

Motivations for Digital Preservation

The Museum of Modern Art, along with its fellow cohorts in Matters in Media Art, have a very clear directive for why digital preservation is an important activity in museum collections. The threats to their digitized or born-digital objects, records, and other corresponding materials that affect long-term collection management is well understood at MoMA. While many museums today will still be grappling with the very definition of digital preservation and the basic need for it in their institutions, MoMA has assembled a critical team of people and collaborators that understands digital preservation on a deeper level, allowing for the forthcoming work that they have done. Glenn Wharton certainly had the foresight to see that assembling a team with cross-disciplinary skills (computer science, library science, conservation, etc) and seeking advice from outside the museum field was the key to making digital preservation feasible for MoMA.

The motivation to gather a team of people with various skills relating to digital preservation was ultimately focused on the Museum's unique collections, as opposed to the preservation of other digital assets collected at the museum within its library, archive, and digital asset management team (images for web, etc). Due to the ethical and legal stewardship obligation of any museum, *collections* tend to be a top priority for museums, especially when threats to the collection's integrity are at hand. However, importantly, the MoMA library is not left out from the scope of digital preservation; its involvement with the New York Art Resources Consortium, which is handling the issue of long-term digital preservation of online resources, albeit separately from the museum, demonstrates the library unit's own commitment to working in digital preservation.

While conservation and collection management is the main driver for the state-of-the-art digital repository at MoMA, ultimately it seems that the DRMC team will slowly integrate more and more features that enhance the repository's capabilities. For example, the Museum's collection management system, TMS (The Museum System) is already integrated with Binder; therefore the descriptive metadata created by the Museum's collection management team is maintained alongside the artworks in the repository. One would not be surprised if, one day in the future, additional resources are also linked into the Binder-Arkivum system; such as library or archive research materials relevant to specific digital artworks. While it is true that the MoMA has created the most impressive, standards-based, trusted digital repository in the museum field to date, because its creation revolves only around artworks, the big picture question for the future of digital preservation at MoMA is - Will the museum eventually commit to extending their digital preservation best practices and system to other museum materials? Regardless of department, museums are creating and receiving a multitude of valuable digital assets that we cannot allow to become lost or deteriorate without expensive and intellectual consequences. The analysis for whether digital preservation efforts will be extended across the whole museum is a positive conclusion. Although it was not discussed in this chapter, MoMA ensures the bit-level preservation needed for basic safeguarding of its library and archive materials, but those materials currently are not included in the DRMC (Sanchez and Smith 2013). As a way to optimize the functionality of the DRMC across more museum departments, finding a way to preserve additional documents, research, and archives alongside the digital artworks in the repository could be a feasible way for MoMA to extend its policy mandate across the greater museum.

Digital Preservation Maturity

Since digital preservation is a comparatively new discipline, models for good practice, including technologies and services, therefore exist at varying levels of maturity. The development of any new capability within an organization will often follow a path that begins with developing awareness of the need for that capability (and the steps required to acquire it), and ends with the realization of that capability, which potentially may vary at level of

sophistication. Levels of maturity can range from achievement of minimum standards to best practice. The narrative of MoMA's digital repository certainly follows this path, and the end result is an institutionalized, mature digital preservation system.

The maturity of the Museum of Modern Art's digital preservation system has certainly exceeded that of basic or minimum process, and is moving upward from a managed process to an optimized process, which is arguably the most mature level any institution can aim for. There are many models for achieving the realization of digital preservation in an institution, and MoMA can be said to adopt a hybrid model of developing a bespoke solution by using open-source software, using outsourced services, and also developing their own tools. A hybrid approach to creating a digital repository, while full of overwhelming choices, can offer the most opportunity to create a cost-effective, long-term solution that is specifically tailored to the museum's needs. By breaking the digital repository down into a three-part system, as opposed to trying to find technology that 'does it all', MoMA was able to deeply evaluate exactly what its digital collection needed for long-term preservation in isolated, focused steps. The technology used for the DRMC was continuously considered as a specific tool to achieve the museum's needs. Such mature assessment during the development of MoMA's repository can be argued to be the result of its staff having a deep understanding of both the minute steps required to preserve digital objects, but also the big picture analysis of what it truly means to be a trusted digital repository.

Within the context of Nancy McGovern's Five Organizational Stages of Digital Preservation, the MoMA's DRMC project would classify it within the fourth stage of maturity, in which institutionalizing of policies, procedures, and techniques creates a robust program that can be rationally managed and scaled, as needs demand (Kenney and McGovern 2003). Within the this maturity model, McGovern also quotes certain foundational documents that are put into consideration for a mature digital preservation program including the OCLC-RLG's *Trusted Digital Repositories: Attributes and Responsibilities* for comprehensive organization requirements, the Open Archival Information System Reference Model for digital archive requirements and object-level digital preservation requirements. MoMA has thoroughly demonstrated its full implementation of these standards. Other attributes that indicate the

mature-status of MoMA's digital preservation are evident in their use of high-level policy, technological infrastructure, and the Museum's achievement in providing optimized access and use of its archived digital collections.

Approach to Standards

This thesis discusses the concepts of standards and best practices a lot, however MoMA interestingly adopts such standards only to the extent that they meet the Museum's context, use, and needs. For example, on the topic of normalization of formats, although is often recommended as a best practice, Fino-Radin advises against normalization when possible. A quote from Ben Fino-Radin best exemplifies the un-rigid strategy towards standards:

> *"Simply put- guidelines are the base level of what to do, best practices as the vetted ideal way of doing it, and standards as the agreed upon and interoperable way of sharing it. The first and most important thing is to understand what your institution needs. What are the problems your collection faces, and what can you do within your means to fix them. I can't tell you how many institutions I see blindly trying to adopt or invent some standard, thinking it will solve their problems, yet not being able to demonstrate why they really need it. In practice, carefully informed action that is sensitive to the needs of your institution is more important than meeting a best practice"* (Sanchez and Smith 2013b, 12).

There is a large amount of truth to this statement; it is one thing to know about the standards practiced among memory institutions, but is is another thing to have a mature, fully-formed understanding of the greater implications of standards. Therefore, following MoMA's model, it is important to know how to use standards as a tool to solve your institution's problems, and not use them as a blanket, or passive, solution.

Standards and best practices are perhaps the only way any professional field can pay-it-forward with tangible solutions and guidelines for emerging institutions to follow; however MoMA's more liberal approach to implementing standards is a worthwhile model to pay some attention to. As viewed by MoMA's staff, formal standards coming out of related fields can be helpful for addressing parts of the preservation puzzle, but there are doubts around any narrow use of formal standards for the museum context. MoMA's staff has had the foresight to see the need for standards, but since many of these standards come from the library field, there is no one universal solution when considering the wide variability of museum collections and the rapid

evolution of the underlying technologies needed to render them. To argue the other side, when dealing with some major digitization projects, for example digitizing oral histories or film slides, a stricter approach to standards may be applicable for such digital collections. Regardless, museums collect a wide variety of materials as an inherent part of its institution-type; the types of digital assets will be less homogenous compared to allied memory institutions like libraries. MoMA's approach to standards fits well within the broad scope of digital preservation best practices and allows the museum to legitimize its preservation processes, however the staff is continuously cautious to assess how to tailor those standards to the unique needs of museum collections.

Continuing on the topic of MoMA's approach to standards, policy is one way in which the Museum exemplifies a more direct following of best practices from allied fields. By using high-level policy as a way to hold the institution accountable for digital preservation mandates, MoMA supports the concept that policies should be governing documents that address the same legal and ethical issues outlined in museum collection management policies. In many ways, the vision-based statements in MoMA's digital preservation policy not only follow the best practices/models of policies from allied fields, but also mimics closely the format recommended by the greater professional museum field. In Chapter 5 of this thesis, the compiled research calls for equal consideration for policy around the care, storage, and documentation of digital collections that museums already apply within its collection management policies. MoMA as a case study, brings this point to life with the implementation of its digital preservation policy.

In addition, it is interesting to note that the Museum did not integrate the issue of digital preservation as a sub-section of its overall collection management policy, but rather made it a separate document. On the one hand, this means the digital preservation policy is not available for public viewing, so few institutions or members of the public would even be aware of its existence. However, having a separate policy for the digital repository allows the Museum to have the text-space to freshly address a detailed account of its mandates, goals, staff roles, and strategies for digital preservation.

Contributions to the Museum Field

The final part of this analysis regards MoMA's contributions to the greater museum field with its good work in digital preservation. MoMA has already exemplified its eagerness to share its tools and strategies for digital preservation openly by making its home-brewed software Binder, free and open-source to the world. Considering this fact, it can be concluded that MoMA is certainly capable of continuing its leadership by extending mechanisms for shared research, lessons learned, and practices for the future museum community. Matters in Media Art encapsulates one example of success in creating a framework for sharing knowledge and developing standards that can be widely applied within the museum-specific context. However, MoMA's approach to open-sourcing tools is not as commonly seen within the museum field; in fact this is a practice much more deeply embedded in that of the tech industry.

The MoMA undeniably committed a large amount of time, staff, and money investment in the creation of Binder because it recognized a need, and had the means to do so. To MoMA, the investment was largely worth to be able to also provide this tool as a free resource to other museums or cultural institutions. Open-source software has historically fostered a broad community of sharing, discussion, and independent product improvements. As an open-source tool, MoMA has created the opportunity for software developers around the world to take the software code for Binder, change it, enhance it, and discover any bugs or challenges; this ultimately encourages community feedback and collaborative discussion on how to use and customize MoMA's product to a wider audience.

The Museum clearly wants to partake in the opportunities posed by the knowledge-sharing environment that has worked for the open-source community for many years. This very same attitude in of itself is a contribution to the museum community, and one that will hopefully take flight amongst other institutions over time. Without the eagerness to foster collaboration and cross-communication, the follow-through of digital preservation will not be feasible in many museums, especially those who do not have the same budget and staff size of the Museum of Modern Art. Digital preservation will be difficult for any museum to achieve alone. However, following the same attitude of openness demonstrated by MoMA's open-source software, making tools necessary for implementation openly available, along with a community

to help along the way, poses hope that accessibility to digital preservation will become possible for more museums in need. Although perhaps only a first scratch on the surface regarding the potential for digital preservation collaboration, MoMA's contribution to the greater museum field is encapsulated in the implications set forth by open-sourcing Binder. MoMA's contributions are also closely tied to its pioneering staff, and leadership in the implementation of the Museum's state-of-the art, standards-based digital repository that can be viewed as a model case study for the greater museum field.

Conclusion

The solution to MoMA's digital preservation system was found by achieving significantly higher levels of end-to-end digital preservation best practices, such as those defined by the US National Digital Stewardship Alliance. With research and integrated knowledge from allied fields in mind, a system was designed around the already available Archivematica file-format preservation system, which could be integrated with the Arkivum data archiving service. This was all then married to an indexing tool that MoMA designed and financed, culminating as the open-source API, Binder; which is now not only available for MoMA, but made free and open to all museums seeking solutions to digital preservation. Having reached the highest levels of the National Digital Stewardship Alliance best practice standards including the implementation of policy, MoMA has raised the bar for practical implementation and standards for digital preservation in the museum field. As concisely stated by MoMA's Digital Repository Manager, Ben Fino-Radin: "I think that many institutions have not truly come to face the facts that they must act ***now*** when it comes to the preservation of born-digital or moving image materials in their collections...We aren't talking about 'someday this will be gone' anymore -- we're talking about 'this will be gone tomorrow' if you don't do something today" (Sanchez and Smith 2013b, 12).

Chapter 10: Discussion

Based on the case studies and literature review presented in this thesis, four key themes relating to digital preservation efforts in museums will be outlined and discussed below. The four key themes are: defining digital preservation; integration of digital preservation technology; collaboration; and policy development. The chapter will conclude by identifying an important challenge facing museums in their digital preservation efforts, one that the research conducted for this thesis revealed: what staff position in the museum is responsible for digital preservation?

1. Understanding and Defining Digital Preservation

In 2015, technology is being increasingly integrated into our everyday lives; it provides tools and resources for information and documentation impact our very way of seeing the world. The rapid integration of technology has revolutionized mass accessibility to the internet and supported a practice important in digital preservation, Linked Open Data. For example, the sheer vastness of size, time, funding, and staff dedicated to The Metropolitan Museum of Art's online accessibility initiative exemplifies how museums recognize the way our society now aims, and even expects, to use the internet to connect with the world. However, as we have developed into an 'information society,' many have assumed that the ephemerality of digital material, or the abstractness of information encoded in the binary form of 1's and 0's, makes digital objects essentially immaterial, and as a result, not subject to any of the physical and environmental threats that we normally associate with physical materials.

Museum professionals, just like any other member of today's technological society, can also be susceptible to underestimating the threats to digital information. The misnomer of digital materials cannot be faulted to any individual, but arguably is the greater result of being part of the transition to a new age of technology-human integration. At the same time, although awareness of computer science and the way technology works is in flux for the current generation of cultural heritage professionals, it is the profession's responsibility to steward cultural heritage collections for the future, be they analog or digital. An important part of addressing threats to digital collections, grappling with broad changes in the use of technology in museum work, and stewardship of digital collections, is to have a shared definition of digital preservation.

'Digital preservation' however is often misunderstood both inside and outside the museum world, because it is often assumed that it means digitization of analog objects. While digitization forms part of digital preservation, it does not include the broad scope of activities encompassed by the term, especially the care and stewardship of digital objects themselves. As the case studies highlight, the San Francisco Museum of Modern Art, the Metropolitan Museum of Art, and the Museum of Modern Art New York have certainly *not* misconceived digital ephemerality within their stewardship practices, however, they *do* reflect varying levels of understanding of digital preservation, and furthermore, varying levels of digital stewardship.

As outlined in Chapter 3, the Digital Preservation Coalition outlines digital preservation to involve both "[a]series of managed activities necessary to ensure continued access to digital materials for as long as necessary" as well as "all of the actions required to maintain access to digital materials beyond the limits of media failure or technological change"(DPC 2015). Specifically, all three case study institutions practice digital preservation as defined, but within classified tiers of digital preservation maturity: *mature*, *established*, and *emerging*.

The Museum of Modern Art encompasses the most 'mature' digital preservation program that follows best practices from both within and outside the museum field. The digital preservation leaders within MoMA's staff had the foresight to understand that the expertise, education, and tools needed to reach their goals had to be found outside the museum field. Rather than reinvent digital preservation for the museum context, MoMA allied with already established communities from the library, archive, and digital preservation technology fields, and simply adjusted those best practices and tools to meet the mission and needs of its collections care. The staff on MoMA's digital repository team never lost sight that technology for digital preservation is nothing more than a tool, and is not something that can be blindly adopted without a deeper evaluation and understanding of how it meets the Museum's needs. As such, MoMA serves as an ideal model for the broader museum field.

On the next tier, the San Francisco Museum of Modern Art has an 'established' digital preservation program and is actively working towards future advancements, good practices, and policy. SFMOMA may not yet have all the details of its digital preservation systems figured out; however as an institution, the importance of digital preservation is understood because they look

to models like MoMA for collaboration, policy, and best practices for their own future. Thus it can be determined that while SFMOMA is still developing its implementation of a digital preservation system, they are very much a museum on the right path towards digital preservation maturity.

The SFMOMA has a distinctive view of the two components of digital collections within its care: time-based media artwork and digital asset management. Digital preservation at SFMOMA is therefore realized within two different systems: the digital art vault, and the digital asset management system. While the two categories remain separate, SFMOMA's team expressed in the interview for this thesis that they understand the importance of working towards preservation for both. SFMOMA's shared understanding for digital preservation within its holdings will guide them towards future success and good work in the practice of collections and records management.

The last tier of 'emerging' is the level of digital preservation that dominates the museum field today and therefore is the most relatable example of the three case studies. The Metropolitan Museum of Art encapsulates an 'emerging' digital preservation program. Despite its size and prestige, the Met's experience thus far with digital preservation proves that museums of all types and sizes will slowly grapple with understanding and achieving digital preservation. The Met's emerging digital preservation efforts exemplifies the way that the digital age is transforming museum work, as well as the importance of digital asset management in the realization of digital preservation.

The Met's earliest 'digital initiative' was massive, and focused on the actual digitization process itself, with over 1,000,000 photos online. Managing those assets using software that aids in metadata capture, organization, and dissemination was an obvious next step in ensuring that the huge amount of work went into digitizing was not wasted. Many other museums will relate to a scenario in which they 'go digital' and begin the challenging process of digitization, which is a major effort to fund and staff. Because of the focus on digitization and access, the other half of

the definition of digital preservation, the emphasis on long-term digital stewardship, has not yet been fully considered by the Met. Therefore, the Met maintains an emerging status in the maturity of its digital preservation program; new developments are on the horizon for them, but their approach stems from the shorter-term needs of a major digitization effort, so the specific trajectory towards trustworthy long-term digital preservation follows a particular pathway.

Digital preservation-specific tools and activities are most evident within the practices of MoMA and the SFMOMA who both use recommended standards such as the Open Archival Information System or preservation-specific metadata. While it is clear that these two museums represent many best practices for digital preservation, interestingly both MoMA and SFMOMA's digital preservation initiatives are very much aligned with *collections* activities. Unsurprisingly, both institutions are also prominent stewards of digital artwork, an artistic medium that is still underrepresented at the Met. Both the MoMA and SFMOMA have a clear understanding of the full set of activities required for digital preservation (bitstream and format preservation), and for both museums, such efforts remain primarily defined within the work of collections.

The efforts of the case study museums here highlight how a museum interprets *digital stewardship*. Traditionally in museums, records management has been treated somewhat separately from collections staff and activities. Collections management traditionally involved a higher level of thoughtful, time-consuming *preservation* activities compared to that of records management. Yet both records management and collections management are interrelated, because the significance of any museum object is only as good as the museum's understanding of the provenance, history, condition, and historical significance of that object, which is derived through the existence of thoughtful and accessible records management.

Finally, defining digital preservation within museums resembles that of the moving continuum model proposed by proposed by Zorich, Waibel, and Erway in 2008, as discussed in Chapter 4. The continuum model suggests that the similarities of digital collections among all three memory institution-types will ultimately bring these separate disciplines together because their collections are starting to resemble each other. The care needed for a Library's digitized book, for example, begins to look similar, if not the same, to an Archive's scan of an original

manuscript, as well as to a Museum's digitized slide collection. This concept of convergence can be applied to activities within individual museums, which traditionally have a variety of separate departments. It seems likely that digital stewardship within museums will move towards a model of convergence as these institutions define digital stewardship initiatives internally and amongst each other.

In conclusion, the changing landscape for museum work posed by the integration of the digital world in the museum profession will not only challenge the field to define digital preservation or stewardship, but will also challenge the traditional definition of collections management and care. Regardless of the level of digital preservation maturity, the collective museum field will find that the more digital assets and collections that make it into institutions, the more museums will need to orient, develop, and act on new understandings of collection management for the digital age. Digital preservation shares many of the philosophies of museum collection management --the actions that make up each are where the major differences lie. Considering the similarities between each practice, the museum field should adopt digital preservation with the same sense of responsibility and need as well-established collection management and care.

2. Integration of Digital Preservation Technology

Throughout this thesis, technological aspects of digital preservation have been referred to quite frequently. The focus on technology here was motivated by the need for a resource in the museum community that summarized much of the literature and tools used by the digital preservation community, as applied specifically to museums. Although technology can be a barrier to museums because of its expense, more and more museums are adopting digital technology in some form or another. As a result, the need to delineate and evaluate technological components of long-term digital preservation is important. Moreover, stewardship professionals must approach digital objects from the perspective of their long-term viability, and ensure that that relevant technological areas are carefully considered when evaluating the long-term value of a particular digital object, so that a technical and social infrastructure that supports preservation over time is in place.

From this perspective, the technology used in at least one of the three case studies can be observed to fall into three categories, as outlined below: *digital asset management systems, OAIS compliant software*, and *storage media.*

First, all three case studies used a *digital asset management systems (DAMS)* to manage the access and organization of their frequently used digital assets. Interestingly, the three case study museums may have communicated with one another about their evaluations of DAMS, because all three used the same software: first Media Bin, and then later, Net Exposure. More importantly, the shared use of DAMS in the case studies raises the question, is a digital *asset management* system the same as a digital *preservation* system?

Put simply, the answer is 'no.' Although DAMS can employ many similar preservation activities, this is only the case if the software is leveraged by its users to do so (Lazorchak 2012). Digital preservation systems ultimately are a set of processes, protocols, and policies that are most often mediated with some technological aspect to aid in creating information packages suitable for long-term storage. With so many software systems appearing in museums today - collection management systems, digital asset management systems, content management systems, and now digital preservation systems - truly understanding the differences in the way they are used can easily become confusing, especially if one does not have experience in understanding computer science and information science. So, once again, to put it in the simplest terms, a digital asset management system is *not* a digital preservation system in of itself, because a DAMS does not usually follow the specific recommendations for metadata, fixity checks, and formats that are put forth by Trusted Digital Repository model, the Open Archival Information System Reference Model, the Library of Congress, or other digital preservation communities such as the Online Computer Library Center (OCLC). However, a DAMS can be used for some preservation activities such as managing the legal and ethical information relating to a digital object, providing a platform for metadata capture upon creation, and can track the use and access of digital assets to avoid complete data loss or tampering.

The point above is further clarified by a 2012 blog post written by Butch Lazorchak for the Library of Congress Digital Preservation blog, *The Signal.* Lazorchak distinguished between the preservation goals of a DAMS and of a cultural institution, such as a library, archive, or in

this context, museum. The major differences between a DAMS and a digital preservation software is based on the type of data being preserved: for a DAMS that data is usually proprietary, whereas preserved data for long-term stewardship is usually open format. In addition, the purpose of preservation for a DAMS is mostly focused on monetization, whereas the goal for long-term digital preservation systems is related to unchanged access over time. Lastly, the time horizon is much shorter for DAMS, while memory institutions are concerned about data viability and accessibility for the long-term (Lazorchak 2012). The technology and infrastructures in a DAMS could be used with the same standards used in "doing" digital preservation however; for example, the specific workflows and methods for media storage would need to be outlined in specific protocols and policy to make a DAMs functional for digital preservation.

Where bitstream preservation is concerned, the Metropolitan Museum of Art and SFMOMA are using best practices to leverage preservation using their DAMS software, when possible. Whether that is through metadata encapsulation, normalization of formats, and disseminating access copies to eliminate human tampering, these methods fulfill the basic levels of preservation. In the context of the National Digital Stewardship Alliance "Levels of Digital Preservation" and that of Nancy McGovern and Ann Kenney's *Five Organizational Stages of Digital Preservation*, both SFMOMA and the Met meet many of the best practices for a level 1 or 2 stage of digital preservation, when considering the institutions' use of a digital asset management system software (DAMS).

The second technology observed in some of the case study museums was software that meets the requirements of the Open Archival Information System Reference Model (OAIS). While this type of software is not yet used at the Metropolitan Museum of Art, it is on the horizon at SFMOMA, and is fully functioning at the Museum of Modern Art. Although the implementation of *OAIS-compliant software* varies among the three case studies, what is evident is the rising awareness of OAIS as a useful model for the museum context. The Museum of Modern Art's decision to use Archivematica has resulted in the de facto use of OAIS as a general infrastructure for their digital repository. Strategically, it worked out well for MoMA that the concept of creating good archival information packages (AIPs) ultimately fulfilled MoMA's long-term digital stewardship goals. While other standards put forth in the digital

preservation community may not always be as easily applicable to the museum context, the foundational and generic requirements of the OAIS *is* a standard that can be applicable to any museum. Software tools that are designed to follow the OAIS standard are perhaps the easiest way for museums to ensure they are following the general steps for digital preservation.

Third, all three case study museums grappled with the technology involved in storage media. Notably, as outlined in the literature review, best practices for storing digital objects for long-term preservation is a complex topic, because there are so many options, which can vary depending on vendor relationships, IT department, size of collection, and access needs. However, only until one observes how actual museums today are handling the storage element of the preservation process do some best practices among museums come to light.

As outlined in Chapter 3 and in the case studies, digital storage options can include cloud-based services, collaborative redundancy systems like LOCKSS, spinning disc, magnetic tape, or hierarchical structures that use both disc and tape. Each case study, working within the best capacity they can, used the storage system that worked best with their means and priorities. Both the Metropolitan Museum of Art and the San Francisco Museum of Modern Art use an internal server and magnetic tape for backup storage of digital assets, for example, with the additional redundancy case in an off-site storage location. The Museum of Modern Art, however, took considerations for digital asset storage a step further through their contract with Arkivum for hierarchical storage that could be integrated directly with Archivematica and Binder. This forward-thinking strategy was based on the assessment of MoMA's collection growth of digital materials, as well as an assessment for the most cost-effective method to sustain a rapid rate of growth. The use of hierarchical storage (tape and disk media) through Arkivum's service increases the MoMA's ability to align with the recommendations set forth by the National Digital Stewardship Alliance. In particular, the hierarchical storage system provided by Arkivum meets the recommendations for scheduled fixity checks, and redundancy of backup copies in three (not just two) locations. MoMA's commitment to the stewardship of digital collections through its technology systems is a best practice within the context of standards supported by the digital preservation community itself.

Finally, it is important to note that each of the case study museums selected their storage systems after careful analysis of their needs; as a result, the best practice for one museum may not match what is necessarily best for another. However, the strategy of employing storage redundancy in *three* locations is a large take away from MoMA's storage system. Since digital objects can be reproducible, taking advantage of their inherent nature can only better ensure that digital objects are kept safe in the instance of natural disaster, malicious activity, or accidental data change.

3. Collaboration

Collaboration in digital preservation efforts in museums is another key theme that emerges from a consideration of the case studies and the literature review, both among different kinds of organizations, such as libraries, archives, and museums, as well as within museums themselves, as a way to create a supportive network.

The three case study museums engaged in a variety of levels of collaboration successfully including relationships outside the museum field with vendors and others in the digital preservation community. For example, Matters in Media Art, involving both SFMOMA and MoMA, is an excellent example of successful inter-museum collaboration that resulted in open discussion and deliverable best practices for the rest of the museum field. The success of Matters in Media Art in jumpstarting the digital preservation programs in two of the three case studies proves it to be a model worth repeating in future museum projects. At the same time, MoMA's effort to open-source its software, Binder, is an act of open collaboration in of itself by making this tool freely accessible to others, and by inviting other institutions to use, modify, and enhance this tool.

Outside of the important external collaborations observed in the case studies, the need for internal collaboration to develop and implement successful digital preservation, as well as to leverage support, is also apparent. SFMOMA fosters a highly collaborative staff environment; this is key to their ability to unite in the work of individual departments within the greater institution. The Metropolitan Museum of Art on the other hand, as such a large-sized institution, has major units (library, archive, and museum) that still maintain a larger degree of separation.

This is a situation that many museums, big or small, may recognize in themselves. However, the Met *is* developing collaborative efforts to synthesize and optimize the digital asset management workflows between the three units, and ultimately to support more collective digital preservation activities. Internal collaboration is necessary for uniting the staff in museums who work in the frontlines of any digital preservation initiatives.

4. Policy

Although technology is an important, and arguably unavoidable, element in digital preservation planning, and collaboration is important, another consistent theme among the case study examples is the importance of institutional attitude and commitment towards digital collections. Digital preservation is not just a technology problem, but as confirmed by many in the digital preservation field, it is a management issue. *Policy* is a tangible method for institutions to outline the management support of their preservation activities; this very same perspective has been posited by the museum community for many years in regards to collection management policies. Now, in the digital age, museums need to support policy development for both analog and digital assets. Digital preservation policy not only holds a museum accountable for its activities and ethical handling of its digital assets, but it also marks the institution's acknowledgement that digital collections must be stewarded on the same level as traditional collections.

If the decision makers in any given institution do not understand the need for a digital preservation system, regardless of how basic or advanced the system is, then the basic framework required will not be in place, and it will be extremely difficult to proceed. The writing of policy is a process for involving leadership within a cultural institution, and supports an understanding of the true value and institutional obligation for digital preservation. Unfortunately, it comes as no surprise that digital preservation policy, let alone digital preservation initiatives, are still very rare within the museum field.

Most institutions that begin to work in digital preservation do not have policy to guide them. Policy continues to be an achievement only after some basic planning and implementation has already taken place. Policy after the fact is not a bad practice, but more often simply the

only course of action an institution has. An equally effective pathway is taking time to evaluate what works and does not work within a digital preservation plan before the formulation of policy. However, this is all based upon having the gusto, advocacy, and support needed to jumpstart a digital preservation program in the first place.

For example, The Metropolitan Museum of Art, although using some digital preservation strategies, does not yet have as clearly a defined digital preservation plan, such as SFMOMA's digital art vault, and certainly does not have a plan of the caliber of the one at the Museum of Modern Art. One can speculate that the Metropolitan Museum of Art will slowly gain awareness around digital preservation as the Museum moves along a continuum of collaboration internally within its library, archive, and museum units. The beginnings of the Met's time-based art collection may also instigate further conversation around digital preservation needs. Regardless of these factors, without some kind of change agent, the conversation around digital preservation may never gain the momentum it needs to instigate advancements in the area of long-term stewardship of digital collections. For museums stuck in stasis or unable to advance to more mature levels of digital preservation, one can argue that digital stewardship policy could be the way to stimulate digital preservation projects.

The team at SFMOMA is acutely aware of the need for policy around digital preservation, and the museum intends to work towards that goal. This is positive, and ultimately, is the only approach one could ask of an institution that is still grappling with this complicated topic. Other factors however can inhibit a museum from implementing policy, such as gaining the attention of the Board of Trustees. Dedicating the time for a Board to create or approve a policy can be an uphill battle, especially when the question of budget is raised. For its part, SFMOMA has implemented action plans and workflows outside of high level policy as a way to create structure and document their progress in the area of digital preservation. These actions will not only support the Museum's ability to implement successful digital preservation activities, but will also help the museum attain funding, such as grants. For other museums that find high-level policy to be out of their reach, starting with preservation plans and protocols, can be very helpful when policy itself remains an out-of-reach goal.

It is unsurprising that the case study with the most mature digital preservation program, MOMA, has the largest digital art collection, and that it also has the only official digital preservation policy. The Museum of Modern Art exemplifies the highest degree of best practice in the field of digital preservation within the museum context, including its development of policy. Not only has the MoMA developed one of the very few digital preservation policies in a U.S. museum to date, but its staff used model digital preservation policies from the allied community to apply best practices to the context of MoMA's own repository.

In the spirit of the calling for best practices in governance and policy by one of the museum field's most important authors, Marie Malaro, museums can only expect to continue doing good work in the legal and ethical handling of assets (digital and analog included) with guiding documents that integrate the museum's mission with the everyday work of its collections. The ultimate goal for any digital preservation program should be to implement policy one day, informed by professional standards suggested by the Library of Congress, the Trusted Digital Repository model, and the Open Archival Information System Reference Model, as well as relying on initiatives such as Planets, the Online Computer Library Center, the Joint Information Systems Coalition, and the Canadian Heritage Information Network. Museums would be best advised to take the advice of the many professional networks that are working to to make digital preservation accessible and understood among cultural heritage institutions, and prioritize the inclusion of digital preservation in current policy.

Concluding Thoughts

The four themes discussed above highlight an important challenge facing museums in their digital preservation efforts: what staff position in the museum is responsible for digital preservation? Whether one is grappling with the meaning of digital stewardship for their museum, the technology needed to implement preservation activities, collaborating for support, or looking to policy for long-term guidance, there remains a level of uncertainty around which individuals are the ones who bear the responsibility to manage, track, and implement digital preservation. Libraries and archives, who thus far lead the field of digital preservation, typically are organized differently than museums and have fewer departments, and therefore, a smaller

variety of staff roles and digital formats exist under their roof. While the role of digital preservation may be clearly defined for librarians and archivists, this is not true yet for the museum field, and this is likely one reason for the disjunction between digital preservation and museum work. Perhaps the most logical starting place for advocating for digital preservation in the museum context is among the collections staff.

Because the Metropolitan Museum of Art's digital initiatives sprouted from major digitization and online access mandates, it seemed a natural progression for the digital asset management team to be responsible for monitoring metadata capture, quality and ingest protocols, and ensuring backups and access to master files. The Met has an awareness of digital preservation and the associated technology, but is perhaps absent is the need to integrate the responsibility for long-term digital preservation into the jobs of the digital asset management team. Instead, long-term digital preservation is a de facto result of the Met's needs for access and dissemination of digital materials.

The San Francisco Museum of Modern Art seeks to hire a time-based media conservator who specializes in digital artwork, but until they can move forward with adding a new staff member, the responsibility of digital preservation falls to contract staff, with additional collaborative support from the Information and Access, IT, and curatorial teams. As a result, permanent responsibility for digital preservation remains somewhat unclear, as digital preservation is not yet as distinctive role as is long-term maintenance of records.

The Museum of Modern Art is the only case study that has a defined team of people dedicated to the management of their digital repository. The Digital Repository Manager, bolstered by their Media Conservators and IT engineers on staff, creates a clear directive for the whole museum regarding who bears the responsibility for digital preservation. However, the Museum of Modern Art's vast number of digitized media and born-digital artworks is perhaps the biggest call to action within the three case studies for digital preservation, and this correlates with MoMA's development into a leading institution that provides staff, money, technology, and policy for stewarding digital collections.

In light of the difficulty of defining who is responsible for digital preservation, a major discussion should be had; otherwise, museums will be motivated to take action only when there

are dire threats to important assets. *Collections* is perhaps the most accessible department any museum can leverage for immediate need of digital preservation, although there are inevitably additional digital assets in the institution that are also important for long-term storage. Since museums typically share the goal of permanent, perpetual stewardship of cultural heritage, the care of its unique collections is perhaps the most accessible way to gain support for digital preservation. Not all museums have the luxury of having full-time conservators on staff, but any museum will have either a collection manager, or a registrar who also manages the safety of collections. These particular museum positions are the suggested way to advocate for digital preservation in the future.

The MoMA and SFMOMA's initiatives that most closely align with the digital preservation community are both focused on the specific care of accessioned collections, and not so focused on that of records management. The Met, which does not yet possess many time-based or digital artworks, is slower in developing a digital preservation system. Although registration and records management may take lower priority at a museum, the relationship between a museum's collection and the records relating to that collection will hopefully lead one day to further integration of digital preservation into other staff positions in the museum.

In conclusion, considering the terrain for museum staff roles in the field today, the role and responsibility of digital preservation may need to be fulfilled by those that advocate for the care of museum collections. Regardless of title, making the case for digital preservation around museum collections will be the most accessible way to make digital preservation mainstream within the museum field.

In the next chapter, several conclusions concerning the state of digital preservation in museums today will be presented.

Chapter 11: Conclusions and Recommendations to the Field

Fifty to a hundred years from now, how will museums access, plug in, or turn on today's cultural heritage objects stored in digital forms? This is the difficult question that digital preservation asks the museum field to consider. The present question resonates in a striking way with the familiar *duty of care* charged unto museums for their traditional collections of artwork, historical artifacts, and scientific specimens. However, this time, we are working within the context of a medium both abstract and unfamiliar to most of the museum profession: the 1's and 0's of source code. Now deeply embedded in a new age of technology, museums must extend their call to duty as caretakers and stewards of cultural memory to the new-age artifacts of today-- those that are born-digital.

Collection management and digital preservation share parallel missions, but it is also true that digital collections and assets cannot remain in preservation stasis for long without succumbing to physical degradation, technological/format obsolescence, bit rot, and complete data loss. It can be concluded then that museums need to expand collections management to include the preservation standards and strategies from the digital preservation community, and do so with quickly. With time and help from outside the museum field, the same good work that is applied in traditional collections management will need to include the new frontier of cultural heritage found on tapes, computers, and discs.

Below, a set of five conclusions concerning the state of digital preservation in museums today are presented: first, preservation is possible; second, standards, guidelines, and best practices are already available, but use wisely; third, embrace new practices in policy; fourth, collaboration will be key for success; and five, embrace change and act now. Digital preservation is certainly possible, and for museums just starting out, this can be done even in small capacities. Recommendations to the field for starting digital preservation for museums of any size will also be presented. The overall goal of this chapter is to inspire and motivate the museum field to give digital preservation the time and priority it deserves, to understand that digital preservation is possible for all museums, and that the time for museums to act is now.

Conclusion 1: Preservation Is Possible

There is no doubt that digital preservation will prove to be a difficult goal for many museums because of the existence of a variety of barriers, including knowledge of key issues, available resources, and staff time; nevertheless, the current situation for the stewardship of digital collections is cause for concern. Although safekeeping digital materials is challenging, the museum profession cannot afford to hold off taking action without serious legal and ethical consequences.

There is also no doubt, however, that digital preservation is possible for the museum field today. More than a decade of research from the library, information science, academic research fields has worked out many of the challenges that faced memory institutions back in the early 2000s. By emphasizing the five areas described below, *metadata capture, assessment and inventory, accessing open source software, recognizing the ongoing nature of digital preservation,* and *planning*, museums can jumpstart their digital preservation efforts.

Metadata Capture: Much of digital preservation involves the important role of metadata capture, or intellectual control of digital assets. Over time, the information about a digital object can be lost or forgotten as staff members come and go. So, implementing protocols for standardized metadata to be captured upon creation and acquisition is therefore an easily accessible starting point for any museum to prepare its digital collections for the greater process of preservation. Using recommended metadata schemas, such as VRA Core and PREMIS, costs virtually nothing, but only requires staff initiative and time.

Assessment and Inventory: An important action to develop momentum for digital preservation activities in museums is to create a robust assessment and inventory of one's digital assets. A well-educated review of what an institution contains that is worth preserving (and inevitably not all digital assets will require long-term preservation) is an important step in assessing the preservation needs of a collection, as well as an effective step in advocating and lobbying for higher-level support (and funding) for digital preservation programs in our institutions.

Free and Open-source Software Tools: The three case studies exemplified a particular degree of commitment to technology systems that aid each museum in achieving a variety of

levels of digital preservation. While these case studies serve as excellent examples or models for the museum field, there is still the very apparent reality that many museums will never have the financial means to invest in technology systems like those at the MoMA or the Metropolitan Museum of Art. Yet, digital preservation is still possible. Many in the cultural sector who have implemented digital preservation systems are acutely aware of the financial burden that digital preservation can present. This is why the digital preservation community has provided a variety of free and open-source software tools that can be adapted to the museum context. With some education, collaboration, and time put into implementation, museums can and should take advantage of these available tools. Some recommendations include Archivematica, BagIt, Digital Record Object Identification (DROID), and EMET (Embedded Metadata Extraction Tool). These and more free tools have been outlined by the IMLS-funded Digital POWRR project, a resource highly recommended to the museum field.

Recognizing the ongoing nature of digital preservation: Although simple physical media storage is not a recommended tactic for the long-term, for some small sized museums, this may be the only feasible storage option available. Luckily external hard-drive storage has become increasingly less expensive, with a terabyte of storage available for a little as $80 and can withstand degradation for a decent number of years. However, this strategy is only recommended with the contingency that museums should regularly heed to the 'digital curation' activities in addition to using hard-drives; one cannot simply put digital material on a hard-drive, and leave it on a shelf for 10 years without consequences.

The area of storage highlights an important recommendation: museums must recognize that digital preservation is ultimately a continuous, ongoing process. As soon as the museum profession understands and accepts the ongoing time commitment required of digital stewardship, the more regular and normalized digital curation will become in our institutions. Any size museum can also implement a basic protocol: have the foresight to choose non-proprietary digital formats for storage, store metadata files in simple text formats (such as XML) to be stored with the digital object, conduct regular fixity checks, refresh files or media occasionally, and create at least three master copies of digital assets, two of which should be

stored off-site and backed up regularly. These recommendations for simple, or 'good enough' digital preservation stems from the work and advice compiled in this thesis from resources like the IMLS-funded Digital POWRR, the National Digital Stewardship Alliance, the Digital Curation Centre, and the Society of American Archivists. And underlying these recommendations is the idea that digital preservation is an ongoing process.

Planning: Digital preservation can range from a set of complicated processes, to very simple steps that ultimately just require action, and ongoing commitment to those actions. To maintain the momentum and fastidiousness that many busy museum professionals will be up against in their work, digital preservation plans, checklists, and policies will be key to defining staff roles, work timelines, metadata standards, quality control, and other digital preservation activities.

In sum, digital preservation is possible in the museum field today, and museums can begin efforts in this area, with a recognition of the good work from allied fields, and by following the simple recommendations outlined above.

Conclusion 2: Use Standards and Best Practices Wisely

As summarized in this thesis, basic digital preservation standards, guidelines, and best practices have now emerged. The material is available for the museum field, and need only be pursued. In addition to books, there will always be academic articles, as well as studies accessible only on the Web, demonstrating the very reason why digital formats are dominating the information highway. The bibliography of this thesis aims to function as a useful starting point for museums to gain some education and perspective for digital preservation.

In particular, as a way to weed through the many relevant resources available to museum professionals on the topic of digital preservation, it is recommended that museums pay particular attention to the Open Archival Information System and the Trusted Digital Repository Model. For museums seeking practical resources, templates for digital collection audits, digital preservation plans, and digital preservation policies are readily available for museums on the Canadian Heritage Information Network (CHIN) website. The resources provided by CHIN can

be liberally adapted to any museum type, although it is recommended that museums be open to amending any features of these templates that do not pertain to its specific needs.

Another highly recommended practical resource for the museum field is the ISO 163163 Trusted Repository Audit and Checklist (TRAC). As a formal standard recognized by the international cultural heritage community, using ISO 16363 as a self-auditing tool can be a useful guideline for any museum committed to making sure their institution is working towards the requirements for trusted digital stewardship.

Cost models are also useful tools for any museum advocating for funding from the management level. For more basic digital preservation programs, simple audits can help determine what funding will be needed for up front digital preservation. However, for museums wishing to progress to more mature and sustained digital preservation initiatives, recommended resources for cost models include the LIFE (Life Cycle Information for E-literature) Model, CMDP (Cost Model for Digital Preservation), and Total Preservation Cost Analysis recommended by the UC Curation Center at the California Digital Library (LIFE 2015; CMDP 2012; Abrams Cruse, and Kunze 2012).

Although building foundations for digital preservation from tools supported by allied disciplines, namely library and archives, will be key to the promotion of digital preservation in the museum field, as exemplified by the MoMA case study, it is important for museums to remember the uniqueness of its collections in opposition to that of allied fields. Formal standards are excellent guideposts and educational tools that will create a foundation for digital preservation, but they do not always provide a one-size-fits all solution for preserving museum collections. The wider variety of digital collection types in museums, including original artworks, will require museums to bend the rules set by formal standards in order to recontextualize the governing uses of digital preservation best practices for the museum world. The Museum of Modern Art tested and *adapted* standards for their own digital repository. Another conclusion and recommendation to the field is to follow the guidelines of the digital preservation community, but using MoMA as a model, to do so with skepticism and openness to adaptation. Museum professionals are advised to be cautioned against lapsing into an unquestioned dependency on best practices and technology solutions developed in other fields.

Museums must always return to the question of how tools and practices from other fields fulfill the mission and digital preservation needs associated with their own institution.

3. Embrace New Practices in Policy

The concept of needing policy to aid in a museum's governance and mission for stewardship is perhaps the most familiar topic to the museum field discussed within this thesis. Therefore, little argument should be needed to convince the field that digital objects --be they historical records, curatorial research, institutional records, collection documentation, or accessioned collections -- are just as important as their analog counterparts that museums already take great care to steward in the name of the public trust. Therefore digital collections require, furthermore deserve, the same amount of institutional commitment and policy.

While some digital assets in museums may require more preservation action than others, the option to do nothing to make sure these materials are viable for our future cultural record, is surely not an option. An important recommendation to museum professionals is to take action to be advocates for digital preservation. Lobbying for future planning and policy for digital preservation can only happen from the humble efforts of museum professionals on the front line. Even if high-level, institution-wide policy remains out of reach, those with the ability to advocate for internal, departmental policies is surely better than remaining in a dangerous stasis. Policy can help guide a wider network of staff to understand the importance of digital preservation, and to instigate more unifying support to act on digital preservation.

4. Collaboration Will Be Key to Success

Education on digital preservation is very much absent within the museum field. Without more professionals in the field who understand the foundations of digital preservation, developing robust digital preservation initiatives may be an uphill battle. As a burgeoning topic in museums, more digital preservation research and more educational resources within the museum field is needed in order to 'catch up' to the library and archive fields. This 'catching up' is therefore going to be reliant on museum professionals being open to collaboration with LAMs by way of participation in professional conferences, workshops, online training, and collaborative

partnerships, all of which already exists within the digital preservation community. Some examples of such digital preservation community resources include the Northeast Document Conservation Center, who offers online workshops and free resources; the Library of Congress Digital Preservation Outreach and Education initiative, which provides a national calendar of digital preservation courses; the Digital Directions conferences, which explores the challenges and best practices surrounding care of digital collections; and the Preservation Archiving Special Interest Group, which hosts webinars and an annual conference-style meeting.

The museum field needs to formulate a more expansive network; the library and archive fields are just as concerned and will share many of the same challenges that museums face regarding digital preservation. Therefore, the museum field must take advantage of the open digital preservation community, a community that is willing to collaborate and help in developing common goals of long-term preservation and stewardship of our cultural memory within the digital age.

There have already been many successful attempts at creating collaborative solutions for digital preservation outside the museum field. Museums have simply not followed suit. However, raised awareness of the need for digital preservation in museums will inspire more museum professionals to seek participation in collaborative projects. Matters in Media Art is an excellent example of a successful museum collaboration, and it serves as a model for opening conversation and creating tangible solutions for digital preservation in museums.

Museums need not reinvent the wheel, and they should take cues from already established models in the digital preservation community. A visionary example can be found in the MetaArchive, a project open to any cultural institution with the mission to alleviate the costs and technology of digital preservation, and spread the work among a network of institutions. Following the model of LOCKSS (Lots of Copies Keep Stuff Safe), the MetaArchive is a model that addresses the major barriers in digital preservation that many museum professionals will need to address: funding, community, and support. Could museums work together to formulate

their own version of the MetaArchive model? Can museums work together to distribute the burden of digital preservation from the individual to the many? While these questions will remain unanswered, they imply a hopeful message for the future as we continue to progress into an evolving age of technology.

No one institution can expect to achieve a mature level of digital preservation alone. Within the current economic landscape, museums need to promote interdisciplinary collaboration and communication in order to create the necessary support system that it will take to achieve widespread digital preservation within memory institutions alike.

Conclusion 5: Embrace Change and Take Action

Digital preservation is an ongoing activity, an agreed set of outcomes, an understood responsibility, a selection process, and a cooperative activity. Digital preservation is a public good. Only time will tell how the museum field will embrace the reality of digital preservation and stewardship. Although the future is uncertain, that reality is embedded in hard work and a professional shift in the way we treat digital collections. The impact of digitization, digital artworks, and other digital media has changed the landscape of cultural heritage and the way we envision social memory forever. Museums are advised to keep up with the changing world around them and strive to always improve their work in fulfilling their valiant missions in the name of the public trust and the lifespan of our cultural memory. In the end, museums must bravely embrace change in the area of digital stewardship, and take action.

As these five conclusions highlight, digital preservation today is possible for museums and does not have to be manifested in a complicated technology system, but must be acted upon nonetheless. Unlike collections that can be stored onto shelves for a later time, cultural heritage and institutional assets in digital forms require immediate and iterative action. Digital preservation as a practice, commands cultural institutions to be proactive in beginning digital preservation from the very creation of digital objects. Having proactive protocols for managing digital assets and collections from creation to dissemination to storage, needs to be as integrated into our professional practice as the habitual condition reporting, collections housing, and environmental regulating of traditional collections management.

In conclusion, "To create a collection, to inherit one, or to be given oversight of a collection, is also to create, inherit, or accept a great responsibility" (AIC 2002). This statement, from the American Institute of Conservation, summarizes the core sense of ethical responsibility towards stewardship shared among all collecting institutions. Digital preservation offers a contemporary extension to the ethos of that statement. Digital preservation truly introduces nothing different from the standards and collections practices of the museum field, but rather offers solutions, even manageable solutions, that museums can use to uphold their responsibility to care for cultural heritage of the digital age.

Works Cited

Abrams, Stephens, Patricia Cruse, and John Kunze. "Total Cost of Preservation (TCP): Cost Modeling for Sustainable Services." *UC Curation Center, California Digital Library*. April 9, 2012.

Allinson, Julie. "OAIS." *JISC Standards Catalogue*. Updated September 23, 2006. Accessed September 10, 2015. https://web.archive.org/web/20130328032330/http://standards.jisc.ac.uk/catalogue/OAIS.phtml

American Alliance of Museums (AAM). "Continuum of Excellence." *Assessment Programs*. Last updated 2015. Accessed October 5, 2015. http://www.aam-us.org/resources/assessment-programs.

_____ (b). "Assessment Types." *Museum Assessment Program* (MAP). Last updated 2015. Accessed October 5, 2015. http://www.aam-us.org/resources/assessment-programs/MAP/assessment-types.

American Institute for Conservation of Historic and Artistic Works. "Position Paper on Conservation and Preservation in Collecting Institutions." June 2002. Accessed September 5. 2015. Digital version. http://www.conservation-us.org/docs/default-source/governance/position-paper-on-conservation-and-preservation-in-collecting-institutions-(june-2002).pdf?sfvrsn=7.

Anderson, Martha. "B is for Bit Preservation," *The Signal* blog, September 7, 2011. Accessed October 2, 2015. http://blogs.loc.gov/digitalpreservation/2011/09/b-is-for-bit-preservation/.

Archive-It. "About Us." *Archive-It blog*. 2014. Accessed October 6, 2015. https://archive-it.org/learn-more/.

_____ (b). "New York Art Resources Consortium (NYARC)." *Nyarc.org*. 2014. Accessed October 6, 2015. https://www.archive-it.org/organizations/484.

Archivematica. "Archivematica: preserving memory since 2009." *Artefactual Systems Inc*. 2015. Accessed October 6, 2015. https://www.archivematica.org/en/.

Archivists' Toolkit. "Archivists' Toolkit." 2009. Accessed October 1, 2015. http://www.archiviststoolkit.org/.

Arkivum. "Case Study - MoMA." *Arkivum Ltd.* September 2015. Provided by request October 2015. Digital document. http://arkivum.com/wp-content/uploads/2015/08/MoMA-Case-Study.pdf.

Artefactual Systems."User Manual -What is Binder?" Last updated May 22, 2015. Accessed July 10, 2015. http://binder.readthedocs.org/en/latest/user-manual/overview/intro.html.

_____ (b)."User Manual -Current Project Status." Last updated May 22, 2015. Accessed July 10, 2015.http://binder.readthedocs.org/en/latest/user-manual/overview/project-status.html.

ArtStor. "The Metropolitan Museum of Art." *Artstor.org.* Accessed September 26, 2015. http://www.artstor.org/content/metropolitan-museum-art-0.

Beagrie, Neil, Najla Semple, Peter William, and Richard Wright. "Part 1: Final Report October 2008." *Digital Preservation Policies Study* prepared by Charles Beagrie and Joint Information Systems Coalition, 2008. Digital PDF maintained by the UK WebArchive, http://www.webarchive.org.uk/wayback/archive/20140615022334/http://www.jisc.ac.uk/media/documents/programmes/preservation/jiscpolicy_p1finalreport.pdf.

Beagrie, Neil, Najla Semple, Peter William, and Richard Wright (b). "Part 2: Appendices – Mappings of Core University Strategies and Analysis of their Links to Digital Preservation." *Digital Preservation Policies Study* prepared by Charles Beagrie and Joint Information Systems Coalition, 2008. Digital PDF maintained by the UK WebArchive, http://www.webarchive.org.uk/wayback/archive/20140615022345/http://www.jisc.ac.uk/media/documents/programmes/preservation/jiscpolicy_p2finalreportappendices.pdf.

Blue Ribbon Task Force on Sustainable Digital Preservation and Access. "Index." 2008. Accessed September 3, 2015. http://brtf.sdsc.edu/index.html

Boudrez, Filip. "Digital Containers for Shipment into the Future." *Expertisecentrum DAVID* (eDavid). Antwerp, Belgium. 2005.. Accessed September 14, 2015. Digital article. http://www.expertisecentrumdavid.be/docs/digital_containers.pdf.

Bowling, Melissa. "Interview with: Jim Moske, The Metropolitan Museum of Art Archives." *Society of American Archivists Museum Archives Section Standards and Best Practices Working Group Electronic Records Project*: 43-46. July 15, 2014.

Brown, Adrian. *Practical Digital Preservation: A how-to guide for organizations of any size*. 2013. American Library Association Neal-Schuman: Chicago, IL.

Byers, Fred R. *Information Technology Care and Handling for the Preservation of CDs and DVDs - a Guide for Librarians and Archivists*. Gaithersburg, Md.: National Institute of Standards and Technology, Technology Administration, U.S. Dept. of Commerce, 2003.

Canadian Heritage Information Network (CHIN). "Digital Preservation Plan Framework for Museums." *Canadian Heritage Information Network*. Last updated June 20, 2013. Accessed May 10, 2015. http://www.rcip-chin.gc.ca/carrefour-du-savoir-knowledge-exchange/cadre_plan_preservation_numerique-digital_preservation_plan_framework-eng.jsp.

_____ [b]. "Digital Preservation Survey: 2011 Preliminary Results." *Digital Preservation Toolkit*. Last modified June 15, 2013. Accessed March 2015. http://www.rcip-chin.gc.ca/carrefour-du-savoir-knowledge-exchange/sondage_preservation_2011-preservation_survey_2011-eng.jsp.

_____ [c]. "CHIN Introduces a Digital Preservation Toolkit." Last modified June 12, 2013. Accessed May 5, 2015. http://www.rcip-chin.gc.ca/sgc-cms/nouvelles-news/anglais-english/?p=6410.

Choi, Jenny. Interviewed by Emma P. James. In-person interview at The Metropolitan Museum of Art, New York, NY. May 19, 2015.

Cohen, Daniel J. and Roy Rosenzweig. "Preserving Digital History: The Fragility of Digital Materials." *Digital History: A Guide to Gathering, Preserving, and Presenting the Past on the Web*. 2005. Accessed August 23, 2015. E-book. http://chnm.gmu.edu/digitalhistory/preserving/1.php.

Consultative Committee for Space Data Systems (CCSDS). *Audit and Certification of Trustworthy Digital Repositories, Recommended Practice, CCSDS 652.0-M-1, Magenta Book*, Issue 1. Washington, D.C., September 2011.

Consultative Committee for Space Data Systems (CCSDS). *Reference Model for an Open Archival Information System (OAIS): Recommended Practice CCSDS 650.0-M-2; Recommendation for Space Data System Practices*. Magenta Book, Recommended Practice, issue 2. Washington D.C.: CCSDS Secretariat, June 2012. Digital version. https://public.ccsds.org/publications/archive/650x0m2.pdf

Corrado, Edward M. and Heather Lea Moulaison. *Digital Preservation for Libraries, Archives, and Museums*. 2014. Rowman & Littlefield: Lanham, MD.

"Cost Model for Digital Preservation (CMDP)." *Danish National Archives*. Last updated 2012. Accessed November 5, 2015. http://www.costmodelfordigitalpreservation.dk/.

Dempsey, Lorcan. "Scientific, Industrial, and Cultural Heritage: a shared approach: a research framework for digital libraries, museums and archives." *Ariadne: web magazine for information professionals*, 22 (January 12, 2000). Accessed September 22, 2015. http://www.ariadne.ac.uk/issue22/dempsey/.

"Digital Asset Management System Policy." Unpublished document provided by Layna White, San Francisco Museum of Modern Art. August 8, 2014.

Digital Curation Centre. "About the DCC." 2004. Accessed July 20, 2015. http://www.dcc.ac.uk/about-us..

_____(b). "What is Digital Curation?" 2004. Accessed April 15, 2015. . http://www.dcc.ac.uk/digital-curation/what-digital-curation.

_____(c). "JSTOR/Harvard Object Validation Environment (JHOVE)." Digital Curation Centre. 2015. Accessed September 20, 2015. http://www.dcc.ac.uk/resources/external/jstorharvard-object-validation-environment-jhove

_____(d). "List of Metadata Standards." *Digital Curation Centre* Resources. 2015. Accessed September 15, 2015. http://www.dcc.ac.uk/resources/metadata-standards/list.

Digital POWRR."Tool Grid." 2013. Accessed February 19, 2015. http://digitalpowrr.niu.edu/tool-grid/.

Digital Preservation Coalition (DPC). "Digital Preservation Handbook." 2015. Accessed August 10, 2015. http://www.dpconline.org/advice/preservationhandbook.

_____. "Glossary." 2015. Accessed July 19, 2015. http://www.dpconline.org/advice/preservationhandbook/glossary#D.

Dillon, Patrick M., and David C. Leonard. *Multimedia and the Web from A to Z.* Phoenix, AR: The Oryx Press, 1998.

Duff, W., Carter, J., Cherry, J. M., MacNeil, H. & Howarth, L.C. "From coexistence to convergence: studying partnerships and collaboration among libraries, archives and museums" *Information Research*, 18, no. 3, paper 585 (2013).

Duranti, Luciana, Jim Suderman, and Malcolm Todd. *A Framework of Principles for the Development of Policies, Strategies and Standards for the Long-term Preservation of Digital Records.*The InterPARES 2 Project, March 2008. Digital version. http://www.interpares.org/public_documents/ip2(pub)policy_framework_document.pdf.

Duranti, Luciana and Randy Preston, eds. *International Research on Permanent Authentic Records in Electronic Systems (InterPARES) 2: Experiential, Interactive and Dynamic Records*. Padova, Italy: Associazione Nazionale Archivistica Italiana, 2008. Digital version. http://www.interpares.org/ip2/display_file.cfm?doc=ip2_book_complete.pdf.

Duryee, Alex. "What is the Real Impact of SHA-256? A Comparison of Checksum Algorithms." AVPreserve. October 2014. Digital version. http://www.avpreserve.com/wp-content/uploads/2014/10/ChecksumComparisons_102014.pdf

Electronic Resource Preservation and Access Network (ERPA). "Digital Preservation Policy Tool." *ERPA Guidance*, September 2003. Accessed July 12, 2015. Digital version. http://www.erpanet.org/guidance/docs/ERPANETPolicyTool.pdf

Erway, Ricky. "Walk This Way: Detailed Steps for Transferring Born-Digital Content From Media You Can Read In-House." June, 2013. Accessed September 13, 2015. http://www.oclc.org/content/dam/research/publications/library/2013/2013-02.pdf.

Final Report of the Blue Ribbon Task Force on Sustainable Digital Preservation and Access. *Sustainable Economics for a Digital Planet: Ensuring Long-Term Access to Digital Information*. February 2010. Digital version. http://brtf.sdsc.edu/biblio/BRTF_Final_Report.pdf.

Fino-Radin, Ben. "Conservation in Collections of Digital Works of Art." Presented at the *Electronic Media Group Session, AIC 40th Annual Meeting*, May 8–11, 2012, Albuquerque, NM. Printed in The *Electronic Media Review*, vol.2 (2013): 101-112.

_____. "Digital Preservation Practices and the Rhizome ArtBase". 2011. http://media.rhizome.org/artbase/documents/Digital-Preservation-Practices-and-the-Rhizome-ArtBase.pdf

_____. Interviewed by Emma P. James. In-person interview at The Museum of Modern Art, New York, NY. May 20, 2015.

_____ (b). "MoMA's Digital Art Vault." *Inside/Out blog* for *Moma.org*, April 14, 2015. Accessed May 12, 2015. http://www.moma.org/explore/inside_out/2015/04/14/momas-digital-art-vault.

_____ (c). "MoMA Media Conservation - Digital Preservation Policy." Unpublished. Email to author on May 22, 2015.

Fino-Radin, Ben, Karen van Malssen, and Dan Gillean. "Open Source and Digital Preservation and Access: The First Digital Preservation Repository for Museum Collections: An Open Source Approach." Presented at the *Association of Moving Image Archivists Conference*, Savannah, Georgia, October 8-11, 2014.

Fleming, Robin and Dan Lipcan. "Digitizing the Library's' Collections: An Introduction," *Now at The Met*, January 5, 2012. Accessed September 27, 2015. http://www.metmuseum.org/about-the-museum/now-at-the-met/features/2012/digitizing-the-library-collections.

Griesinger, Peggy. "Preserving the Technical History of Media Works." *Inside/Out* blog for *Moma.org.* May 20, 2015. Accessed July 10, 2015. http://www.moma.org/explore/inside_out/2015/05/20/preserving-the-technical-history-of-media-works

Gwinn, Nancy E. "LAMMS and International Collaboration." Paper presented at the *ICOMOS Scientific Symposium: Changing World, Changing Views of Heritage: the impact of global change on cultural heritage - Technological Change*, Valletta, Malta, 7 October 2009.

Habing, Thomas."METS, Mods, and PREMIS, Oh My!" *Library of Congress, MODS Official Website*. Presentation for American Library Association 2007 Meeting. Last Updated 2010. Accessed September 13, 2015.

Harvey, Ross and Martha R. Mahard. *The Preservation Management Handbook: A 21st-Century Guide for Libraries, Archives, and Museums*. Rowman & Littlefield: Lanham, MD, 2014.

Hedstrom, Margaret and Sheon Montgomery. "Digital Preservation Needs and Requirements in RLG Member Institutions." *A Study Commissioned by the Research Libraries Group.* Mountain View: CA, 1998. PDF e-publishing. http://www.oclc.org/content/dam/research/activities/digpresneeds/digpres.pdf.

Higgins, Sarah. "Digital Curation: The Challenge Driving Convergence across Memory Institutions." Paper presented at the *UNESCO Memory of the World Conference: The Memory of the World in the Digital Age: Digitization and Preservation*, Vancouver, BC, Canada, 26-28 September 2012.

Hirtle, Peter. "The History and Current State of Digital Preservation in the United States." *Metadata and Digital Collections, a Festschrift in Honor of Tom Turner*. 2008. Accessed August 12, 2015. Digital version. http://cip.cornell.edu/cul.pub/1238609304.

Humanities Advanced Technology and Information Institute, University of Glasgow and National Initiative for a Networked Cultural Heritage (NINCH). *The NINCH Guide to Good Practice in the Digital Representation and Management of Cultural Heritage Materials*. 2003. Digital version. http://www.ninch.org/guide.pdf.

Hutchings, Matthew. "Testing Software Tools of Potential Interest for Digital Preservation Activities at the National Library of Australia." *Information and Technology Division of the National Library of Australia*. July 30, 2012. Digital version. http://openpreservation.org/system/files/Digital%20Preservation%20Project%20Report%20-%20Testing%20Software%20Tools.pdf

Iljon, Ariane. "Convergence of Archives, Libraries, and Museums: A European Perspective." From *Delivering Diversity; Promoting Participation Conference Proceedings* hosted by MDA, ICOM, and CIDOC, London, England, 6-10 September 1999: 23-25.

Ippolito, Jon and Richard Rinehart. *Re-Collection: Art, New Media, and Social Memory*. 2014. Leonardo Series. The MIT Press: Cambridge, MA.

International Council on Archives and InterPARES."Digital Records Pathways: Topics in Digital Preservation, Module 2 Developing Policies and Procedures for Digital Preservation." *Interpares.org*. Drafted July 2012. Digital document. Accessed August 3, 2015. http://www.interpares.org/ip3/display_file.cfm?doc=ip3_canada_gs12_module_2_july-2012_DRAFT.pdf.

International Council of Museums. *ICOM Code of Ethics for Museums*. ICOM, 2013. ISBN-978-92-9012-407-8.

Internet Archive. "Metropolitan Museum of Art - Gallery Images." *Archive.org*. 2014. Accessed October 1, 2015. https://archive.org/details/metropolitanmuseumofart-gallery&tab=collection.

_____ "The Metropolitan Museum of Art, Library." *Archive.org*. 2015. Accessed October 1, 2015. https://archive.org/details/metmuseumlibraries&tab=collection.

_____ (b). "About the Internet Archive." *Archive.org*. 2015. Accessed October 1, 2015. https://archive.org/about/.

InterPARES 2 Project. "Terminology Database." 2015. Accessed September 20, 2015. http://www.interpares.org/ip2/ip2_terminology_db.cfm.

"JHOVE2, The Next-Generation Architecture for Format-Aware Characterization." *GitHub*. 2015. Accessed September 20, 2015. https://github.com/opf-labs/jhove2.

Joint Information Systems Committee. "Definition of Digital Preservation," in *JISC Beginner's Guide to Digital Preservation*. Last updated 2012. Accessed 20 July, 2015. http://blogs.ukoln.ac.uk/jisc-beg-dig-pres/.

Kenney, Anne R. and Nancy Y. McGovern. "The Five Organizational Stages of Digital Preservation." In *Digital libraries: a vision for the 21st century : a festschrift in honor of Wendy Lougee on the occasion of her departure from the University of Michigan,* edited by Patricia Hodges, Maria Bonn, Mark Sandler, and John Price Wilkin. Ann Arbor, MI: Michigan Publishing, University of Michigan Library, 2003. Digital version. http://dx.doi.org/10.3998/spobooks.bbv9812.0001.001.

Kirchhoff, Thomas, Werner Schweibenz, and Jörn Sieglerschmidt. "Archives, Libraries, Museums and the Spell of Ubiquitous Knowledge." *Archival Science* 8, no. 4 (2009): 251-66.

Kott, Katherine and Paula Jabloner. *Computer History Museum Digital Repository Policies 0.1.* March 2012. Digital document. Accessed February 2015. http://www.computerhistory.org/atchm/wp-content/uploads/2012/09/Digital_Repository_Policy-DR1.0.pdf.

Lavoie, Brian, and Günter Waibel. *An Art Resource in New York: The Collective Collection of the NYARC Art Museum Libraries*. Report produced by OCLC Programs and Research. 2008. Published online at: www.oclc.org/programs/publications/reports/2008-02.pdf.

Lazorchak, Butch. "DAMs vs. LAMs: It's On!" *The Signal* blog, October 17, 2012. Accessed October 24, 2015. http://blogs.loc.gov/digitalpreservation/2012/10/dams-vs-lams-its-on/.

_____. "Digital Preservation, Digital Curation, Digital Stewardship: What's in (Some) Names?" *The Signal Blog* of the Library of Congress, August 23, 2011. Accessed July 24, 2015. http://blogs.loc.gov/digitalpreservation/2011/08/digital-preservation-digital-curation-digital-stewardship-what's-in-some-names/.

LeFurgy, Bill. "Facing Off with Digital Preservation Policy." *The Signal Blog* from The Library of Congress, July 6, 2011, http://blogs.loc.gov/digitalpreservation/2011/07/facing-off-with-digital-preservation-policy/.

Library of Congress. "About." Digital Preservation (Library of Congress). 2015a. Accessed July 19, 2015. http://www.digitalpreservation.gov/about/.

_____(b). "Recommended File Formats." *Library of Congress Preservation Resources*. Last updated 2015. Accessed August 2, 2015. http://www.loc.gov/preservation/resources/rfs/index.html.

"Life Cycle Information for E-Literature Project (LIFE)." 2015. Accessed November 5, 2015. http://www.life.ac.uk/.

Lipcan, Dan. Interview by Emma P. James. Digital interview conducted Summer 2015.

MacDonald, George F. and Stephen Alsford. "The Museum as Information Utility." In *Museums in the Digital Age*, edited by Ross Parry, 72-79. New York: Routledge. 2010.

Malaro, Marie C. and Ildiko Pogany DeAngelis. *A Legal Primer on Managing Museum Collections Third Edition*. Washington D.C: Smithsonian Books, 2012.

Malaro, Marie C. *Museum Governance: Mission, Ethics, Policy*. Washington, District of Columbia: Smithsonian Institution Press, 1994.

Manus, Susan. "At the Museum: An Interview with Marla Misunas of SFMOMA, Pt.1," *The Signal* blog. April 2, 2014. Accessed March 3, 2015. http://blogs.loc.gov/digitalpreservation/2014/04/at-the-museum-an-interview-with-marla-misunas-of-sfmoma-pt-1/.

_____ (b). "At the Museum: An Interview with Marla Misunas (and Friends) of SFMOMA, Pt.2," *The Signal* blog. June 19, 2014. Accessed March 3, 2015. http://blogs.loc.gov/digitalpreservation/2014/06/at-the-museum-an-interview-with-marla-misunas-and-friends-of-sfmoma-pt-two/.

Martin, Robert S. "Intersecting Missions, Converging Practice." *RBM: A Journal of Rare Books, Manuscripts, and Cultural Heritage* 8, no. 1 (2007): 80-88.

McGovern, Nancy."Digital Preservation Policy Framework: Development Guideline Version 2.1." *Digital Preservation Toolkit* for the *Canadian Heritage Information Network*. Last modified April 25, 2013. Accessed March 2015. http://www.rcip-chin.gc.ca/carrefour-du-savoir-knowledge-exchange/digital_preservation_policy_guidelines-ligne_directrice_strategique_preservation_numerique-eng.jsp.

_____. *Version 2.0 Digital Preservation Policy Framework: Outline*. Prepared for the Inter-University Consortium for Political and Social Research (ICPSR). Last revised October 2007. Accessed July 13, 2015. Digital version. http://www.icpsr.umich.edu/files/ICPSR/curation/preservation/policies/dp-policy-outline.pdf.

McHugh, Andrew and Perla Innocenti, Seamus Ross, Raivo Ruuselepp. "Risk management foundations for digital libraries: DRAMBORA (Digital Repository Audit Method Based on Risk Assessment)." 2008.

MetaArchive Cooperative. "The Cooperative: A brief history of the first private digital preservation network." 2014. Accessed October 12, 2015. http://www.metaarchive.org/the-cooperative.

_____ (b). "Methodology: don't put all your eggs in one basket." 2014. Accessed October 12, 2015. http://www.metaarchive.org/methodology.

_____ (c). "Costs." 2014. Accessed October 12, 2015. http://www.metaarchive.org/costs.

_____ (d). "Our Members: Membership Map." 2014. Accessed October 12, 2015. http://www.metaarchive.org/members.

_____. "Preservation Policy Template: Digital Preservation Policy & Planning Workshop." *Metaarchive.org*. October 15, 2010. Accessed October 2015. Digital version. http://metaarchive.org/public/resources/pres_comm/policy_planning/Digital_Preservation_Policy_Template.pdf.

"Metadata Encoding & Transmission Standard (METS)." *Library of Congress*. Last updated November 23, 2015. Accessed September 12, 2015. http://www.loc.gov/standards/mets/.

MIT Libraries. Digital Preservation Workshop. "Chamber of Horrors: Obsolete and Endangered Media." *Digital Preservation Management: Implementing Short term Strategies for Long Term Management*. Last updated 2012. Accessed April 15, 2015. http://www.dpworkshop.org/dpm-eng/oldmedia/tapes.html

_____ (b). Digital Preservation Workshop. "Timeline." *Digital Preservation Management: Implementing Short term Strategies for Long Term Management*. Last updated 2012. Accessed April 15, 2015. http://www.dpworkshop.org/dpm-eng/timeline/viewall.html#1980.

_____ (c). Digital Preservation Workshop. "Terms and Concepts: Digital Preservation." *Digital Preservation Management: Implementing Short term Strategies for Long Term Management*. Last updated 2012. Accessed April 15, 2015. http://dpworkshop.org/dpm-eng/terminology/preservation.html.

_____ (d). Digital Preservation Workshop. "Digital Preservation Strategies." *Digital Preservation Management: Implementing Short term Strategies for Long Term Management*. Last updated 2012. Accessed June 20, 2015. http://dpworkshop.org/dpm-eng/terminology/strategies.html.

Museum of Modern Art. "About MoMA - Curatorial Departments." *Press.moma.org*. 2011. Accessed October 5, 2015. http://press.moma.org/curatorial-departments/.

_____. "Museum History." *Moma.org.* 2015. Accessed October 5, 2015. http://www.moma.org/about/history.

_____ (b). "Archives." *Moma.org.* 2015. Accessed October 5, 2015. http://www.moma.org/learn/resources/archives/index.

_____ (c). "Film Preservation Center." *Moma.org.* 2015. Accessed October 5, 2015. http://www.moma.org/learn/resources/filmpreservation#historypreservcenter.

_____ (d). "About MoMA." *Moma.org.* 2015. Accessed October 5, 2015. http://www.moma.org/about/.

_____(e). "The Collection." *Moma.org.* 2015. Accessed October 5, 2015. http://www.moma.org/collection.

_____(f). "About The Collection." *Moma.org.* 2015. Accessed October 5, 2015. http://www.moma.org/collection/about

_____(g). "Library." *Moma.org.* 2015. Accessed October 5, 2015. http://www.moma.org/learn/resources/library/index.

_____ (h). "Dadabase FAQs." *Moma.org.* 2015. Accessed October 5, 2015. http://www.moma.org/learn/resources/library/faq_dadabase#digitized.

Murray, Kate. "Preserving Digital and Software-Based Artworks: Recap of a NDSA Discussion," *The Signal* Blog, June 13, 2014. Accessed September 15, 2015. http://blogs.loc.gov/digitalpreservation/2014/06/preserving-digital-and-software-based-artworks-recap-of-a-ndsa-discussion/.

National Digital Stewardship Alliance (NDSA). "Digital Preservation in a Box." Last updated 2015. Accessed September 22, 2015. http://dpoutreach.net/.

_____. "NDSA Levels of Digital Preservation." *Library of Congress Digital Preservation.* Accessed 15 September 2015. http://www.digitalpreservation.gov/ndsa/activities/levels.html.

NDSA Standards & Practice and Infrastructures Working Group. "Checking Your Digital Content: What is Fixity, and When Should I be Checking It?" 2014. Accessed September 20, 2015. Digital version. http://hdl.loc.gov/loc.gdc/lcpub.2013655117.1.

National Initiative for a Networked Cultural Heritage (b). "About NINCH." 2003. Accessed October 12, 2015..http://www.ninch.org/about/.

National Information Standards Organization (NISO). *Understanding Metadata.* Bethesda, MD: NISO Press, 2004.

National Library of Australia. "Encapsulation." *Preserving Access to Digital Information.* 2001. Accessed September 14, 2015. http://pandora.nla.gov.au/pan/10691/20110824-1153/www.nla.gov.au/padi/topics/20.html.

New York Art Resources Consortium. "Web Archiving." *Nyarc.org.* 2015. Accessed October 6, 2015. http://www.nyarc.org/content/web-archiving.

NetX. "Museum Digital Asset Management at MOMA NY." Accessed October 1, 2015. http://netx.net/portfolio/museum-digital-asset-management-moma/.

_____. "New Functions and Features." *Net Exposure Blog*, January 28, 2013. Accessed October 1, 2015. http://netx.net/media-asset-management/.

"New Art Trust Names John R. Lane as President and CEO." ArtDaily.com. October 8, 2008. Accessed September 13, 2015. http://artdaily.com/news/26451/New-Art-Trust-Names-John-R--Lane-as-President-and-CEO#.VfW2c2RViko.

Northeast Document Conservation Center. "NEDCC Digital Preservation Policy Template." *Digital Preservation Resources.* 2008. Digital version. https://www.nedcc.org/assets/media/documents/SoDAExerciseToolkit.pdf.

Novak, Audrey. "Fixity Checks: Checksums, Message Digests, and Digital Signatures." From the Digital Preservation Committee at Yale University Library. November 2006. http://www.library.yale.edu/iac/DPC/AN_DPC_FixityChecksFinal11.pdf .

Novia, Jennifer. "Library, Archive, and Museum (LAM) Collaboration: Driving Forces and Recent Trends." *Endnotes: The Journal of the New Members Round Table*, 3, no. 1 (October 2012). Accessed September 25, 2015. Digital version. http://www.ala.org/nmrt/sites/ala.org.nmrt/files/content/oversightgroups/comm/schres/endnotesvol3no1/2lamcollaboration.pdf.

"OAIS 2: the Information Package." *Alan's Thoughts on Digital Preservation Blog.* January 16, 2008. Accessed September 10, 2015. https://alanake.wordpress.com/2008/01/16/oais-2-the-information-package/.

"OAIS 3: the Submission Information Package." *Alan's Thoughts on Digital Preservation Blog.* January 16, 2008 (b). Accessed September 10, 2015. https://alanake.wordpress.com/2008/01/16/oais-3-the-submission-information-package/

"OAIS 9: Information Flow Processes." *Alan's Thoughts on Digital Preservation Blog.* February 1, 2008(c). Accessed September 10, 2015. https://alanake.wordpress.com/2008/02/01/oais-9-information-flow-processes/.

Ockerbloom, John Mark. "What Repositories Do: The OAIS Model." *Everybody's Libraries Blog*, October 13, 2008. Accessed August 25, 2015. http://everybodyslibraries.com/2008/10/13/what-repositories-do-the-oais-model/.

Online Computer Library Center (OCLC). "Contentdm." 2015. Accessed October 1, 2015. http://www.oclc.org/en-US/contentdm/features.html.

_____. "Library, Archive, and Museum Collaboration." *OCLC Research.* Last updated November 30, 2011. Accessed September 30, 2015. http://www.oclc.org/research/activities/lamsurvey.html.

The OCLC/RLG Working Group on Preservation Metadata. *Preservation Metadata and the OAIS Reference Model: A Metadata Framework to Support the Preservation of Digital Objects.* Dublin, Ohio: OCLC, 2002. http://www.oclc.org/research/projects/pmwg/.

Online Computer Library Center, Inc., and the Center for Research Libraries. "Trustworthy Repositories Audit and Certification: Criteria and Checklist." Chicago and Dublin, Ohio: *OCLC and the Center for Research Libraries*, 2007. Digital version. http://www.crl.edu/PDF/trac.pdf

Oleksik, Peter. "Digitizing MoMA's Video Collection." *Inside/Out blog* for *Moma.org*. April 8, 2015. Accessed August 23, 2015. http://www.moma.org/explore/inside_out/2015/04/08/digitizing-momas-video-collection.

Owens, Trevor. "Archivematica and the Open Source Mindset for Digital Preservation Systems." *The Signal* blog, October 16, 2012. Accessed October 6, 2015. http://blogs.loc.gov/digitalpreservation/2012/10/archivematica-and-the-open-source-mindset-for-digital-preservation-systems/.

Personal Archives Accessible in Digital Media (Paradigm). "Selecting the Right Preservation Strategy, Other preservation Approaches: Encapsulation." *Paradigm.ac.uk*. Last updated January 2, 2008. Accessed September 14th, 2015. http://www.paradigm.ac.uk/workbook/preservation-strategies/selecting-other.html.

Pogrebin, Robin."The Met's Director Looks Ahead." *The New York Times*, March 19, 2014. Accessed November 19, 2015. http://www.nytimes.com/2014/03/20/arts/artsspecial/the-mets-director-looks-ahead.html?_r=0.

PREMIS Editorial Committee. "PREMIS Data Dictionary for Preservation Metadata, version 2.0." The Library of Congress. March 2008. Digital version. http://www.loc.gov/standards/premis/v2/premis-2-0.pdf.

Prom, Christopher J. "Preserving Email - DPC Technology Watch Report 11-01." *Digital Preservation Coalition* and *Charles Beagrie*. 2011, 1-41. doi:http://dx.doi.org/10.7207/twr11-01.

Renga, Alan and Carin Riney. *SDASM Digital Preservation Policy*. San Diego Air and Space Museum. December 2012.

Rinehart, Amanda Kay, Patrice-Andre Prud'homme, and Andrew Reid Huot. "Overwhelmed to Action: digital preservation challenges at the under-resourced institution." *OCLC Systems & Services* 20, no.1 (2014): 28-42.

Rinehart, Richard. "MOAC - A Report on Integrating Museum and Archive Access in the Online Archive of California." *D-Lib Magazine* 9, no. 1, 2003. Accessed October 20, 2015. doi:10.1045/january2003-rinehart.

RLG-NARA Task Force on Digital Repository Certification. "Trustworthy Repositories Audit and Certification: Criteria and Checklist." February 2007. Accessed September 12, 2015. Digital version. http://www.crl.edu/sites/default/files/d6/attachments/pages/trac_0.pdf.

RLG-OCLC Report. "Trusted Digital Repositories: Attributes and Responsibilities." Research Libraries Group. May 2002. Accessed September 12, 2015. http://www.oclc.org/content/dam/research/activities/trustedrep/repositories.pdf.

Sanchez, Crystal and James Smith. "Interview with Glenn Wharton." *The Smithsonian Interview Project: Questions on Technical Standards in the Care of Time-Based and Digital Art*, April 24, 2013. Accessed October 6, 2015. Published online at: http://www.si.edu/content/tbma/documents/transcripts/GlennWharton_130424.pdf.

_____. "Interview with Mark Heller." *The Smithsonian Interview Project: Questions on Technical Standards in the Care of Time-Based and Digital Art*, June 14, 2013. Accessed 15 September 2015. Digital version. http://www.si.edu/content/tbma/documents/transcripts/MarkHellar_130614.pdf.

_____ (b). "Interview with Ben Fino-Radin." *The Smithsonian Interview Project: Questions on Technical Standards in the Care of Time-Based and Digital Art*, April 26, 2013. Accessed October 6, 2015. Published online at: http://www.si.edu/content/tbma/documents/transcripts/benFino-Radin_130426.pdf.

San Francisco Museum of Modern Art. "About SFMOMA." *Sfmoma.org*. Accessed September 12, 2015. http://www.sfmoma.org/about/about_sfmoma#ixzz3lXQwlI7q.

_____ (b). "Photography Collection." *Sfmoma.org*. Accessed September 12, 2015. http://www.sfmoma.org/explore/collection/photography#ixzz3lXS8nMiX.

_____ (c). "Explore The Collection." *Sfmoma.org*. Accessed September 12, 2015. http://www.sfmoma.org/explore/collection#ixzz3lXTdv2Jf.

_____ (d). "Matters in Media Art." *Sfmoma.org*. Accessed September 13, 2015. http://www.sfmoma.org/about/research_projects/research_projects_matters_in_media#ixzz3ldqoZ6HF.

_____ (e). "Research and Projects." *Sfmoma.org*. Accessed September 15, 2015. https://www.sfmoma.org/about/research_projects.

Sawyer, Don. "The Open Archival Information System and the NSSDC." *NSSDC News* for *NASA* 16, no.4, December 2000. Accessed October 12, 2015. Digital version. http://nssdc.gsfc.nasa.gov/nssdc_news/dec00/oais.html.

"SCAPE Catalogue of Digital Preservation Policy Elements." Last modified May 19, 2014. Accessed October 14, 2015. http://wiki.opf-labs.org/display/SP/Catalogue+of+Preservation+Policy+Elements.

"SCAPE Published Preservation Policies." Last modified December 1, 2015. Accessed October 14, 2015. http://wiki.opf-labs.org/display/SP/Published+Preservation+Policies.

Sheldon, Madeline. *Analysis of Current Digital Preservation Policies: Archives, Libraries, and Museums*. Report created for The Library of Congress, July 22, 2013.

Sinclair, Pauline. *A Planets White Paper: The Digital Divide, Assessing Organizations Preparations for Digital Preservation*. Tessella. March 2010. Digital document. http://www.planets-project.eu/docs/reports/planets-market-survey-white-paper.pdf.

Smith, Abby. "The Digital Preservation Conundrum, Part 1." In *The Serials Librarian* 46, no. 1/2 (2004): 107-13.

Smithsonian Institution. *Digitization and Digital Asset Management Policy, Smithsonian Directive 610*. March 31, 2011. Digital document. Accessed October 10, 2015. http://www.si.edu/content/pdf/about/sd/SD610.pdf.

_____. *Smithsonian Institution Information Technology Plan, FY 2012 to FY 2016*. Office of the Chief Information Officer. 2012. Accessed October 10, 2015. Digital document. http://www.si.edu/content/ocio/pdfs/SITP.pdf.

Society of American Archivists (SAA). "Appraisal." *Glossary of Archival and Records Terminology*. 2015. http://www2.archivists.org/glossary/terms/a/appraisal.

_____(b). "Encapsulation." *Glossary of Archival and Records Terminology* from *Archivists.org*. 2015. Accessed September 13, 2015. http://www2.archivists.org/glossary/terms/e/encapsulation

_____. "The Preservation of Digitized Reproductions." SAA: The Preservation of Digitized Reproductions. 1997. Accessed August 10, 2015. http://www.archivists.org/statements/digitize.asp.

"Standards at the Library of Congress." *The Library of Congress*. Last Updated July 13, 2015. Accessed September 15, 2015. http://www.loc.gov/standards/.

Staudeman, Sarah and Paul Messier. "Video Format Identification Guide." *Video Preservation Website (VPW)* of Stanford University. 2007. Accessed April 15, 2015. http://videopreservation.conservation-us.org/vid_id/index.html

Sterling, Bruce."Introduction: Digital Decay." In *Permanence through Change the Variable Media Approach*. 2003. New York, New York: Guggenheim Museum Publications, 10-22.

Task Force on Archiving Digital Information. "Preserving Digital Information: Report of the Task Force on Archiving of Digital Information." Commissioned by the *Commission on Preservation and Access and the Research Libraries Group*, May 1, 1996. Digital version. http://www.oclc.org/content/dam/research/activities/digpresstudy/final-report.pdf

Tate Modern. "Matters in Media Art." *Tate.org*. Last updated December 2015. Accessed September 13, 2015. http://www.tate.org.uk/about/projects/matters-media-art.

_____ (b). "Post-Acquisitions." *Matters in Media Art from Tate.org*. Last updated December 2015. Accessed September 13, 2015. http://www.tate.org.uk/about/projects/matters-media-art/acquisitions/post-acquisitions

_____ (c). "Lending Time-Based Media Art." *Matters in Media Art from Tate.org*. Last updated December 2015. Accessed September 13, 2015. http://www.tate.org.uk/about/projects/matters-media-art/lending-time-based-media-2005.

_____ (d). "External Resources." *Matters in Media Art* from *Tate.org*. Accessed September 14, 2015. http://www.tate.org.uk/about/projects/matters-media-art/acquisitions/external-resources.

_____. "Tate Digital Preservation and Continuity Policy and Strategy 2011." Cited in the *Tate Records Management Policy*: p.3, November 2013. Accessed October 2015. Digital version. http://www.tate.org.uk/download/file/fid/37165.

The Metropolitan Museum of Modern Art (MET). "Main Building." *History of the Museum* from *Metmuseum.org*. 2015. Accessed September 27, 2015. http://www.metmuseum.org/about-the-museum/history-of-the-museum/main-building.

_____ (b). "Museum Mission Statement." *Metmuseum.org*. 2015. Accessed September 27, 2015. .http://www.metmuseum.org/about-the-museum/mission-statement.

_____ (c) "Thomas J. Watson Library." *Museum Departments, Office of the Director* from *Metmuseum.org*. 2015. Accessed September 27, 2015. http://www.metmuseum.org/about-the-museum/museum-departments/office-of-the-director/thomas-j-watson-library

_____ (d) "Museum Archives." *Museum Departments, Office of the Director* from *Metmuseum.org*. 2015. Accessed September 27, 2015. http://www.metmuseum.org/about-the-museum/museum-departments/office-of-the-president/archives.

_____ (e) "The Collections Online." *Metmuseum.org*. 2015. Accessed September 27, 2015.http://www.metmuseum.org/collection/the-collection-online.

_____ (f) "Collection." *Metmuseum.org*. 2015. Accessed September 27, 2015.http://www.metmuseum.org/collection.

Thomas J. Watson Library. "Digital Collections from The Metropolitan Museum of Art Libraries." *Digital Collections, Thomas J. Watson Library, The Metropolitan Museum of Art*. 2015. http://libmma.contentdm.oclc.org/cdm.

Trant, Jennifer. "Emerging Convergence? Thoughts on museums, archives, and libraries, and professional training." In *Museum Management and Curatorship*, 24, no. 4 (December 2009): 369-387.

Tristram, Claire. "Data Extinction." *MIT Technology Review*, October 1, 2002. Accessed September 14, 2015. http://www.technologyreview.com/featuredstory/401682/data-extinction/.

UC Berkeley School of Information Management and Systems. "Executive Summary." *How Much Information? 2003.* http://groups.ischool.berkeley.edu/archive/how-much-info-2003/execsum.htm#summary

The UK National Archives."Download DROID: file format identification tool." *The National Archives*. 2015. Accessed September 10, 2015. http://www.nationalarchives.gov.uk/information-management/manage-information/preserving-digital-records/droid/.

Wall, Helen D. "Picturing Met Museum through Visitor's Eyes." *Digital Underground* blog, April 16, 2015. Accessed September 28, 2015. http://www.metmuseum.org/about-the-museum/museum-departments/office-of-the-director/digital-media-department/digital-underground/2015/picture-met-museum.

Wheatley, P., Andy Jackson, and Andy Tester (contributors). "Main Page: Community Owned digital Preservation Tool Registry." *COPTR*. Last updated November 28, 2014. Accessed March 9, 2015. http://coptr.digipres.org/Main_Page.

White, Layna. Interviewed by Emma P. James. In-person interview at the San Francisco Museum of Modern Art, San Francisco, CA. May 6, 2015.

"World Wide Web Consortium Process Document." World Wide Web Consortium (W3C). September 1, 2015. Accessed September 10, 2015. http://www.w3.org/2015/Process-20150901/.

Yarrow, Alexandra, and Barbara Clubb, Jennifer-Lynn Draper. "Public Libraries, Archives and Museums: Trends in Collaboration and Cooperation." From *IFLA Professional Reports*: 108. The Hague: International Federation of Library Associations and Institutions, 2008. Accessed September 22, 2015. http://archive.ifla.org/VII/s8/pub/Profrep108.pdf.

Yeung, Tim Au. "Digital Preservation: Best Practice for Museums."(commissioned by the Canadian Heritage Information Network). Minister of Public Works and Government Services, Canada, 2004

Zorich, D., G. Waibel, and R. Erway. *Beyond the Silos of the LAMs: Collaboration Among Libraries, Archives and Museums*. 2008. Accessed October 16, 2015. Digital version. http://www.oclc.org/programs/reports/2008-05.pdf.

Appendix A: Glossary of Useful Terms

AIP (Archival Information Package) - an information package that is preserved within an Open Archival Information System (OAIS) digital repository

API (application programming interface) - a specification for an interface that allows software components to communicate, and typically used by software developers to enable different software tools to interoperate.

Authenticity - the quality of trustworthiness of a record - in this context a digital object. Authenticity provides the assurance that a record is what it purports to be and has demonstrably not been tampered with or otherwise corrupted.

Bit - the fundamental unit of digital information storage, which can have a binary value of either 1 or 0

Bitstream - a sequence of bytes, which has a meaningful common properties for the purposes of preservation. A bitstream may be a file or a component of a file.

Bitstream preservation - the aspect of preservation management that is concerned with maintaining the integrity of every bitstream ingested into the digital repository, by ensuring that a demonstrably bit-perfect copy can be retrieved on demand, for as long as required.

Byte - a unit of digital information and measure of data volume, normally equivalent to **eight bits**

Characterization - the aspect of logical preservation that is concerned with understanding the nature of digital objects, including their technical and significant properties.

Checksum - a value calculated by an algorithm based on the the bit-level content of a file, such that any change to that content will result in a different checksum value. Checksums can therefore be used to detect changes to data, and hence perform integrity checks

Digital asset register - a record of an organization's digital information assets, which quantifies the value and risk of loss in each one

Digital linear tape - a common format of magnetic tape data storage technology

Digital repository - a combination of people, processes, and technologies, which together provide the means to capture, preserve, and provide access to digital objects

DIP (dissemination information package) - an information package, derived from one or more AIPs and supplied to an end-user by an OAIS digital repository as a result of a request for access

Disk image - a bit-level copy of a digital storage device, such as a hard disk, usually encoded in a single file

DTD (document type definition) - a formal syntax for defining a document type in XML or HTML

Emulation - the class of preservation actions that entail transforming a technology environment to allow a digital object to be accessed in its original form

Endianness - the ordering or sequencing of bytes of a word of digital data in computer memory storage or during transmission. Words may be represented in big-endian or little-endian manner. Big-endian systems store the most significant byte of a word at the smallest memory address and the least significant byte at the largest. A little-endian system, in contrast, stores the least significant byte at the smallest address.

Exabyte - a unit of measurement of data volume, equivalent to 1000 petabytes

Extensible Markup Language (XML) - a markup language for encoding information in human-readable and machine-readable form

File - a bitstream which is managed by a file system as a single, named entity

File Transfer Protocol (FTP) - a protocol for transferring digital files across a network

Fixity (integrity) - the aspect of an information object's authenticity that depends on it being protected against unauthorized or accidental alteration

Format - a predefined structure for organizing a file or bitstream

Information Package - a logical container defined by OAIS, and composed of an information object (content information) and associated preservation description information

Ingest - The final stage of accession, in which one or more AIPs are generated from a Submission Information Package (SIP) and stored in a digital repository. Physically this requires the files to be moved into a permanent storage location within repository control, and the metadata to be incorporated into the relevant metadata management regime

Integrity checking - the process of testing the integrity of a data object, typically using a checksum. this is key aspect of bitstream preservation

Linked open data - a method of publishing structured data using standard web technologies, so that it can be linked together for machine processing

Logical preservation - the aspect of preservation management that is concerned with ensuring the continued usability of meaningful information content, by ensuring the existence of a usable manifestation of an information object

Manifestation - a specific data object that instantiates and information object. Multiple manifestations can exist for any given information object.

Metadata - the set of information required to enable content to be discovered, managed, and used by both human agents and automated systems. Literally "data about data"

METS (metadata encoding and transmission standard) - a widely adopted metadata standard for encoding descriptive, administrative and structural metadata

Migration - the class of preservation actions that entail transforming a digital object into a form which can be accessed in a new technology environment

Migration pathway - a specific migration process for transforming between a source and target format of a data object

Normalization - the process of migrating digital objects to new formats at the point of ingest, in order to minimize the number of formats to be managed within a repository

Open Archives Information System (OAIS) reference model : an international standards (ISO 14721:2003) defining a high level functional model for a digital repository

Persistent identifier - a reference to a digital object which uniquely refers to it, and can be relied on to remain meaningful (capable of being interpreted as referring to that object) for at least as long as the object itself exists

Petabyte - a unit of measurement of data volume, equivalent to 1,000 terabytes

PREMIS - a preservation metadata scheme, now an international de facto standard

Preservation action - the process of enacting and validating a preservation plan. This forms the final stage of logical preservation, and results in the generation of a new AIP. Two major classes of preservation action are migration and emulation.

Preservation description information - the info that is required to preservat an information object in an OAIS digital repository, and which comprises provenance, reference, fixity, context and access rights information

Preservation planning - the aspect of logical preservation that is concerned with identifying threats to the continued availability and usability of authentic digital objects and if such threats are identified, determining appropriate countermeasures. It incorporates the process of technology watch.

Pipeline - a set of data processing elements connected in series, where the output of one element is the input of the next one. The elements of a pipeline are often executed in parallel or in time-sliced fashion.

Quarantine - A process that occurs during accession, whereby a SIP is isolated from other systems until it has been confirmed to be free from any malicious software.

Refreshing - The process of copying data from one storage device to another, of the same or different type, for the purposes of *bitstream preservation* (see above).

Reliability - The aspect of an information object's authenticity that depends on it being a full and accurate representation of the cultural or business activity to which it attests. This requires the establishment of trust in the curatorial processes used to manage the object throughout its lifecycle, and the continued ability to place the object within its original context.

Representation Information - The set of information required to interpret a data object as a meaningful information object, or a component of a technical environment that supports interpretation of that object.

Submission Information Package - An information package that is supplied for ingest into an OAIS digital repository. The ingest process results in the creation of one or more AIPs from the SIP.

Uniform Resource Identifier (URI) - A protocol for identifying networked resources such as web content.

Uniform Resource Locator (URL) - A type of URI that identifies the resource and its location. It therefore acts as an address for networked resources such as web content.

Definitions Sourced From:

Brown, Adrian. *Practical Digital Preservation: A How-to Guide for Organizations of Any Size*. 2013. Chicago, IL: Neal-Schuman, p. xi-xvi.

Wikipedia. "Endianness." 2015. Accessed July 20, 2015. https://en.wikipedia.org/wiki/Endianness.

Wikipedia. "Pipeline (computing)." 2015. Accessed July 20, 2015. https://en.wikipedia.org/wiki/Pipeline_(computing).

Appendix B: National Digital Stewardship Alliance, Levels of Digital Preservation

Table 1: Version 1 of the Levels of Digital Preservation

	Level 1 (Protect your data)	Level 2 (Know your data)	Level 3 (Monitor your data)	Level 4 (Repair your data)
Storage and Geographic Location	- Two complete copies that are not collocated - For data on heterogeneous media (optical discs, hard drives, etc.) get the content off the medium and into your storage system	- At least three complete copies - At least one copy in a different geographic location - Document your storage system(s) and storage media and what you need to use them	- At least one copy in a geographic location with a different disaster threat - Obsolescence monitoring process for your storage system(s) and media	- At least three copies in geographic locations with different disaster threats - Have a comprehensive plan in place that will keep files and metadata on currently accessible media or systems
File Fixity and Data Integrity	- Check file fixity on ingest if it has been provided with the content - Create fixity info if it wasn't provided with the content	- Check fixity on all ingests - Use write-blockers when working with original media - Virus-check high risk content	- Check fixity of content at fixed intervals - Maintain logs of fixity info; supply audit on demand - Ability to detect corrupt data - Virus-check all content	- Check fixity of all content in response to specific events or activities - Ability to replace/repair corrupted data - Ensure no one person has write access to all copies
Information Security	- Identify who has read, write, move and delete authorization to individual files - Restrict who has those authorizations to individual files	- Document access restrictions for content	- Maintain logs of who performed what actions on files, including deletions and preservation actions	- Perform audit of logs
Metadata	- Inventory of content and its storage location - Ensure backup and non-collocation of inventory	- Store administrative metadata - Store transformative metadata and log events	- Store standard technical and descriptive metadata	- Store standard preservation metadata
File Formats	- When you can give input into the creation of digital files encourage use of a limited set of known open formats and codecs	- Inventory of file formats in use	- Monitor file format obsolescence issues	- Perform format migrations, emulation and similar activities as needed

National Digital Stewardship Alliance, "Levels of Digital Preservation." *Library of Congress*. 2014. Accessed May 2, 2015. http://www.digitalpreservation.gov/ndsa/activities/levels.htm.

Appendix C: Case Study Interview Questions

1. How did your digital preservation program start?
2. How many staff members work in digital preservation and/or digital asset management in your unit?
3. What resources does your unit draw upon in dealing with plans/policies for managing the unique digital assets that museums possess?
4. Does your museum follow (or strive to follow) TRAC or ISO 16363 requirements, or something similar? How was the decision made in your unit?
5. Can you briefly describe your institution's selection process for digital preservation? How does your unit prioritize collections to be preserved?
6. Do you use a digital preservation system of any kind? If so please describe it and outline why your institution chose that particular solution.

a. Do you migrate digital objects from original media? Describe the process.
b. What metadata schema does your museum use to describe digital objects?

7. Does your unit normalize files to preservation and access formats upon ingest? What formats do you use?
8. Storage: Where do you store your digital objects/assets? How many backup copies do you keep? Who has access?
9. DAMS and Digital Preservation: How does your museum's digital preservation system interface with your museum's digital asset management system?
10. How does your unit communicate its work/efforts in the area of digital preservation to other departments in the museum? In addition, how do other departments communicate with your unit to provide digital materials for the preservation system?
11. Does your museum preserve emails, museum website, OPAC, or social media?
12. How does your museum (or do you plan to?) provide access to digital collections?
13. How is your digital preservation program funded? Grants or from the museum's overhead budget?

14. Do you have a digital preservation policy or statement? If so, what resources did you use to create it? Could I have a copy?
15. What are your museum's future plans for digital preservation?

Appendix D: Alphabetical List of Websites Cited

D.1 Allinson, Julie. "OAIS." *JISC Standards Catalogue*. Updated September 23, 2006.

D.2 American Alliance of Museums (AAM). "Continuum of Excellence." *Assessment Programs*. Last updated 2015.

D.3 American Alliance of Museums (AAM) (b). "Assessment Types." *Museum Assessment Program* (MAP). Last updated 2015.

D.4 Anderson, Martha. "B is for Bit Preservation," *The Signal* blog, September 7, 2011.

D.5 Archive-It. "About Us." *Archive-It blog*. 2014.

D.6 Archive-it (b). "New York Art Resources Consortium (NYARC)." *Nyarc.org*. 2014.

D.7 Archivematica. "Archivematica: preserving memory since 2009." *Artefactual Systems Inc*. 2015.

D.8 Archivists' Toolkit. "Archivists' Toolkit." 2009.

D.9 Artefactual Systems."User Manual -What is Binder?" Last updated May 22, 2015.

D.10 Artefactual Systems."User Manual -Current Project Status." Last updated May 22, 2015.

D.11 ArtStor. "The Metropolitan Museum of Art." *Artstor.org*. Accessed September 26, 2015.

D.12 Blue Ribbon Task Force on Sustainable Digital Preservation and Access. "Index." 2008.

D.13 Canadian Heritage Information Network (CHIN). "Digital Preservation Plan Framework for Museums." *Canadian Heritage Information Network*. Last updated June 20, 2013.

D.14 Canadian Heritage Information Network, 2011 Survey Results. 2015 [b].

D.15 Canadian Heritage Information Network (CHIN) [c]. "CHIN Introduces a Digital Preservation Toolkit." Last modified June 12, 2013.

D.16 Cohen, Daniel J. and Roy Rosenzweig. "Preserving Digital History: The Fragility of Digital Materials." *Digital History: A Guide to Gathering, Preserving, and Presenting the Past on the Web*. 2005.

D.17 Digital Curation Centre (a). "About the DCC." 2004.

D.18 Digital Curation Centre (b). "What is Digital Curation?" 2015.

D.19 Digital Curation Centre (c). "JSTOR/Harvard Object Validation Environment (JHOVE)." Digital Curation Centre. 2015.

D.20 Digital Curation Centre (d). "List of Metadata Standards." *Digital Curation Centre* Resources. 2015.

D.21 Digital POWRR."Tool Grid." 2013.

D.22 Digital Preservation Coalition (DPC). "Digital Preservation Handbook." 2015.

D.23 Fino-Radin, Ben (b). "MoMA's Digital Art Vault." *Inside/Out blog* for *Moma.org*, April 14, 2015.

D.24 Fleming, Robin and Dan Lipcan. "Digitizing the Library's' Collections: An Introduction," *Now at The Met*, January 5, 2012.

D.25 Griesinger, Peggy. "Preserving the Technical History of Media Works." *Inside/Out* blog for *Moma.org*. May 20, 2015.

D.26 Habing, Thomas."METS, Mods, and PREMIS, Oh My!" *Library of Congress, MODS Official Website*. Presentation for American Library Association 2007 Meeting. Last Updated 2010.

D.27 Internet Archive. "Metropolitan Museum of Art - Gallery Images." *Archive.org*. 2014.

D.28 Internet Archive. "The Metropolitan Museum of Art, Library." *Archive.org*. 2015.

D.29 Internet Archive (b). "About the Internet Archive." *Archive.org*. 2015.

D.30 InterPARES 2 Project. "Terminology Database." 2015.

D.31 "JHOVE2, The Next-Generation Architecture for Format-Aware Characterization." *GitHub*. 2015

D.32 JISC. "Definition of Digital Preservation," in *JISC Beginner's Guide to Digital Preservation*. Last updated 2012.

D.33 LeFurgy, Bill. "Facing Off with Digital Preservation Policy." *The Signal Blog* from The Library of Congress, July 6, 2011.

D.34 Library of Congress. "About." Digital Preservation (Library of Congress). 2015.

D.35 Library of Congress (b). "Recommended File Formats." *Library of Congress Preservation Resources*. Last updated 2015.

D.36 "Life Cycle Information for E-Literature Project (LIFE)." 2015.

D.37 Manus, Susan. "At the Museum: An Interview with Marla Misunas (and Friends) of SFMOMA, Pt.2," *The Signal* blog. June 19, 2014.

D.38 Manus, Susan. "At the Museum: An Interview with Marla Misunas of SFMOMA, Pt.1," *The Signal* blog. April 2, 2014.

D.39 McGovern, Nancy."Digital Preservation Policy Framework: Development Guideline Version 2.1." *Digital Preservation Toolkit* for the *Canadian Heritage Information Network*. Last modified April 25, 2013.

D.40 MetaArchive Cooperative. "The Cooperative: A brief history of the first private digital preservation network." 2014.

D.41 MetaArchive Cooperative (b). "Methodology: don't put all your eggs in one basket." 2014.

D.42 MetaArchive Cooperative (c). "Costs." 2014.

D.43 MetaArchive Cooperative (d). "Our Members: Membership Map." 2014.

D.44 "Metadata Encoding & Transmission Standard (METS)." *Library of Congress*. Last updated November 23, 2015.

D.45 MIT Libraries. Digital Preservation Workshop. "Chamber of Horrors: Obsolete and Endangered Media." *Digital Preservation Management: Implementing Short term Strategies for Long Term Management*. Last updated 2012.

D.46 MIT Libraries (b). Digital Preservation Workshop. "Timeline." *Digital Preservation Management: Implementing Short term Strategies for Long Term Management*. Last updated 2012.

D.47 MIT Libraries (c). Digital Preservation Workshop. "Terms and Concepts: Digital Preservation." *Digital Preservation Management: Implementing Short term Strategies for Long Term Management*. Last updated 2012.

D.48 Murray, Kate. "Preserving Digital and Software-Based Artworks: Recap of a NDSA Discussion," *The Signal* Blog, June 13, 2014.

D.49 Museum of Modern Art. "Museum History." *Moma.org.* 2015.

D.50 Museum of Modern Art. "About MoMA - Curatorial Departments." *Press.moma.org.* 2011.

D.51 Museum of Modern Art (b). "Archives." *Moma.org.* 2015.

D.52 Museum of Modern Art (c). "Film Preservation Center." Moma.org. 2015

D.53 Museum of Modern Art (d). "About MoMA." *Moma.org.* 2015.

D.54 Museum of Modern Art (e). "The Collection." *Moma.org.* 2015.

D.55 Museum of Modern Art (f). "About The Collection." *Moma.org.* 2015.

D.56 Museum of Modern Art (g). "Library." *Moma.org.* 2015.

D.57 Museum of Modern Art (h). "Dadabase FAQs." Moma.org. 2015.

D.58 National Digital Stewardship Alliance (NDSA). "Digital Preservation in a Box." Last updated 2015.

D.59 National Initiative for a Networked Cultural Heritage (b). "About NINCH." 2003.

D.60 National Library of Australia. "Encapsulation." *Preserving Access to Digital Information.* 2001.

D.61 NetX. "Museum Digital Asset Management at MOMA NY."

D.62 NetX. "New Functions and Features." *Net Exposure Blog*, January 28, 2013.

D.63 "New Art Trust Names John R. Lane as President and CEO." ArtDaily.com. October 8, 2008.

D.64 "OAIS 2: the Information Package." *Alan's Thoughts on Digital Preservation Blog.* January 16, 2008.

D.65 "OAIS 3: the Submission Information Package." *Alan's Thoughts on Digital Preservation Blog.* January 16, 2008 (b).

D.66 "OAIS 9: Information Flow Processes." *Alan's Thoughts on Digital Preservation Blog.* February 1, 2008(c).

D.67 Ockerbloom, John Mark. "What Repositories Do: The OAIS Model." *Everybody's Libraries Blog*, October 13, 2008

D.68 New York Art Resources Consortium. "Web Archiving." *Nyarc.org.* 2015.

D.69 Oleksik, Peter. "Digitizing MoMA's Video Collection." *Inside/Out blog* for *Moma.org.* April 8, 2015.

D.70 Online Computer Library Center (OCLC). "Contentdm." 2015.

D.71 Online Computer Library Center (OCLC). "Library, Archive, and Museum Collaboration." *OCLC Research.* Last updated November 30, 2011.

D.72 Owens, Trevor. "Archivematica and the Open Source Mindset for Digital Preservation Systems." *The Signal* blog for *The Library of Congress*, October 16, 2012.

D.73 Personal Archives Accessible in Digital Media (Paradigm). "Selecting the Right Preservation Strategy, Other preservation Approaches: Encapsulation." *Paradigm.ac.uk*. Last updated January 2, 2008.

D.74 Pogrebin, Robin."The Met's Director Looks Ahead." *The New York Times*, March 19, 2014.

D.75 "SCAPE Catalogue of Digital Preservation Policy Elements." Last modified May 19, 2014.

D.76 "SCAPE Published Preservation Policies." Last modified December 1, 2015.

D.77 San Francisco Museum of Modern Art. "About SFMOMA." *Sfmoma.org*.

D.78 San Francisco Museum of Modern Art (b). "Photography Collection." *Sfmoma.org*.

D.79 San Francisco Museum of Modern Art (c). "Explore The Collection." *Sfmoma.org*

D.80 San Francisco Museum of Modern Art (d). "Matters in Media Art." *Sfmoma.org*.

D.81 San Francisco Museum of Modern Art (e). "Research and Projects." *Sfmoma.org*.

D.82 Society of American Archivists (SAA). "Appraisal." *Glossary of Archival and Records Terminology*. 2015.

D.83 Society of American Archivists (SAA) [b]. "Encapsulation." *Glossary of Archival and Records Terminology* from *Archivists.org*. 2015.

D.84 "Standards at the Library of Congress." *The Library of Congress*. Last Updated July 13, 2015.

D.85 Staudeman, Sarah and Paul Messier. "Video Format Identification Guide." *Video Preservation Website (VPW)* of Stanford University. 2007.

D.86 Tate Modern. "Matters in Media Art." *Tate.org*. Last updated December 2015.

D.87 Tate Modern (b). "Post-Acquisitions." *Matters in Media Art from Tate.org*. Last updated December 2015.

D.88 Tate Modern (c). "Lending Time-Based Media Art." *Matters in Media Art from Tate.org*. Last updated December 2015.

D.89 Tate Modern (d). "External Resources." *Matters in Media Art* from *Tate.org*.

D.90 The Metropolitan Museum of Modern Art (MET). "Main Building." *History of the Museum* from *Metmuseum.org*. 2015.

D.91 The Metropolitan Museum of Modern Art (b). "Museum Mission Statement." *Metmuseum.org*. 2015.

D.92 The Metropolitan Museum of Art (c) "Thomas J. Watson Library." *Museum Departments, Office of the Director* from *Metmuseum.org*. 2015.

D.93 The Metropolitan Museum of Art (d) "Museum Archives." *Museum Departments, Office of the Director* from *Metmuseum.org*. 2015.

D.94 The Metropolitan Museum of Art (e) "The Collections Online." *Metmuseum.org*. 2015.

D.95 The Metropolitan Museum of Art (f) "Collection." *Metmuseum.org*. 2015. Accessed September 27, 2015.

D.96 Thomas J. Watson Library. "Digital Collections from The Metropolitan Museum of Art Libraries." *Digital Collections, Thomas J. Watson Library, The Metropolitan Museum of Art*. 2015.

D.97 The UK National Archives."Download DROID: file format identification tool." *The National Archives*. 2015.

D.98 Tristram, Claire. "Data Extinction." *MIT Technology Review*, October 1, 2002.

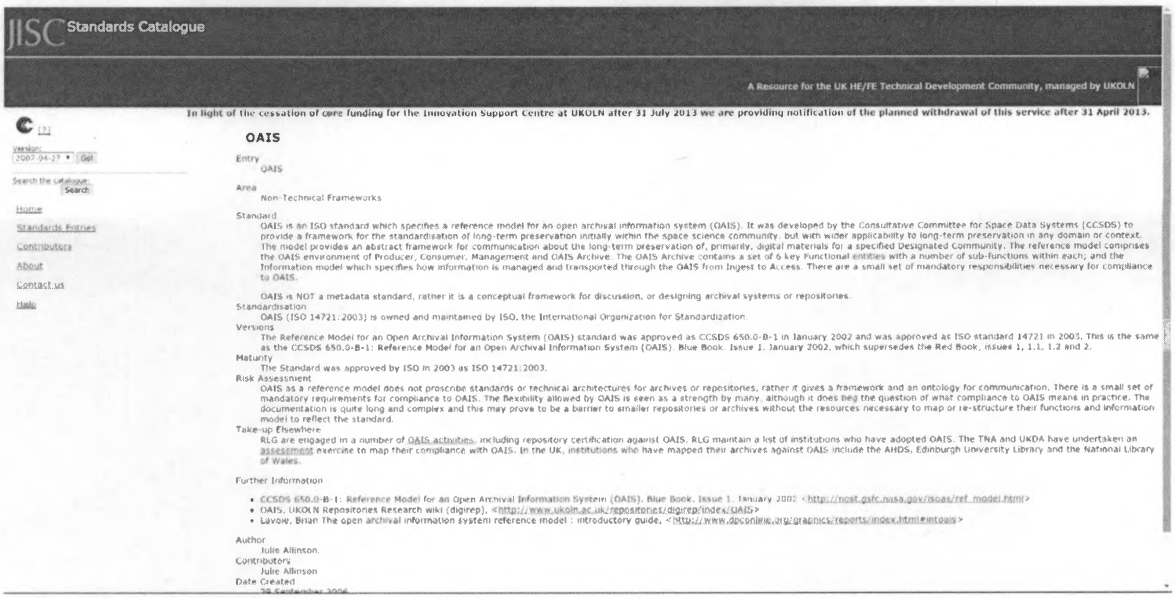
D.99 UC Berkeley School of Information Management and Systems. "Executive Summary." *How Much Information? 2003*.

D.100 Wall, Helen D. "Picturing Met Museum through Visitor's Eyes." *Digital Underground* blog, April 16, 2015.

D.101 Wheatley, P., Andy Jackson, and Andy Tester (contributors). "Main Page: Community Owned digital Preservation Tool Registry." *COPTR*. Last updated November 28, 2014.

List of Screenshots/Excerpts of Cited Websites:

D.1 Allinson, Julie. "OAIS." *JISC Standards Catalogue*. Updated September 23, 2006.



JISC Standards Catalogue

A Resource for the UK HE/FE Technical Development Community, managed by UKOLN

In light of the cessation of core funding for the Innovation Support Centre at UKOLN after 31 July 2013 we are providing notification of the planned withdrawal of this service after 31 April 2013.

Version: 2007-04-27 Go!

Search the catalogue: Search

Home
Standards Entries
Contributors
About
Contact us
Help

OAIS

Entry
OAIS

Area
Non-Technical Frameworks

Standard
OAIS is an ISO standard which specifies a reference model for an open archival information system (OAIS). It was developed by the Consultative Committee for Space Data Systems (CCSDS) to provide a framework for the standardisation of long-term preservation initially within the space science community, but with wider applicability to long-term preservation in any domain or context. The model provides an abstract framework for communication about the long-term preservation of, primarily, digital materials for a specified Designated Community. The reference model comprises the OAIS environment of Producer, Consumer, Management and OAIS Archive. The OAIS Archive contains a set of 6 key Functional entities with a number of sub-functions within each; and the Information model which specifies how information is managed and transported through the OAIS from Ingest to Access. There are a small set of mandatory responsibilities necessary for compliance to OAIS.

OAIS is NOT a metadata standard, rather it is a conceptual framework for discussion, or designing archival systems or repositories.

Standardisation
OAIS (ISO 14721:2003) is owned and maintained by ISO, the International Organization for Standardization

Versions
The Reference Model for an Open Archival Information System (OAIS) standard was approved as CCSDS 650.0-B-1 in January 2002 and was approved as ISO standard 14721 in 2003. This is the same as the CCSDS 650.0-B-1: Reference Model for an Open Archival Information System (OAIS): Blue Book. Issue 1. January 2002, which supersedes the Red Book, issues 1, 1.1, 1.2 and 2.

Maturity
The Standard was approved by ISO in 2003 as ISO 14721:2003.

Risk Assessment
OAIS as a reference model does not proscribe standards or technical architectures for archives or repositories, rather it gives a framework and an ontology for communication. There is a small set of mandatory requirements for compliance to OAIS. The flexibility allowed by OAIS is seen as a strength by many, although it does beg the question of what compliance to OAIS means in practice. The documentation is quite long and complex and this may prove to be a barrier to smaller repositories or archives without the resources necessary to map or re-structure their functions and information model to reflect the standard.

Take-up Elsewhere
RLG are engaged in a number of OAIS activities, including repository certification against OAIS. RLG maintain a list of institutions who have adopted OAIS. The TNA and UKDA have undertaken an assessment exercise to map their compliance with OAIS. In the UK, institutions who have mapped their archives against OAIS include the AHDS, Edinburgh University Library and the National Library of Wales.

Further Information

- CCSDS 650.0-B-1: Reference Model for an Open Archival Information System (OAIS). Blue Book. Issue 1. January 2002 <http://nost.gsfc.nasa.gov/isoas/ref_model.html>
- OAIS. UKOLN Repositories Research wiki (digirep). <http://www.ukoln.ac.uk/repositories/digirep/index/OAIS>
- Lavoie, Brian The open archival information system reference model : introductory guide. <http://www.dpconline.org/graphics/reports/index.html#entoais>

Author
Julie Allinson.
Contributors
Julie Allinson
Date Created
28 September 2006

D.2 American Alliance of Museums (AAM). "Continuum of Excellence." *Assessment Programs*. Last updated 2015

D.3 American Alliance of Museums (AAM) (b). "Assessment Types." *Museum Assessment Program* (MAP). Last updated 2015.

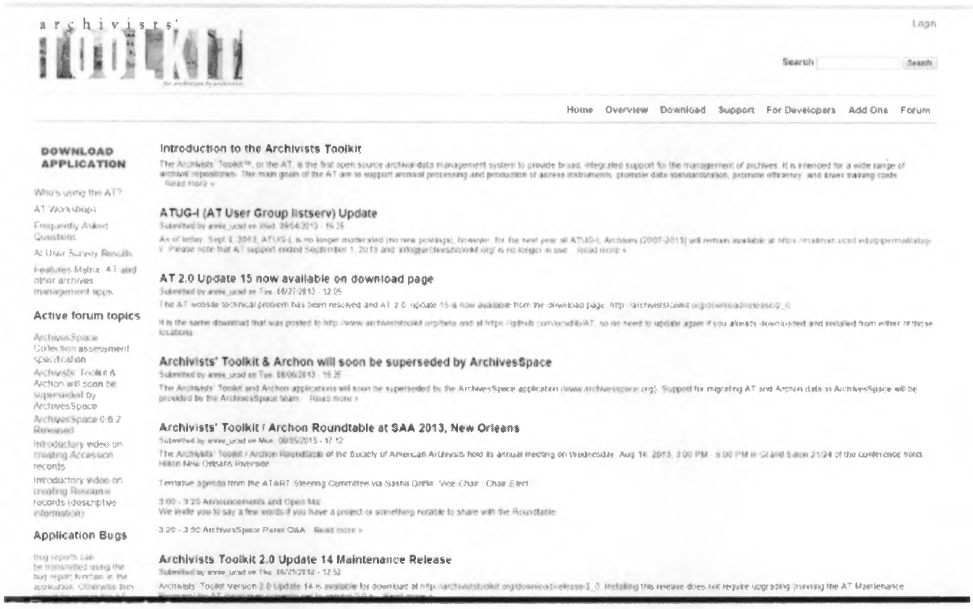
D.4 Anderson, Martha. "B is for Bit Preservation," *The Signal* blog, September 7, 2011.

D. 5 Archive-It. "About Us." *Archive-It blog*. 2014.

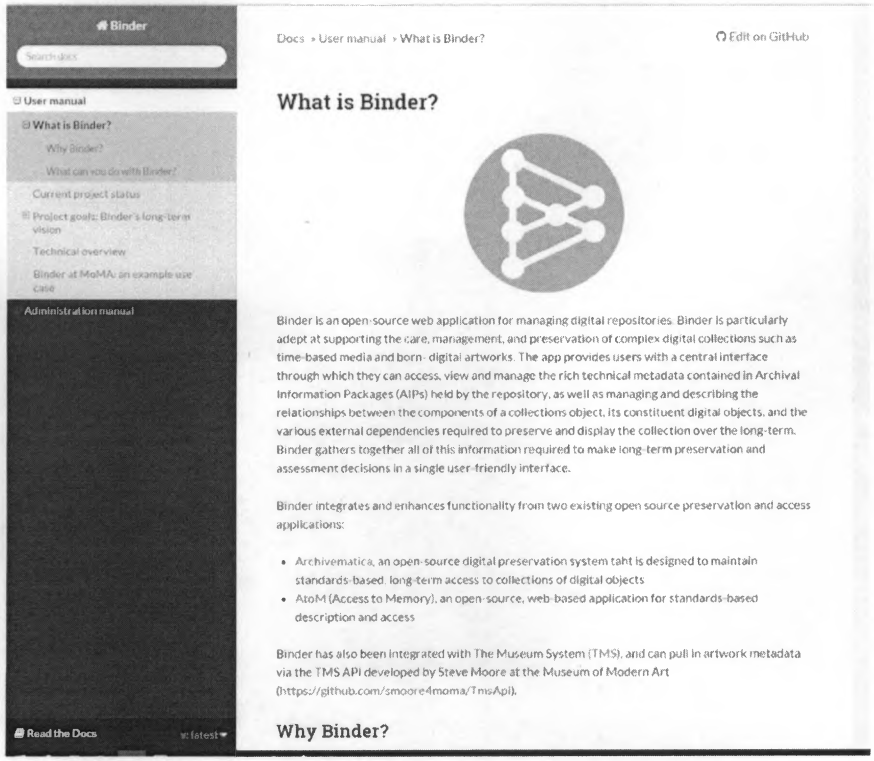
D.6 Archive-it (b). "New York Art Resources Consortium (NYARC)." *Nyarc.org*. 2014.

D.7 Archivematica. "Archivematica: preserving memory since 2009." *Artefactual Systems Inc*. 2015.

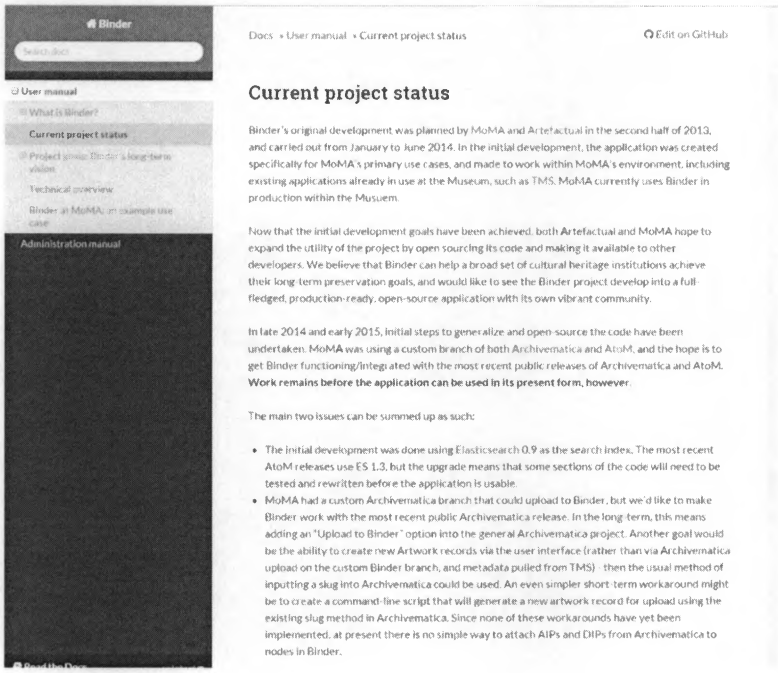
D.8 Archivists' Toolkit. "Archivists' Toolkit." 2009.



D.9 Artefactual Systems."User Manual -What is Binder?" Last updated May 22, 2015.



D.10 Artefactual Systems."User Manual -Current Project Status." Last updated May 22, 2015.

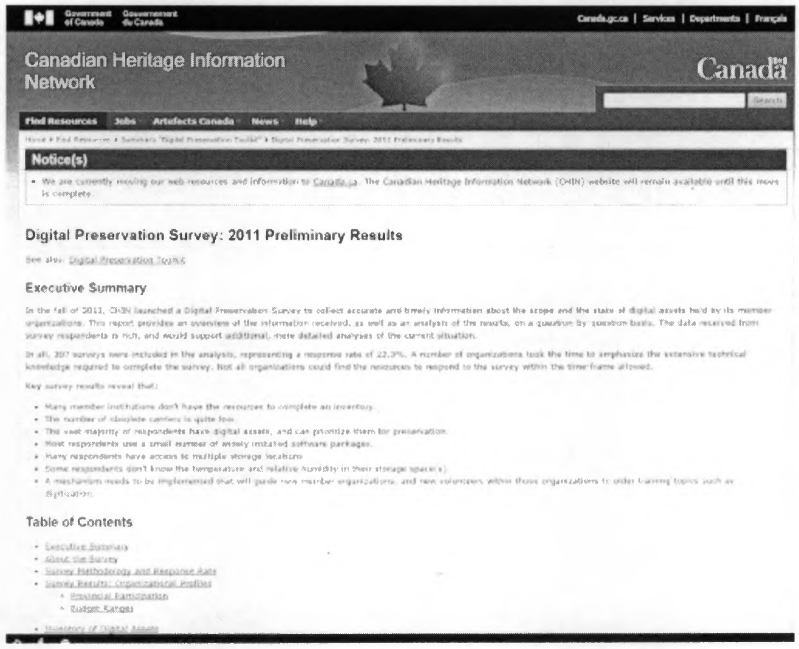


D.11 ArtStor. "The Metropolitan Museum of Art." *Artstor.org*. Accessed September 26, 2015.

D.12 Blue Ribbon Task Force on Sustainable Digital Preservation and Access. "Index." 2008.

D.13 Canadian Heritage Information Network (CHIN). "Digital Preservation Plan Framework for Museums." *Canadian Heritage Information Network*. Last updated June 20, 2013.

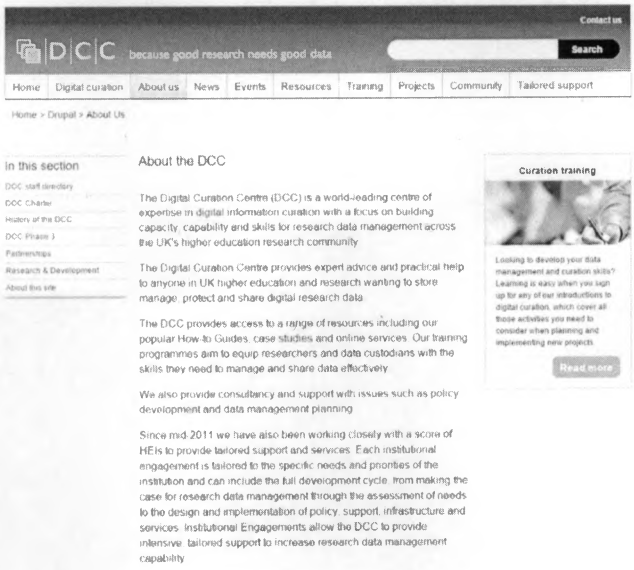
D.14 Canadian Heritage Information Network, 2011 Survey Results. 2015 (b).



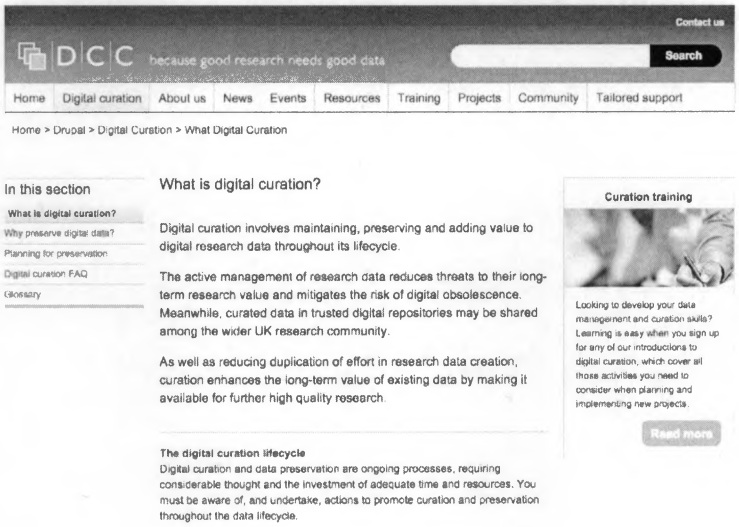
D.15 Canadian Heritage Information Network (CHIN) [c]. "CHIN Introduces a Digital Preservation Toolkit." Last modified June 12, 2013.

D.16 Cohen, Daniel J. and Roy Rosenzweig. "Preserving Digital History: The Fragility of Digital Materials." *Digital History: A Guide to Gathering, Preserving, and Presenting the Past on the Web*. 2005.

D.17 Digital Curation Centre (a). "About the DCC." 2004.

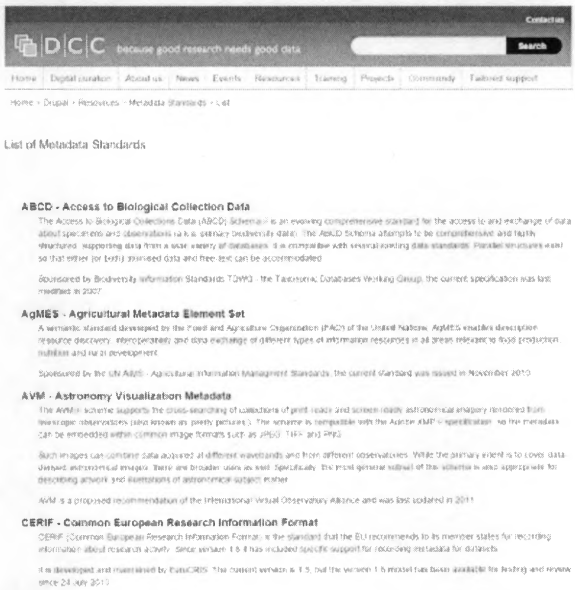


D.18 Digital Curation Centre (b). "What is Digital Curation?" 2015.

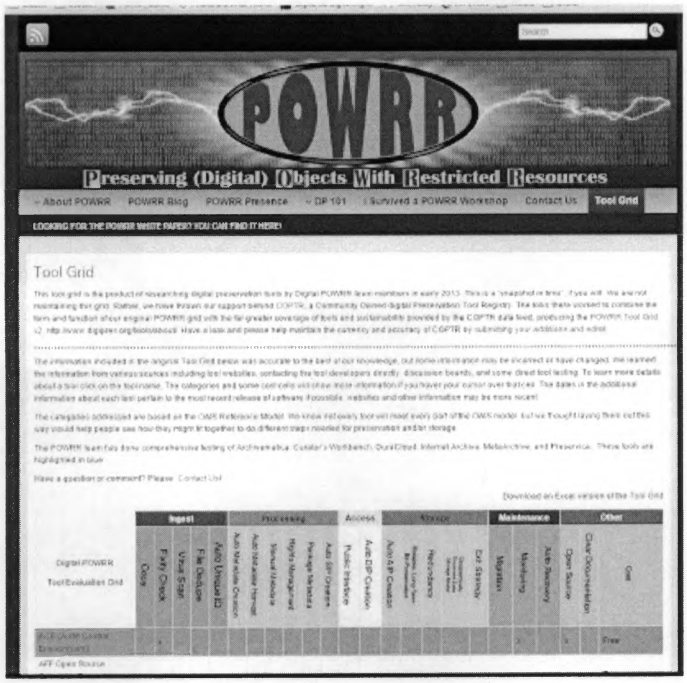


D.19 Digital Curation Centre (c). "JSTOR/Harvard Object Validation Environment (JHOVE)."

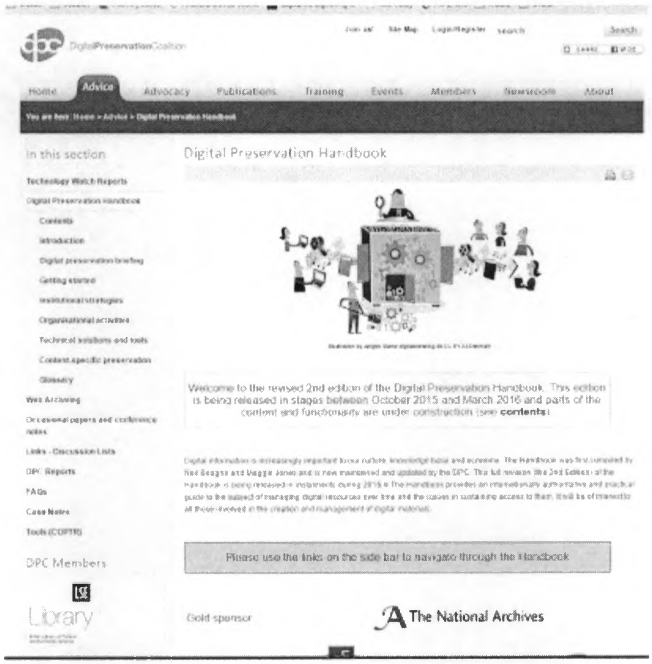
D.20 Digital Curation Centre (d). "List of Metadata Standards." *Digital Curation Centre* Resources. 2015.



D.21 Digital POWRR."Tool Grid." 2013.



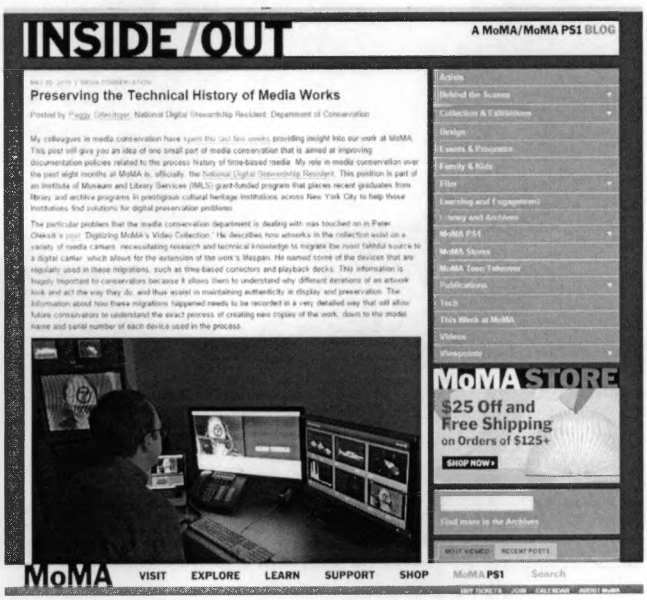
D.22 Digital Preservation Coalition (DPC). "Digital Preservation Handbook." 2015.



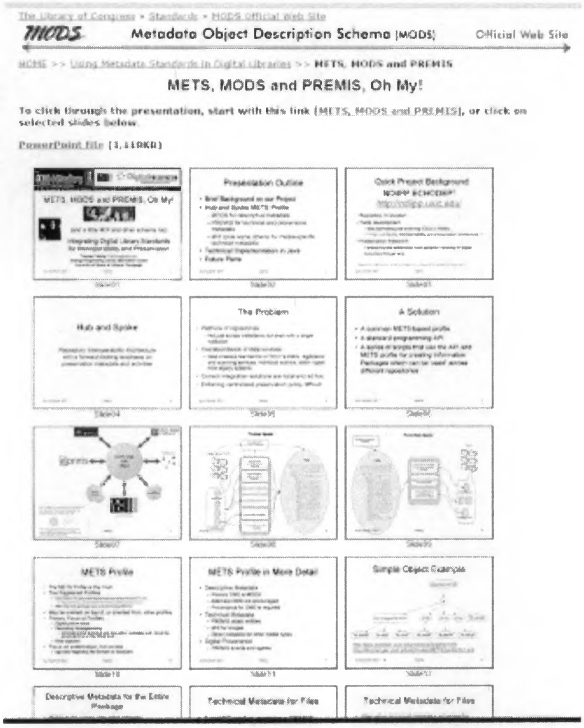
D.23 Fino-Radin, Ben (b). "MoMA's Digital Art Vault." *Inside/Out blog* for *Moma.org*, April 14, 2015.

D.24 Fleming, Robin and Dan Lipcan. "Digitizing the Library's' Collections: An Introduction," *Now at The Met*, January 5, 2012.

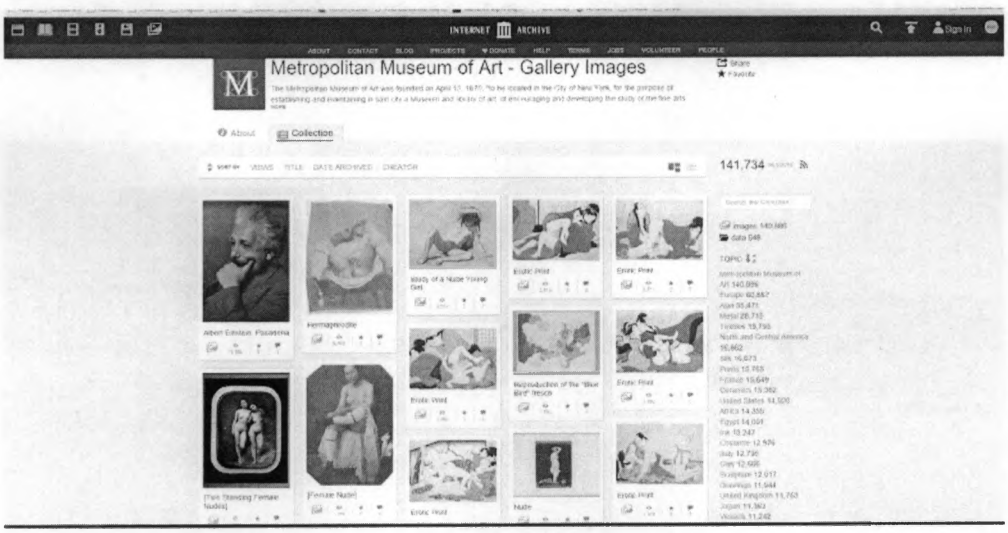
D.25 Griesinger, Peggy. "Preserving the Technical History of Media Works." *Inside/Out* blog for *Moma.org.* May 20, 2015.



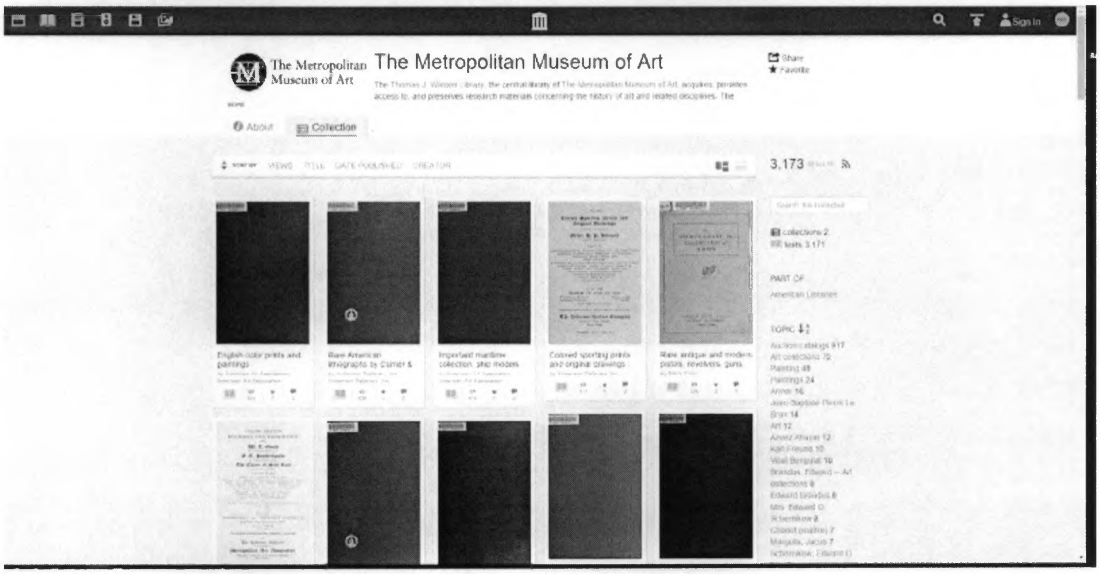
D.26 Habing, Thomas."METS, Mods, and PREMIS, Oh My!" *Library of Congress, MODS Official Website*. Presentation for American Library Association 2007 Meeting.



D.27 Internet Archive. "Metropolitan Museum of Art - Gallery Images." *Archive.org*. 2014.



D.28 Internet Archive. "The Metropolitan Museum of Art, Library." *Archive.org*. 2015.

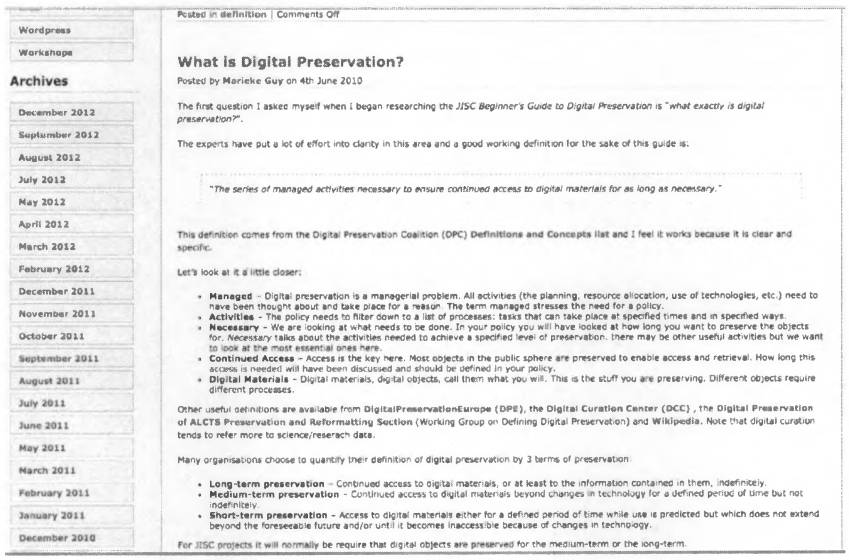


D.29 Internet Archive (b). "About the Internet Archive." *Archive.org*. 2015.

D.30 InterPARES 2 Project. "Terminology Database." 2015.

D.31 "JHOVE2, The Next-Generation Architecture for Format-Aware Characterization." *GitHub*. 2015.

D.32 JISC. "Definition of Digital Preservation," in *JISC Beginner's Guide to Digital Preservation.* Last updated 2012.



Wordpress
Workshops

Archives

December 2012
September 2012
August 2012
July 2012
May 2012
April 2012
March 2012
February 2012
December 2011
November 2011
October 2011
September 2011
August 2011
July 2011
June 2011
May 2011
March 2011
February 2011
January 2011
December 2010

Posted in definition | Comments Off

What is Digital Preservation?

Posted by Marieke Guy on 4th June 2010

The first question I asked myself when I began researching the *JISC Beginner's Guide to Digital Preservation* is "*what exactly is digital preservation?*".

The experts have put a lot of effort into clarity in this area and a good working definition for the sake of this guide is:

> *"The series of managed activities necessary to ensure continued access to digital materials for as long as necessary."*

This definition comes from the Digital Preservation Coalition (DPC) Definitions and Concepts list and I feel it works because it is clear and specific.

Let's look at it a little closer:

- **Managed** – Digital preservation is a managerial problem. All activities (the planning, resource allocation, use of technologies, etc.) need to have been thought about and take place for a reason. The term managed stresses the need for a policy.
- **Activities** – The policy needs to filter down to a list of processes: tasks that can take place at specified times and in specified ways.
- **Necessary** – We are looking at what needs to be done. In your policy you will have looked at how long you want to preserve the objects for. *Necessary* talks about the activities needed to achieve a specified level of preservation. There may be other useful activities but we want to look at the most essential ones here.
- **Continued Access** – Access is the key here. Most objects in the public sphere are preserved to enable access and retrieval. How long this access is needed will have been discussed and should be defined in your policy.
- **Digital Materials** – Digital materials, digital objects, call them what you will. This is the stuff you are preserving. Different objects require different processes.

Other useful definitions are available from **DigitalPreservationEurope (DPE)**, the **Digital Curation Center (DCC)** , the **Digital Preservation of ALCTS Preservation and Reformatting Section** (Working Group on Defining Digital Preservation) and **Wikipedia**. Note that digital curation tends to refer more to science/reserach data.

Many organisations choose to quantify their definition of digital preservation by 3 terms of preservation:

- **Long-term preservation** – Continued access to digital materials, or at least to the information contained in them, indefinitely.
- **Medium-term preservation** – Continued access to digital materials beyond changes in technology for a defined period of time but not indefinitely.
- **Short-term preservation** – Access to digital materials either for a defined period of time while use is predicted but which does not extend beyond the foreseeable future and/or until it becomes inaccessible because of changes in technology.

For JISC projects it will normally be require that digital objects are preserved for the medium-term or the long-term.

D.33 LeFurgy, Bill. "Facing Off with Digital Preservation Policy." *The Signal Blog* from The Library of Congress, July 6, 2011.

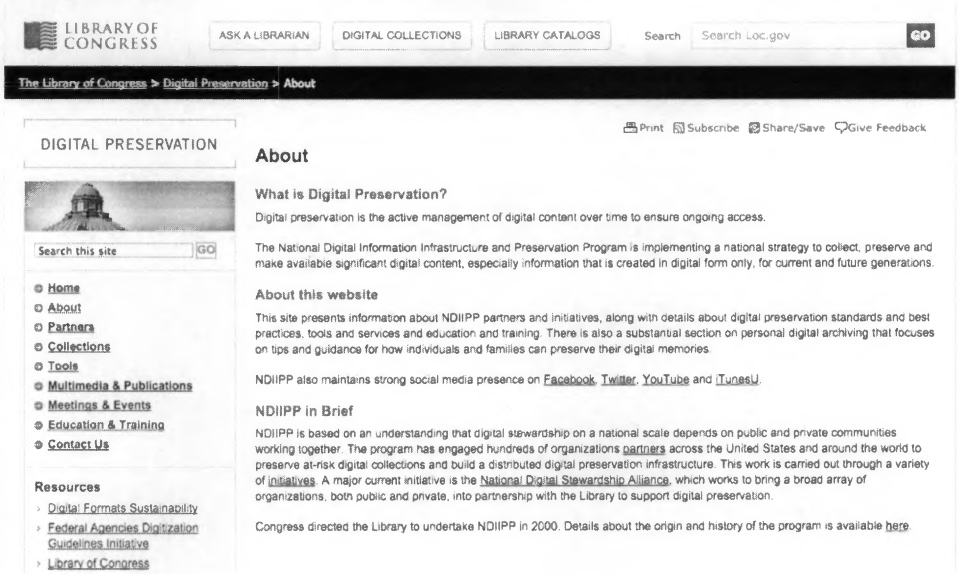
LIBRARY OF CONGRESS

Search

THE SIGNAL
DIGITAL PRESERVATION

Facing Off with Digital Preservation Policy

D.34 Library of Congress. "About." Digital Preservation (Library of Congress). 2015.



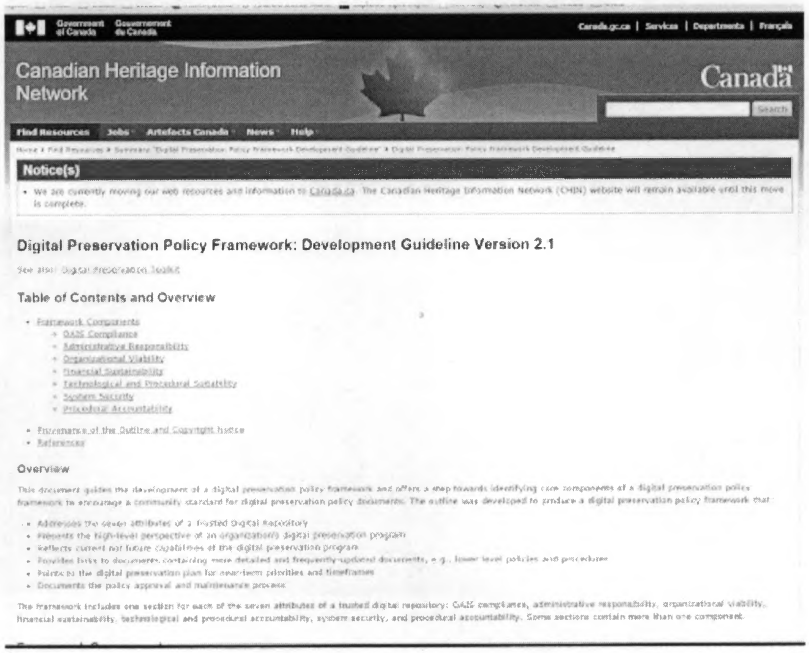
D.35 Library of Congress (b). "Recommended File Formats." *Library of Congress Preservation Resources*. Last updated 2015.

D.36 "Life Cycle Information for E-Literature Project (LIFE)." 2015.

D.37 Manus, Susan. "At the Museum: An Interview with Marla Misunas (and Friends) of SFMOMA, Pt.2," *The Signal* blog. June 19, 2014.

D.38 Manus, Susan. "At the Museum: An Interview with Marla Misunas of SFMOMA, Pt.1," *The Signal* blog. April 2, 2014.

D.39 McGovern, Nancy."Digital Preservation Policy Framework: Development Guideline Version 2.1." *Digital Preservation Toolkit* for the *Canadian Heritage Information Network*. Last modified April 25, 2013.



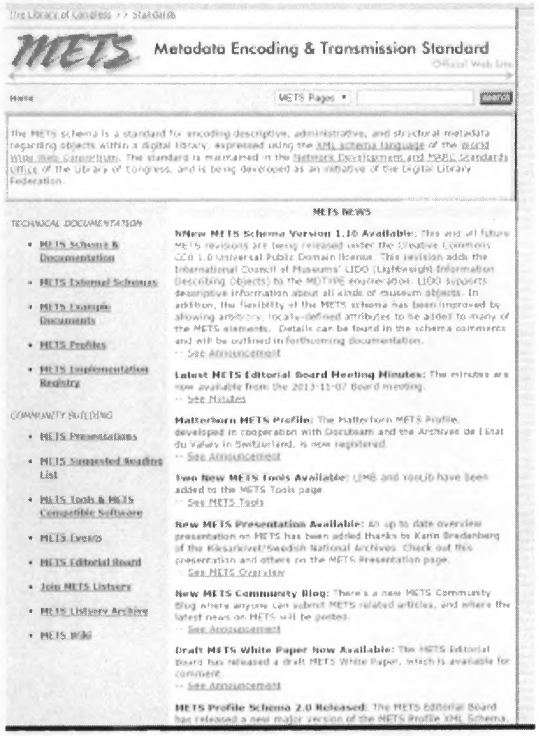
D.40 MetaArchive Cooperative. "The Cooperative: A brief history of the first private digital preservation network." 2014.

D.41 MetaArchive Cooperative (b). "Methodology: don't put all your eggs in one basket." 2014.

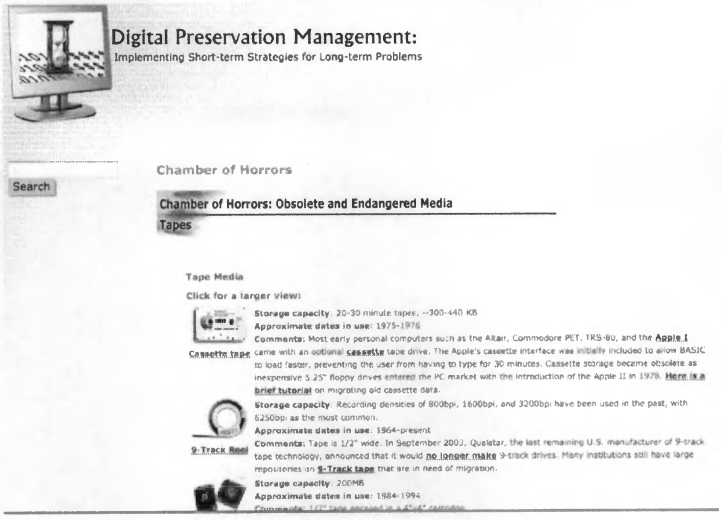
D.42 MetaArchive Cooperative (c). "Costs." 2014.

D.43 MetaArchive Cooperative (d). "Our Members: Membership Map." 2014.

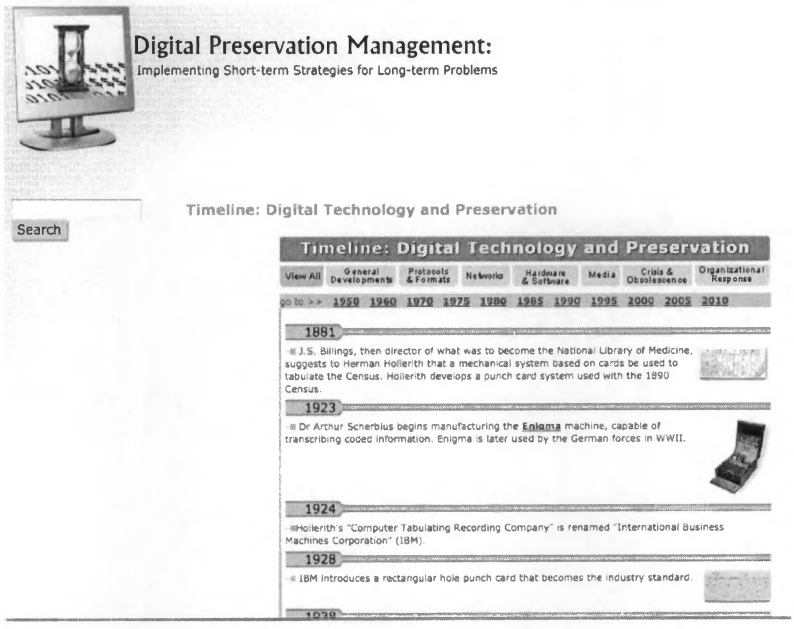
D.44 "Metadata Encoding & Transmission Standard (METS)." *Library of Congress*. Last updated November 23, 2015.



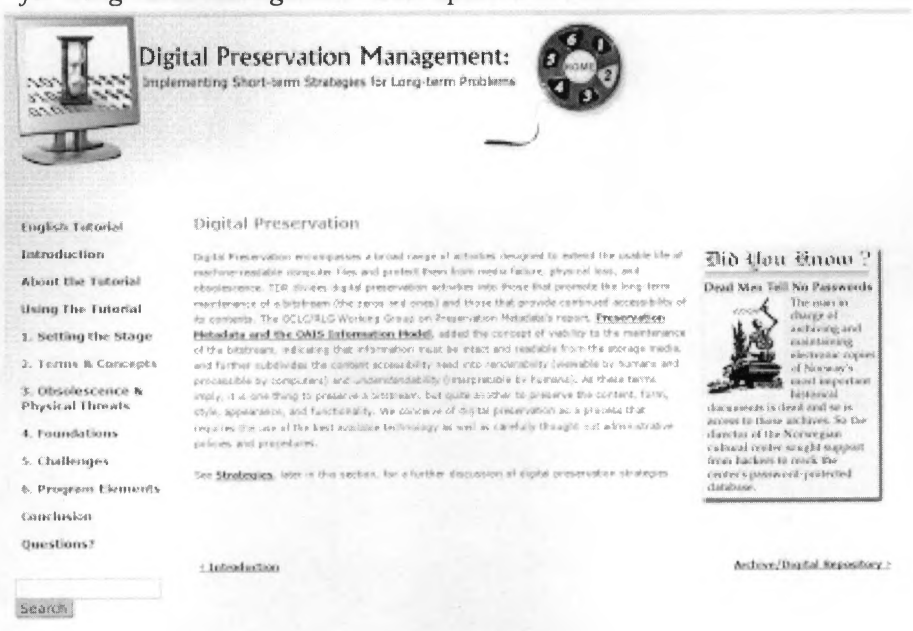
D.45 MIT Libraries. Digital Preservation Workshop. "Chamber of Horrors: Obsolete and Endangered Media." *Digital Preservation Management: Implementing Short term Strategies for Long Term Management*. Last updated 2012.



D.46 MIT Libraries (b). Digital Preservation Workshop. "Timeline." *Digital Preservation Management: Implementing Short term Strategies for Long Term Management.* Last updated 2012.

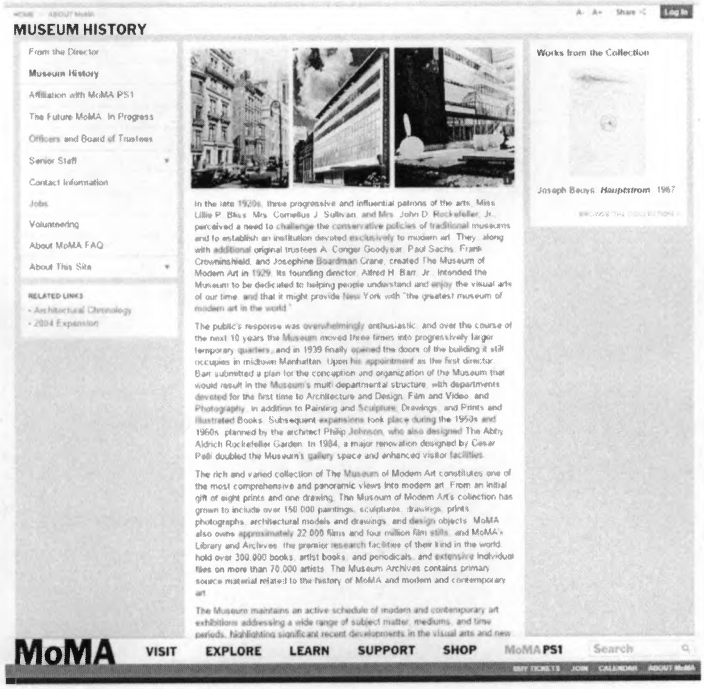


D.47 MIT Libraries (c). Digital Preservation Workshop. "Terms and Concepts: Digital Preservation." *Digital Preservation Management: Implementing Short term Strategies for Long Term Management.* Last updated 2012.



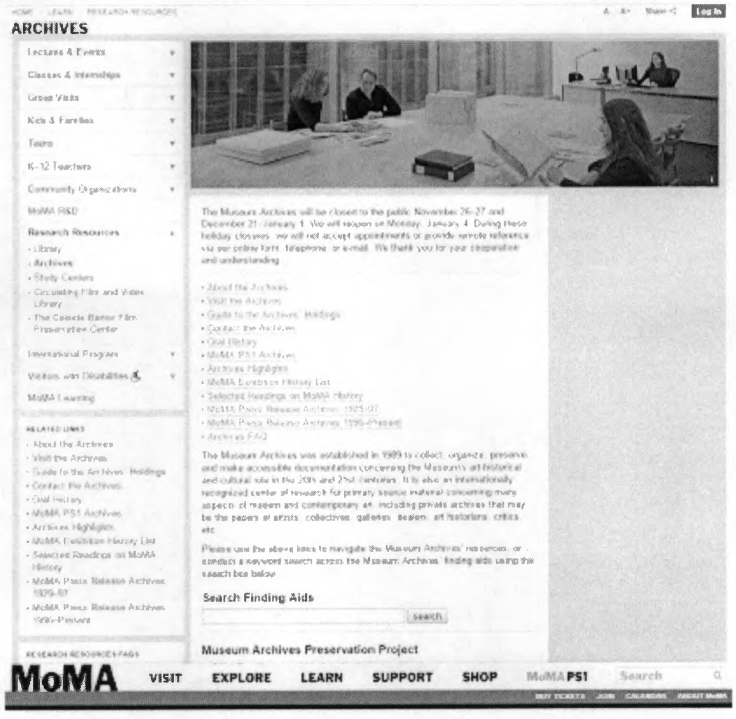
D.48 Murray, Kate. "Preserving Digital and Software-Based Artworks: Recap of a NDSA Discussion," *The Signal* Blog, June 13, 2014.

D.49 Museum of Modern Art. "Museum History." *Moma.org*. 2015.



D.50 Museum of Modern Art. "About MoMA - Curatorial Departments." *Press.moma.org.* 2011.

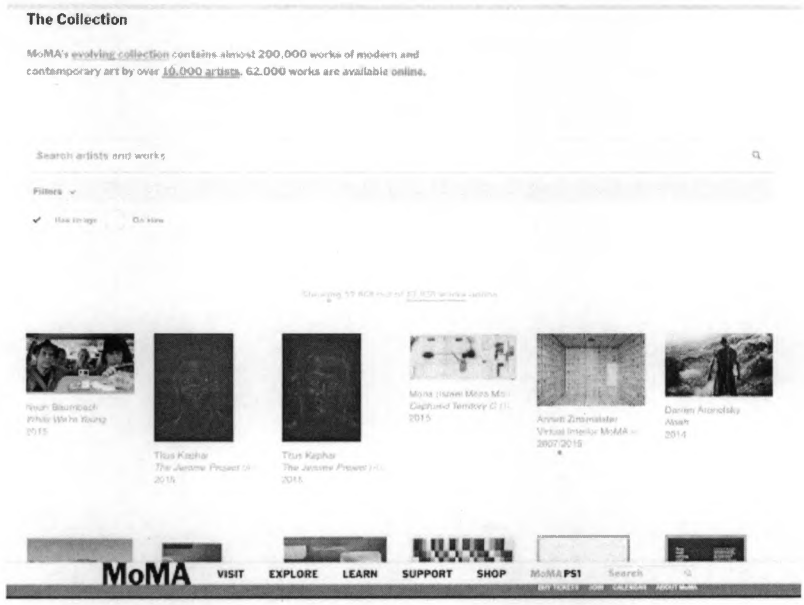
D.51 Museum of Modern Art (b). "Archives." *Moma.org.* 2015.



D.52 Museum of Modern Art (c). "Film Preservation Center." Moma.org. 2015

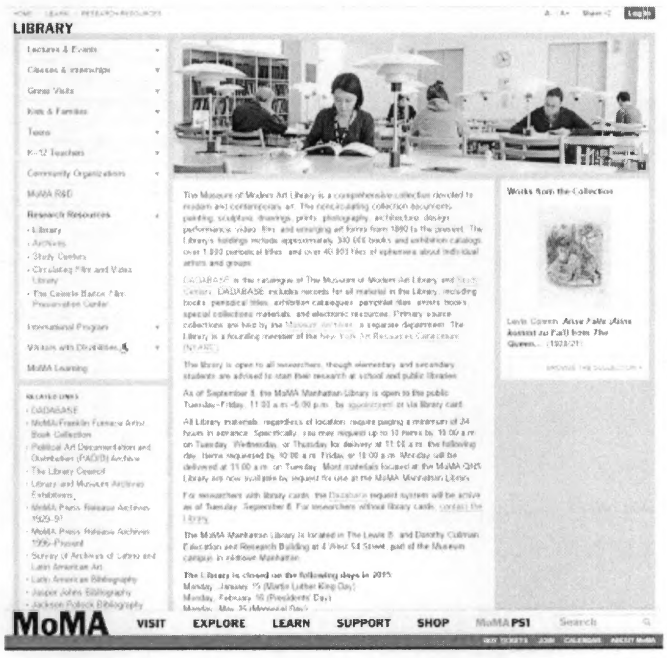
D.53 Museum of Modern Art (d). "About MoMA." *Moma.org*. 2015.

D.54 Museum of Modern Art (e). "The Collection." *Moma.org*. 2015.

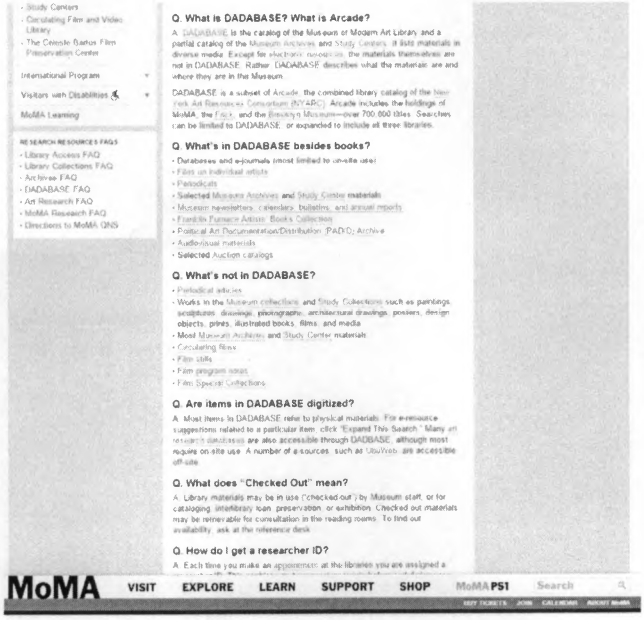


D.55 Museum of Modern Art (f). "About The Collection." *Moma.org*. 2015.

D.56 Museum of Modern Art (g). "Library." *Moma.org*. 2015.



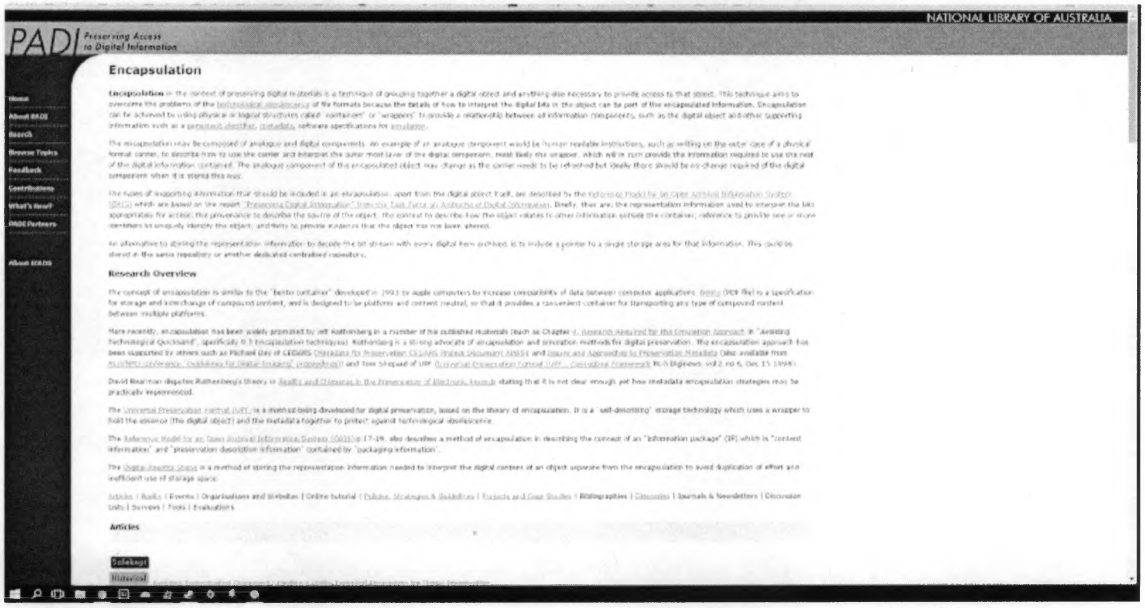
D.57 Museum of Modern Art (h). "Dadabase FAQs." Moma.org. 2015.



D.58 National Digital Stewardship Alliance (NDSA). "Digital Preservation in a Box." Last updated 2015.

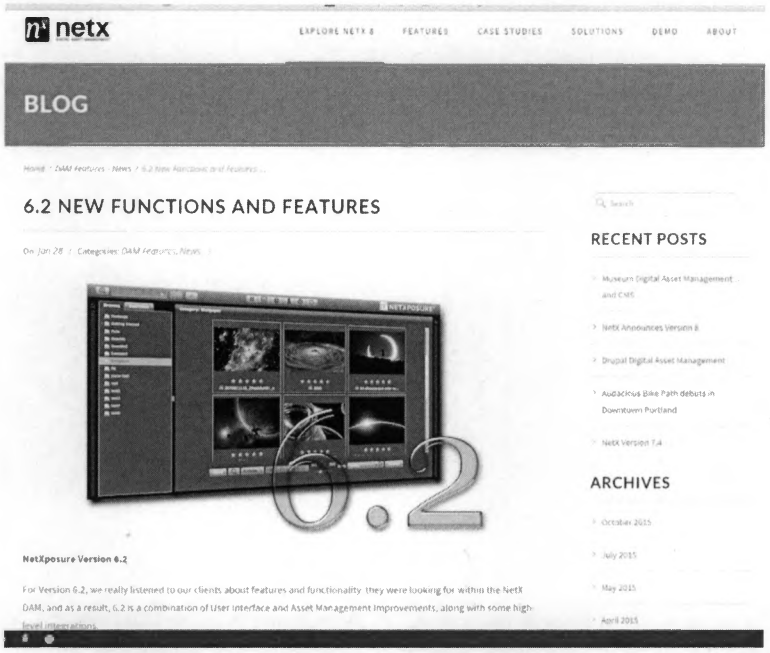
D.59 National Initiative for a Networked Cultural Heritage (b). "About NINCH." 2003.

D.60 National Library of Australia. "Encapsulation." *Preserving Access to Digital Information.* 2001.



D.61 NetX. "Museum Digital Asset Management at MOMA NY."

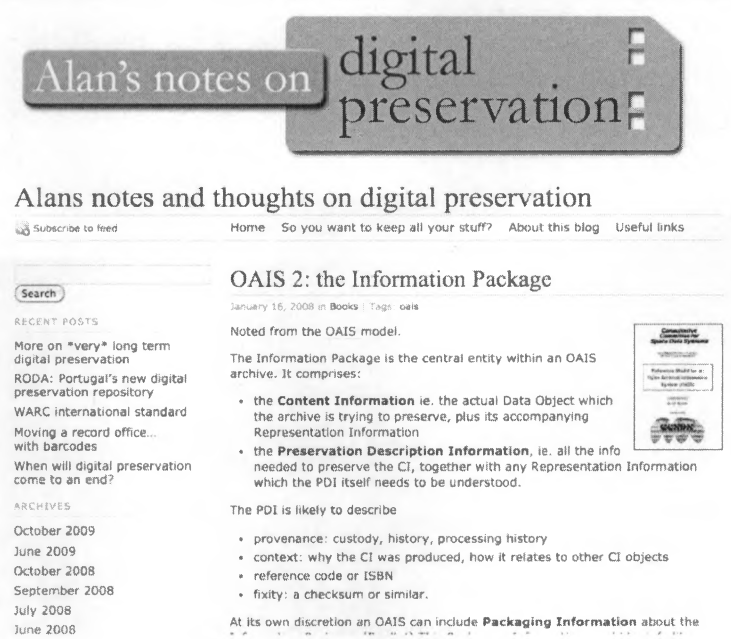
D.62 NetX. "New Functions and Features." *Net Exposure Blog*, January 28, 2013.



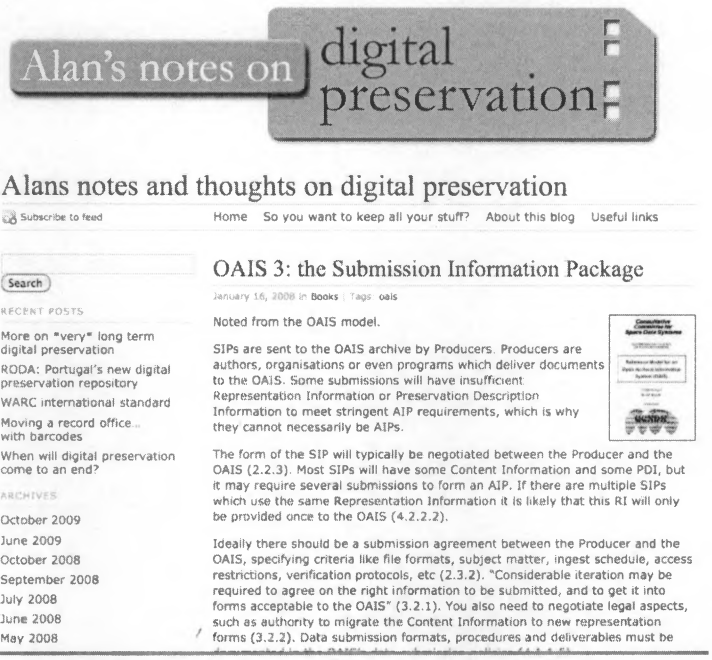
D.63 "New Art Trust Names John R. Lane as President and CEO." ArtDaily.com. October 8, 2008.

D.64 New York Art Resources Consortium. "Web Archiving." *Nyarc.org*. 2015.

D.65 "OAIS 2: the Information Package." *Alan's Thoughts on Digital Preservation Blog.* January 16, 2008.



D.66 "OAIS 3: the Submission Information Package." *Alan's Thoughts on Digital Preservation Blog*. January 16, 2008 (b).


Alan's notes on digital preservation

Alans notes and thoughts on digital preservation

Subscribe to feed | Home | So you want to keep all your stuff? | About this blog | Useful links

Search

RECENT POSTS

More on "very" long term digital preservation
RODA: Portugal's new digital preservation repository
WARC international standard
Moving a record office... with barcodes
When will digital preservation come to an end?

ARCHIVES

October 2009
June 2009
October 2008
September 2008
July 2008
June 2008
May 2008

OAIS 3: the Submission Information Package

January 16, 2008 in Books | Tags: oais

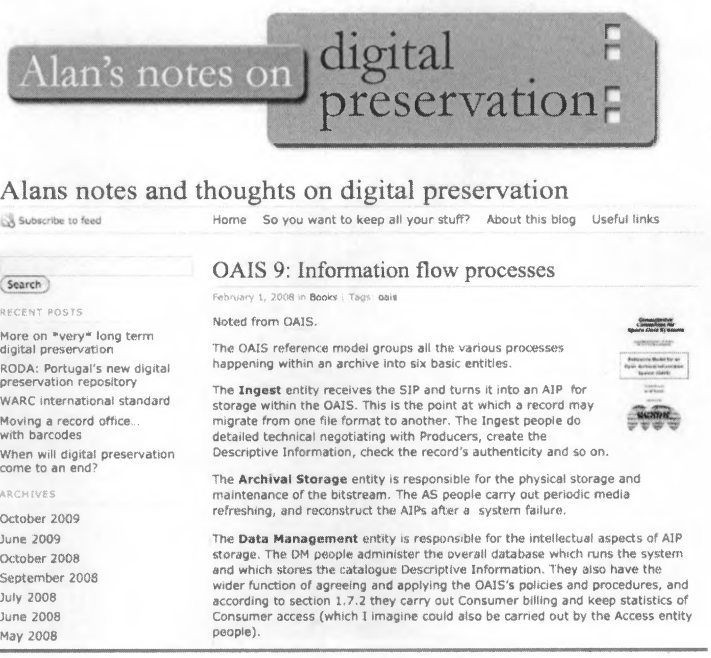
Noted from the OAIS model.

SIPs are sent to the OAIS archive by Producers. Producers are authors, organisations or even programs which deliver documents to the OAIS. Some submissions will have insufficient Representation Information or Preservation Description Information to meet stringent AIP requirements, which is why they cannot necessarily be AIPs.

The form of the SIP will typically be negotiated between the Producer and the OAIS (2.2.3). Most SIPs will have some Content Information and some PDI, but it may require several submissions to form an AIP. If there are multiple SIPs which use the same Representation Information it is likely that this RI will only be provided once to the OAIS (4.2.2.2).

Ideally there should be a submission agreement between the Producer and the OAIS, specifying criteria like file formats, subject matter, ingest schedule, access restrictions, verification protocols, etc (2.3.2). "Considerable iteration may be required to agree on the right information to be submitted, and to get it into forms acceptable to the OAIS" (3.2.1). You also need to negotiate legal aspects, such as authority to migrate the Content Information to new representation forms (3.2.2). Data submission formats, procedures and deliverables must be

D.67 "OAIS 9: Information Flow Processes." *Alan's Thoughts on Digital Preservation Blog*. February 1, 2008(c).


Alan's notes on digital preservation

Alans notes and thoughts on digital preservation

Subscribe to feed | Home | So you want to keep all your stuff? | About this blog | Useful links

Search

RECENT POSTS

More on "very" long term digital preservation
RODA: Portugal's new digital preservation repository
WARC international standard
Moving a record office... with barcodes
When will digital preservation come to an end?

ARCHIVES

October 2009
June 2009
October 2008
September 2008
July 2008
June 2008
May 2008

OAIS 9: Information flow processes

February 1, 2008 in Books | Tags: oais

Noted from OAIS.

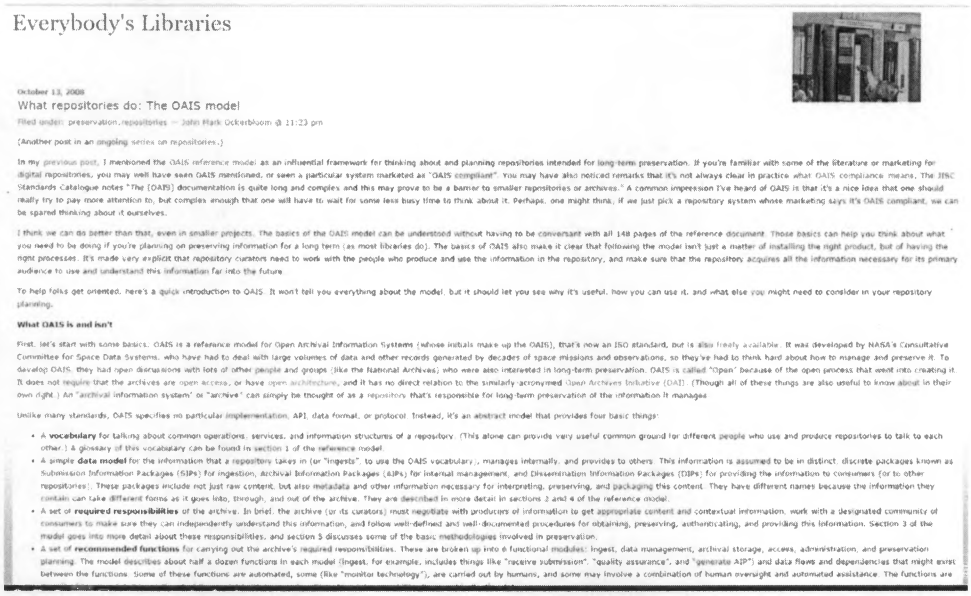
The OAIS reference model groups all the various processes happening within an archive into six basic entities.

The **Ingest** entity receives the SIP and turns it into an AIP for storage within the OAIS. This is the point at which a record may migrate from one file format to another. The Ingest people do detailed technical negotiating with Producers, create the Descriptive Information, check the record's authenticity and so on.

The **Archival Storage** entity is responsible for the physical storage and maintenance of the bitstream. The AS people carry out periodic media refreshing, and reconstruct the AIPs after a system failure.

The **Data Management** entity is responsible for the intellectual aspects of AIP storage. The DM people administer the overall database which runs the system and which stores the catalogue Descriptive Information. They also have the wider function of agreeing and applying the OAIS's policies and procedures, and according to section 1.7.2 they carry out Consumer billing and keep statistics of Consumer access (which I imagine could also be carried out by the Access entity people).

D.68 Ockerbloom, John Mark. "What Repositories Do: The OAIS Model." *Everybody's Libraries Blog*, October 13, 2008



Everybody's Libraries

October 13, 2008

What repositories do: The OAIS model

Filed under: preservation, repositories — John Mark Ockerbloom @ 11:23 pm

(Another post in an ongoing series on repositories.)

In my previous post, I mentioned the OAIS reference model as an influential framework for thinking about and planning repositories intended for long-term preservation. If you're familiar with some of the literature or marketing for digital repositories, you may well have seen OAIS mentioned, or seen a particular system marketed as "OAIS compliant". You may have also noticed remarks that it's not always clear in practice what OAIS compliance means. The JISC Standards Catalogue notes "The [OAIS] documentation is quite long and complex and this may prove to be a barrier to smaller repositories or archives." A common impression I've heard of OAIS is that it's a nice idea that one should really try to pay more attention to, but complex enough that one will have to wait for some less busy time to think about it. Perhaps, one might think, if we just pick a repository system whose marketing says it's OAIS compliant, we can be spared thinking about it ourselves.

I think we can do better than that, even in smaller projects. The basics of the OAIS model can be understood without having to be conversant with all 148 pages of the reference document. Those basics can help you think about what you need to be doing if you're planning on preserving information for a long term (as most libraries do). The basics of OAIS also make it clear that following the model isn't just a matter of installing the right product, but of having the right processes. It's made very explicit that repository curators need to work with the people who produce and use the information in the repository, and make sure that the repository acquires all the information necessary for its primary audience to use and understand this information far into the future.

To help folks get oriented, here's a quick introduction to OAIS. It won't tell you everything about the model, but it should let you see why it's useful, how you can use it, and what else you might need to consider in your repository planning.

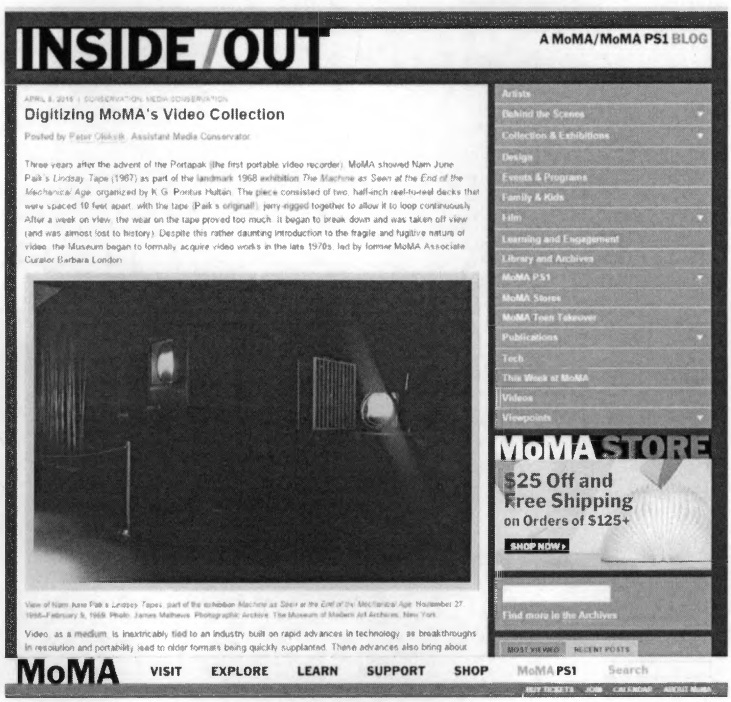
What OAIS is and isn't

First, let's start with some basics. OAIS is a reference model for Open Archival Information Systems (whose initials make up the OAIS), that's now an ISO standard, but is also freely available. It was developed by NASA's Consultative Committee for Space Data Systems, who have had to deal with large volumes of data and other records generated by decades of space missions and observations, so they've had to think hard about how to manage and preserve it. To develop OAIS, they had open discussions with lots of other people and groups (like the National Archives) who were also interested in long-term preservation. OAIS is called "Open" because of the open process that went into creating it. It does not require that the archives are open access, or have open architecture, and it has no direct relation to the similarly acronymed Open Archives Initiative (OAI). (Though all of these things are also useful to know about in their own right.) An "archival information system" or "archive" can simply be thought of as a repository that's responsible for long-term preservation of the information it manages.

Unlike many standards, OAIS specifies no particular implementation, API, data format, or protocol. Instead, it's an abstract model that provides four basic things:

- A **vocabulary** for talking about common operations, services, and information structures of a repository. (This alone can provide very useful common ground for different people who use and produce repositories to talk to each other.) A glossary of this vocabulary can be found in section 1 of the reference model.
- A simple **data model** for the information that a repository takes in (or "ingests", to use the OAIS vocabulary), manages internally, and provides to others. This information is assumed to be in distinct, discrete packages known as Submission Information Packages (SIPs) for ingestion, Archival Information Packages (AIPs) for internal management, and Dissemination Information Packages (DIPs) for providing the information to consumers (or to other repositories). These packages include not just raw content, but also metadata and other information necessary for interpreting, preserving, and packaging this content. They have different names because the information they contain can take different forms as it goes into, through, and out of the archive. They are described in more detail in sections 2 and 4 of the reference model.
- A set of **required responsibilities** of the archive. In brief, the archive (or its curators) must negotiate with producers of information to get appropriate content and contextual information, work with a designated community of consumers to make sure they can independently understand this information, and follow well-defined and well-documented procedures for obtaining, preserving, authenticating, and providing this information. Section 3 of the model goes into more detail about these responsibilities, and section 5 discusses some of the basic methodologies involved in preservation.
- A set of **recommended functions** for carrying out the archive's required responsibilities. These are broken up into 6 functional modules: Ingest, data management, archival storage, access, administration, and preservation planning. The model describes about half a dozen functions in each model (Ingest, for example, includes things like "receive submission", "quality assurance", and "generate AIP") and data flows and dependencies that might exist between the functions. Some of these functions are automated, some (like "monitor technology") are carried out by humans, and some may involve a combination of human oversight and automated assistance. The functions are

D.69 Oleksik, Peter. "Digitizing MoMA's Video Collection." *Inside/Out blog* for *Moma.org*. April 8, 2015.



INSIDE/OUT

A MoMA/MoMA PS1 BLOG

APRIL 8, 2015 | CONSERVATION, MEDIA CONSERVATION

Digitizing MoMA's Video Collection

Posted by Peter Oleksik, Assistant Media Conservator

Three years after the advent of the Portapak (the first portable video recorder), MoMA showed Nam June Paik's *Lindsay Tape* (1967) as part of the landmark 1968 exhibition *The Machine as Seen at the End of the Mechanical Age*, organized by K.G. Pontus Hultén. The piece consisted of two half-inch reel-to-reel decks that were spaced 10 feet apart, with the tape (Paik's original!) jerry-rigged together to allow it to loop continuously. After a week on view, the wear on the tape proved too much; it began to break down and was taken off view (and was almost lost to history). Despite this rather daunting introduction to the fragile and fugitive nature of video, the Museum began to formally acquire video works in the late 1970s, led by former MoMA Associate Curator Barbara London.

View of Nam June Paik's *Lindsay Tape*, part of the exhibition *Machine as Seen at the End of the Mechanical Age*, November 27 1968–February 9, 1969. Photo: James Mathews Photographic Archive. The Museum of Modern Art Archives, New York.

Video, as a medium, is inextricably tied to an industry built on rapid advances in technology; as breakthroughs in resolution and portability lead to older formats being quickly supplanted. These advances also bring about

Artists
Behind the Scenes
Collection & Exhibitions
Design
Events & Programs
Family & Kids
Film
Learning and Engagement
Library and Archives
MoMA PS1
MoMA Stores
MoMA Teen Takeover
Publications
Tech
This Week at MoMA
Videos
Viewpoints

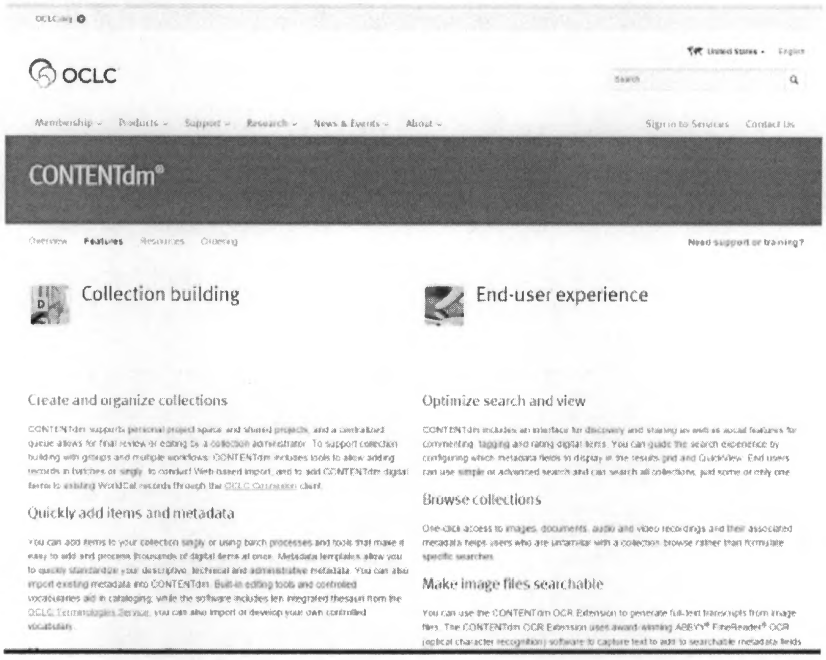
MoMA STORE
$25 Off and Free Shipping on Orders of $125+
SHOP NOW ▸

Find more in the Archives

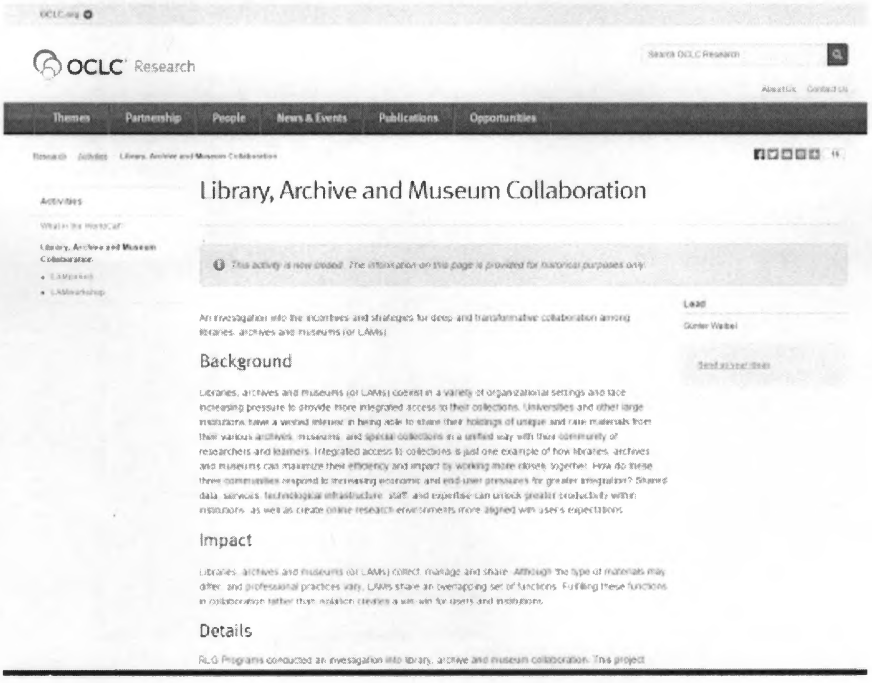
MOST VIEWED | RECENT POSTS

MoMA VISIT EXPLORE LEARN SUPPORT SHOP MoMA PS1 Search

D.70 Online Computer Library Center (OCLC). "Contentdm." 2015.



D.71 Online Computer Library Center (OCLC). "Library, Archive, and Museum Collaboration." *OCLC Research.* Last updated November 30, 2011.



D.72 Owens, Trevor. "Archivematica and the Open Source Mindset for Digital Preservation Systems." *The Signal* blog for *The Library of Congress*, October 16, 2012.

D.73 Personal Archives Accessible in Digital Media (Paradigm). "Selecting the Right Preservation Strategy, Other preservation Approaches: Encapsulation." *Paradigm.ac.uk*. Last updated January 2, 2008.

D.74 Pogrebin, Robin."The Met's Director Looks Ahead." *The New York Times*, March 19, 2014.

The New York Times

ARTS MUSEUMS SPECIAL SECTION

The Met's Director Looks Ahead

IN the five years since Thomas P. Campbell became director and chief executive of the Metropolitan Museum of Art, the museum world has had to grapple with declining philanthropic dollars, mushrooming entertainment options and a rapidly developing digital world. But in an interview, Mr. Campbell said he was optimistic about the public's appetite for museums in general and the Met's prospects in particular. This interview has been edited and condensed.

Q. *What is the biggest issue facing museums today?*

A. Impacting all of us is technology. We've made a huge investment in transitioning from being an analog museum to a digital museum and there are great opportunities in that to see the collections on the whole, to deliver the information to our audiences in new ways. Still, at the end of the day, the core values remain the same: It's about bringing people face to face with works of art and stimulating their curiosity.

Q. *Can museums really hope to compete with computers for people's attention and leisure time?*

Don't miss a single story. Try a Times digital subscription for 99¢

D.75 "SCAPE Catalogue of Digital Preservation Policy Elements." Last modified May 19, 2014.

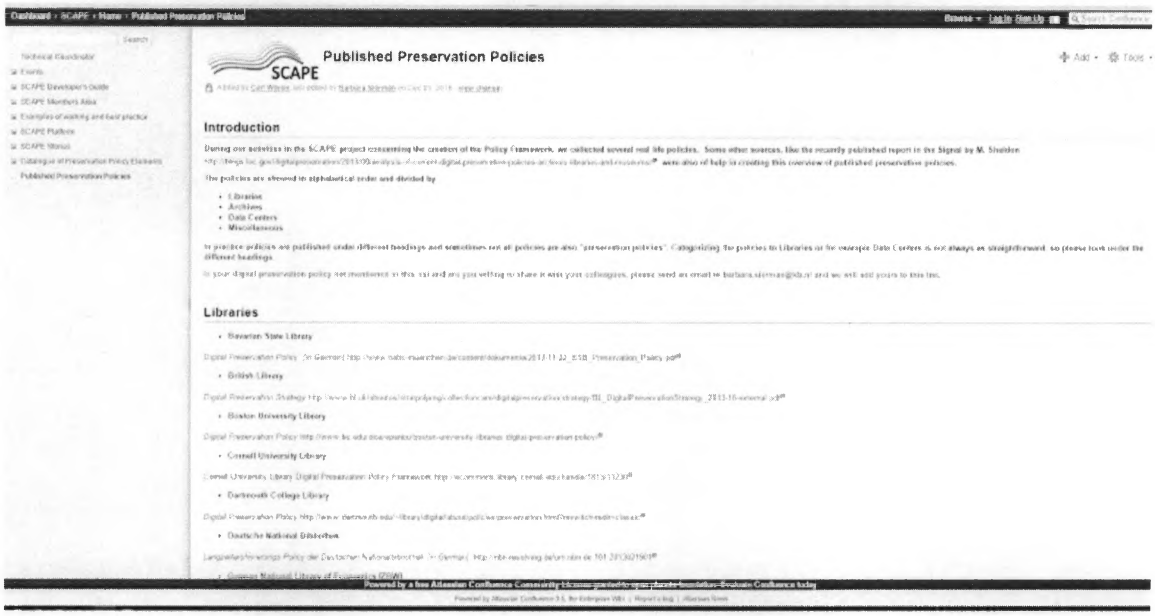
SCAPE

Catalogue of Preservation Policy Elements

About the Catalogue of Preservation Policy Elements

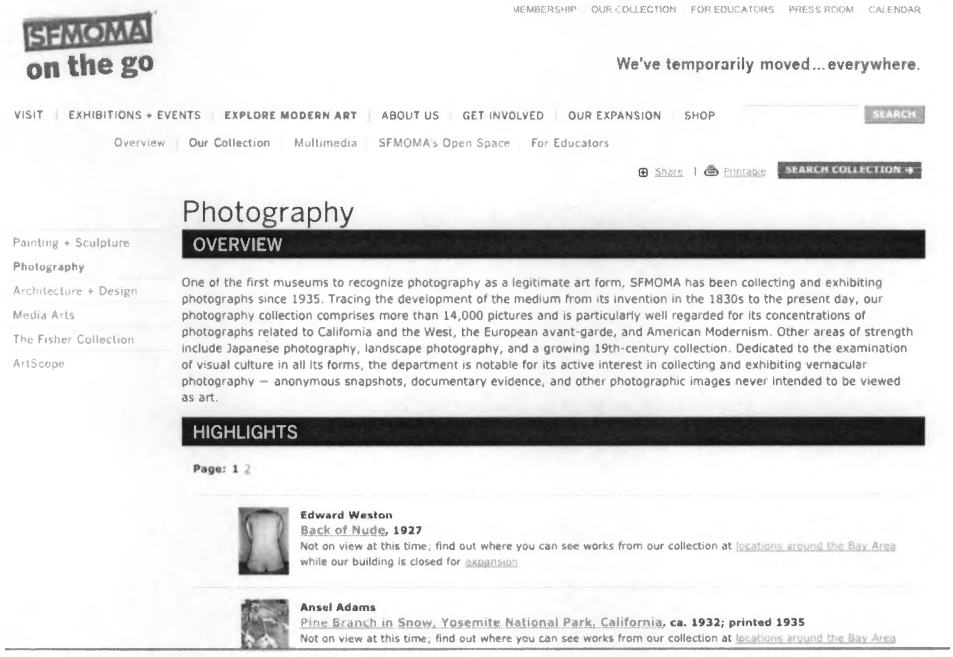
Structure of the Catalogue of Preservation Policy Elements

D.76 "SCAPE Published Preservation Policies." Last modified December 1, 2015.

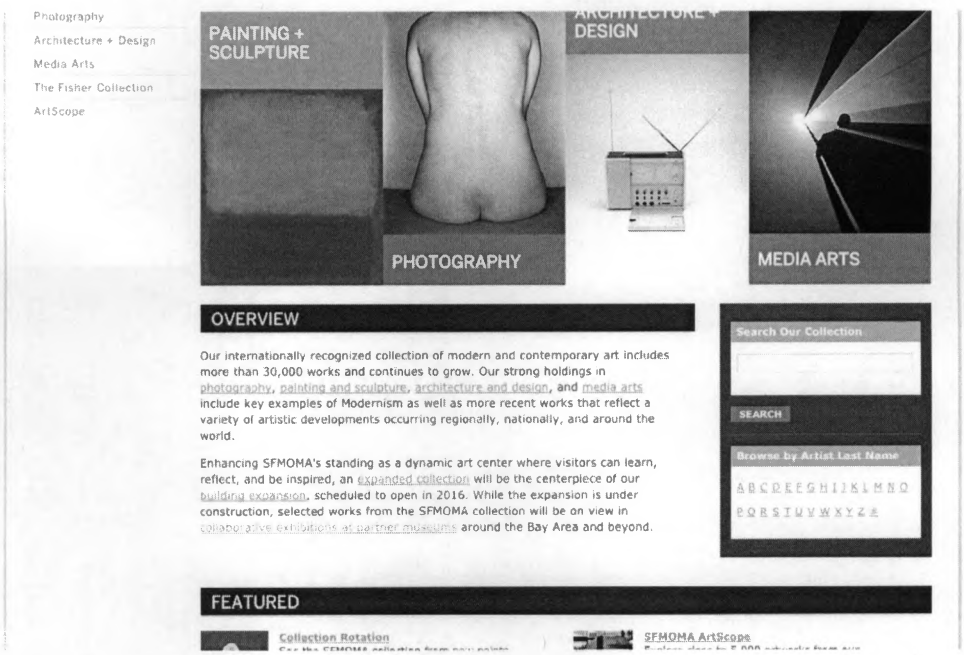


D.77 San Francisco Museum of Modern Art. "About SFMOMA." *Sfmoma.org*.

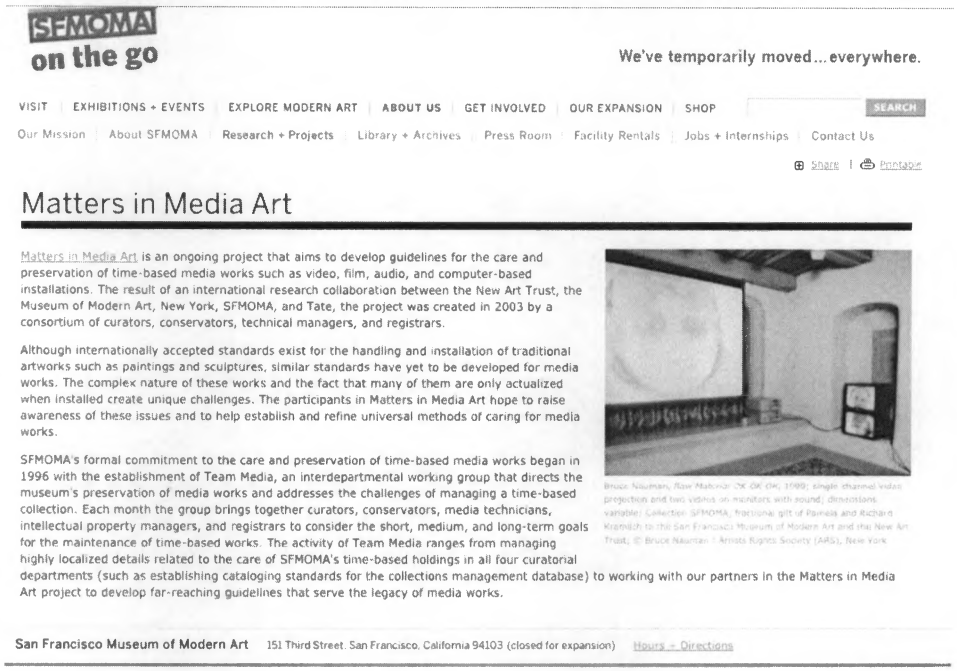
D.78 San Francisco Museum of Modern Art (b). "Photography Collection." *Sfmoma.org.*



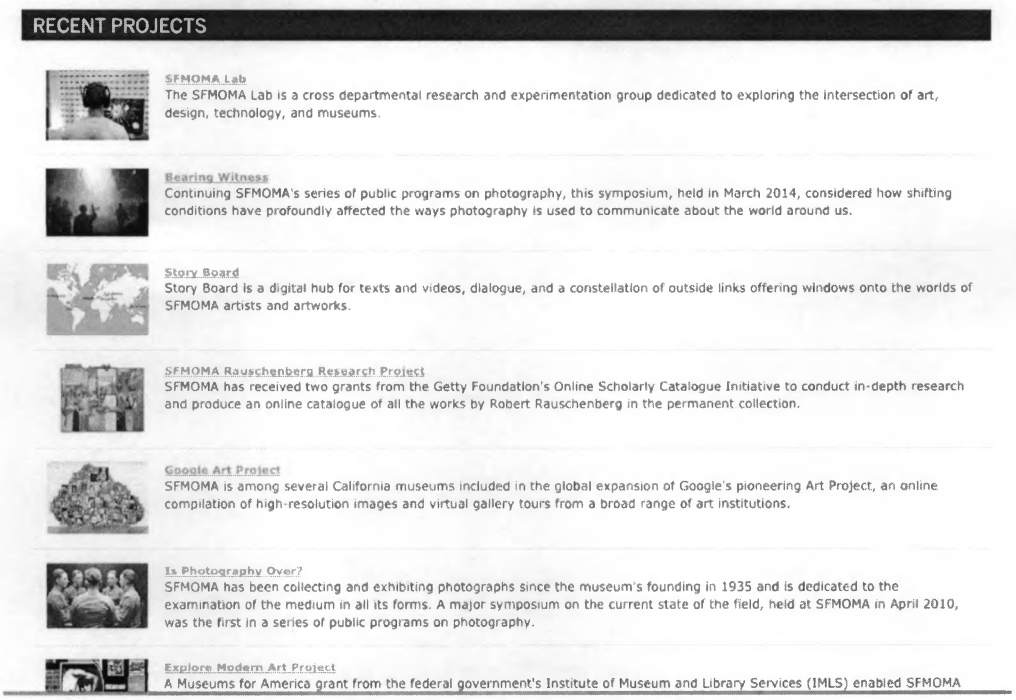
D.79 San Francisco Museum of Modern Art (c). "Explore The Collection." *Sfmoma.org*



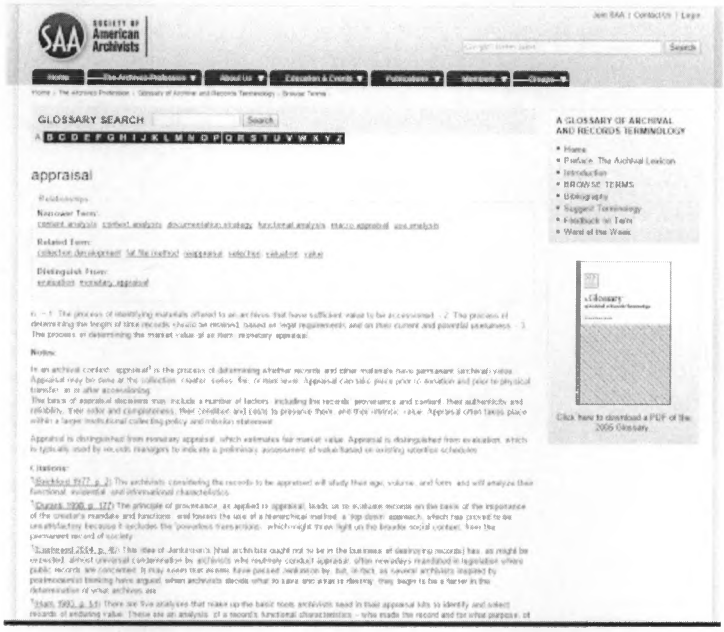
D.80 San Francisco Museum of Modern Art (d). "Matters in Media Art." *Sfmoma.org*.



D.81 San Francisco Museum of Modern Art (e). "Research and Projects." *Sfmoma.org*.

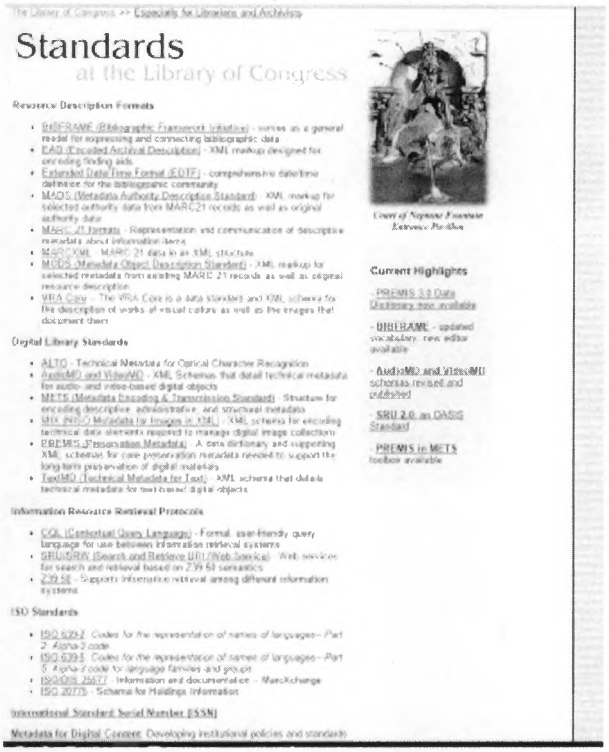


D.82 Society of American Archivists (SAA). "Appraisal." *Glossary of Archival and Records Terminology*. 2015.

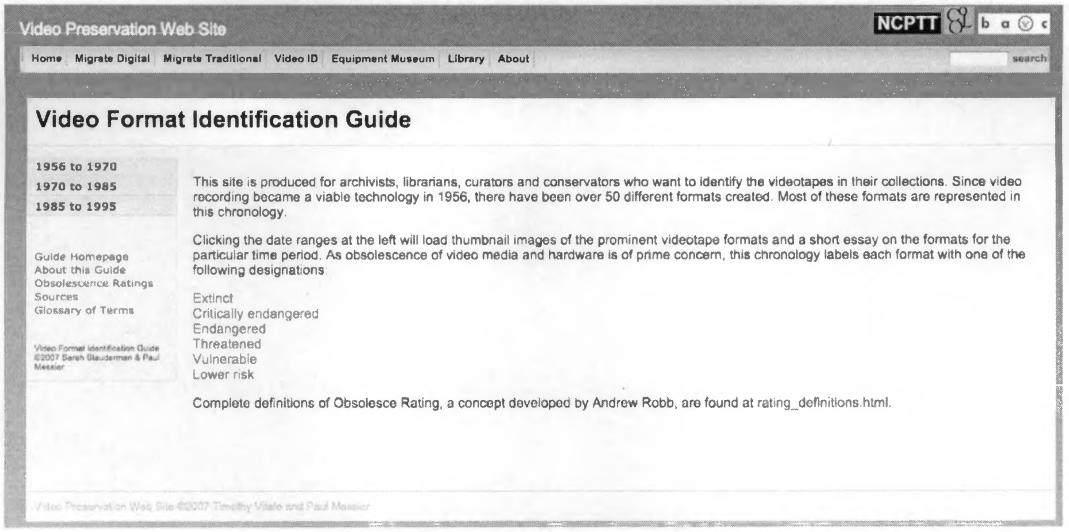


D.83 Society of American Archivists (SAA) [b]. "Encapsulation." *Glossary of Archival and Records Terminology* from *Archivists.org*. 2015.

D.84 "Standards at the Library of Congress." *The Library of Congress*. Last Updated July 13, 2015.



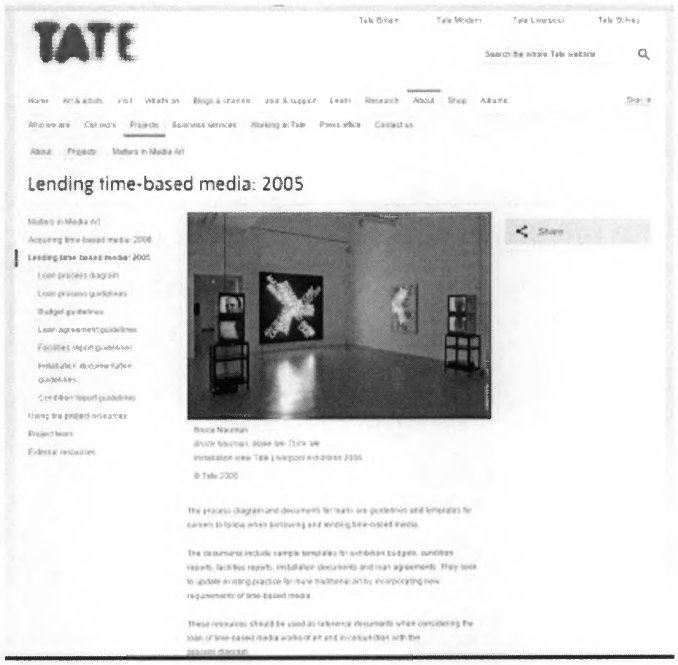
D.85 Staudeman, Sarah and Paul Messier. "Video Format Identification Guide." *Video Preservation Website (VPW)* of Stanford University. 2007.



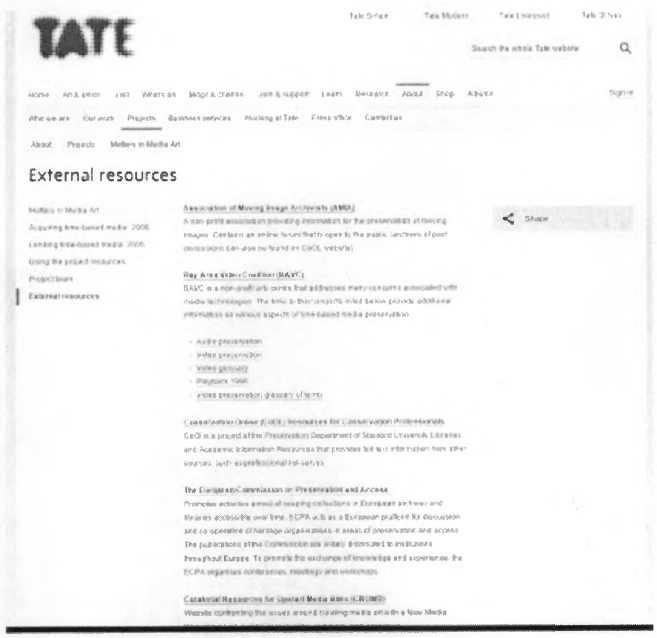
D.86 Tate Modern. "Matters in Media Art." *Tate.org.* Last updated December 2015.

D.87 Tate Modern (b). "Post-Acquisitions." *Matters in Media Art from Tate.org.* Last updated December 2015.

D.88 Tate Modern (c). "Lending Time-Based Media Art." *Matters in Media Art from Tate.org.* Last updated December 2015.



D.89 Tate Modern (d). "External Resources." *Matters in Media Art* from *Tate.org.*



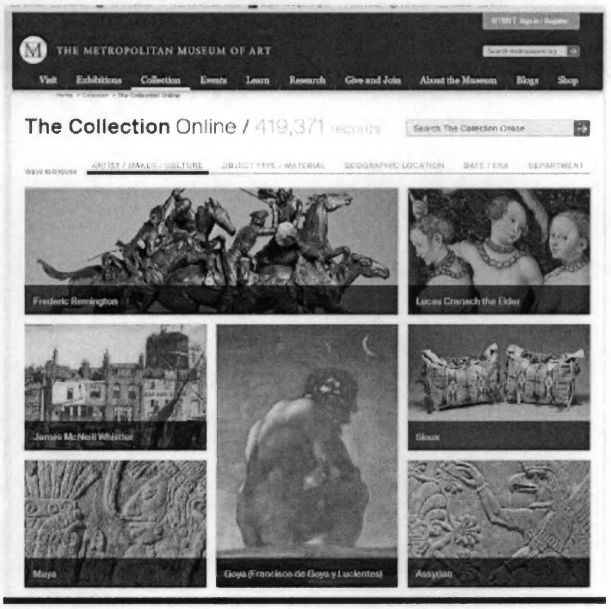
D.90 The Metropolitan Museum of Modern Art (MET). "Main Building." *History of the Museum* from *Metmuseum.org*. 2015.

D.91 The Metropolitan Museum of Modern Art (b). "Museum Mission Statement." *Metmuseum.org*. 2015.

D.92 The Metropolitan Museum of Art (c) "Thomas J. Watson Library." *Museum Departments, Office of the Director* from *Metmuseum.org*. 2015.

D.93 The Metropolitan Museum of Art (d) "Museum Archives." *Museum Departments, Office of the Director* from *Metmuseum.org*. 2015.

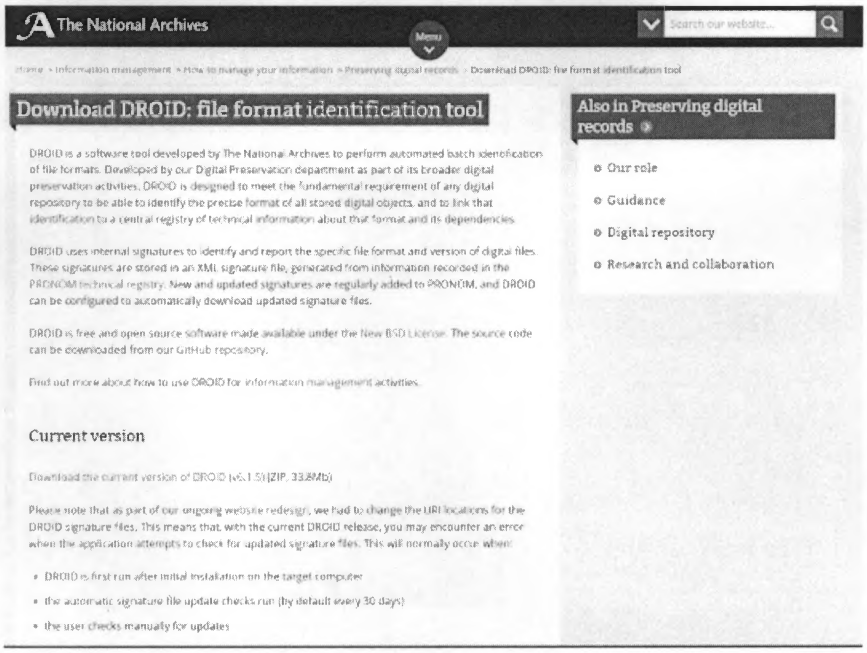
D.94 The Metropolitan Museum of Art (e) "The Collections Online." *Metmuseum.org.* 2015.



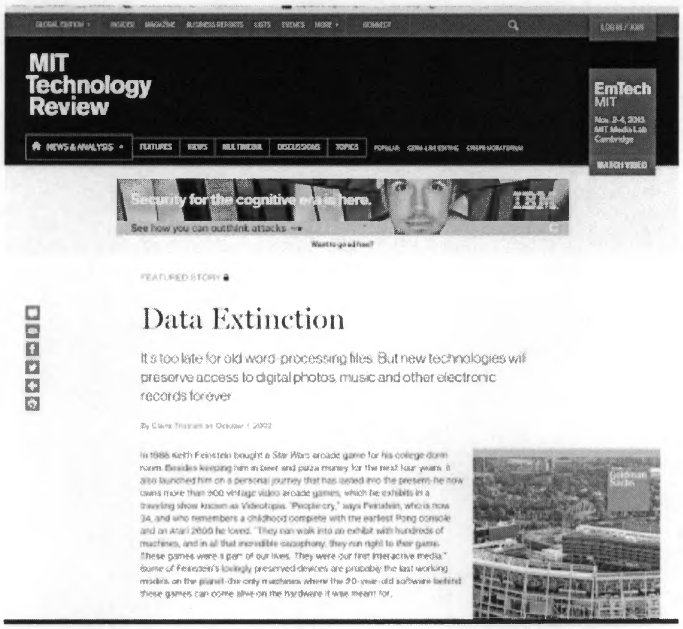
D.95 The Metropolitan Museum of Art (f) "Collection." *Metmuseum.org.* 2015. Accessed September 27, 2015.

D.96 Thomas J. Watson Library. "Digital Collections from The Metropolitan Museum of Art Libraries." *Digital Collections, Thomas J. Watson Library, The Metropolitan Museum of Art*. 2015.

D.97 The UK National Archives."Download DROID: file format identification tool." *The National Archives*. 2015.



D.98 Tristram, Claire. "Data Extinction." *MIT Technology Review*, October 1, 2002.



D.99 UC Berkeley School of Information Management and Systems. "Executive Summary." *How Much Information? 2003.*

D.100 Wall, Helen D. "Picturing Met Museum through Visitor's Eyes." *Digital Underground* blog, April 16, 2015.

D.101 Wheatley, P., Andy Jackson, and Andy Tester (contributors). "Main Page: Community Owned digital Preservation Tool Registry." *COPTR*. Last updated November 28, 2014.