ENCODING POWER: THE SCRIPTING OF ARCHIVAL STRUCTURES IN DIGITAL SPACES USING THE OPEN ARCHIVAL INFORMATION SYSTEM (OAIS) REFERENCE MODEL

BY

RHIANNON STEPHANIE BETTIVIA

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Library and Information Science
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2016

Urbana, Illinois

Doctoral Committee:

Associate Professor Jerome McDonough, Chair and Research Director
Associate Professor Kevin Hamilton
Assistant Professor Anita Say Chan
Associate Professor Kari Kraus, University of Maryland, College Park

ABSTRACT

The Open Archival Information System (OAIS) Reference Model is a corner stone of the evolving discipline of digital preservation. It undergirds many of the systems that are used in daily practice in organizations engaging in digital preservation, and directly or indirectly influences commercial and open source tools, as well as administrative and personnel functions. Yet it also undergirds the field of digital preservation more generally as its design and revisions have taken place concurrent and in partnership with the growth of the profession into one that has boundaries, curricula, and standards of practice. There is closure around OAIS: it is ever assumed as background or foundational knowledge in new endeavors. It is a black box. Given the pervasiveness of OAIS, this research examines the values scripted into the sociotechnical object it represents. Using discourse analysis, this research traces the power relationships that arise as a result of the discourses that OAIS produces. I also explore the effects on professional practice that occur because of the discourses OAIS brings with it from its scientific origins and archive-informed terminology. The dissertation investigates terms like *Designated Community* and *significant properties* in order to lay bare the imperial tendencies scripted within OAIS as well as to expose the resistive and recuperative potential of this technology.

ACKNOWLEDGMENTS

I would like to thank the Graduate School of Library and Information Science at the University of Illinois, Urbana-Champaign for its generous support of my research over the past five years. I especially thank Linda Smith for unendingly immediate assistance with all matters administrative and research-related. Without her, this would not have been possible.

I also thank my academic mentors for their assistance with this project and many others. Leigh Estabrook, I thank you for your extreme generosity with your time and hospitality. Your pragmatic writing and career advice have made these last two years possible. I thank my committee members for their gracious spirits. Anita Chan, I so appreciate the boundless enthusiasm you have for all of my projects. Kevin Hamilton, I am grateful for the support of the community you allowed me to be a part of through your Systems grant. Kari Kraus, thank you so much for your kind words and thoughtful criticism on this project and during my early years as a graduate student when I benefitted from your wisdom during PVWII. Your kindness in that space was invaluable.

Last and most important of all, I extend my thanks to Jerome McDonough. Jerry, you are everything an advisor ought to be. Thank you for your time, for the research opportunities, for your care, for your books, for your professional contacts, for the use of your office and truly epic headphones, for your collegiality and friendship. You supported me and my research before I even set foot on this campus and it has been an unmitigated pleasure.

I also thank my peers who have supported this journey. Thanks to Ana Lucic for taking care of my spouse and son when I was away doing doctoral research. To Karen

Baker, your eternal optimism is something that I aspire to in life. To Tiffany Chao for sending me links, contacts, articles, and the most beautiful copy of *Le Petit Prince* in my collection. I thank Nikki Weickum and Paul McKean for being my writing companions and for reading draft after draft of this dissertation and everything else I've written for the last two years. I literally could not have written this without you.

Finally, I thank my family and friends. Gail Ferguson, you are truly my comadre in every sense of the word. Thank you for offering help when I didn't know I needed it, for your constant calm and support, and for letting me be Mommy Rhiannon to your beautiful son. Thank you to my parents Angharad Valdivia and Cameron McCarthy for your unending support and for being the model of thoughtful, respectful collaborators with your students and colleagues and teaching me about the kind of academic I want to be. I am so thankful for the good fortune that brought me home to you a decade after I grew up and moved away. Thank you to my sister Ailín and my son Lucas for inspiring me every day: watching you grow is the greatest pleasure in life. Thank you to Alexis Meza for being my partner through graduate school, parenthood, and life. Thank you for the weekends of work time, for the sleep-in days, for trips to Europe, for supporting me in quitting my job in 2009 to pursue what I was really passionate about, and for agreeing to leave New York City to follow me to the Midwest. It has been a crazy adventure and I will always be grateful for it.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Introduction

In 2015, Marcia Bates helpfully, and perhaps not uncontroversially, mapped information professions within spectra of research disciplines. She begins by positing[1] that information disciplines, like education and communication, are meta-disciplines: they exist outside the spectrum of research domains that spans from arts and humanities through social sciences to natural sciences and mathematics.

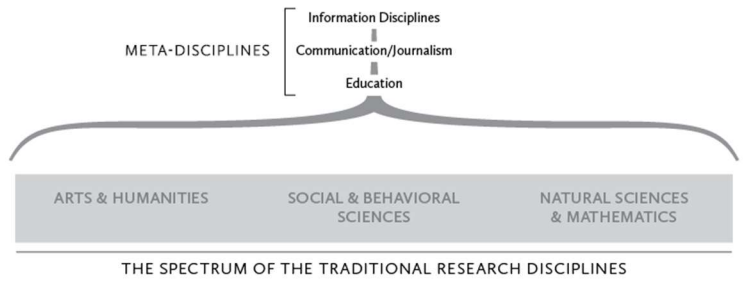

Figure 2. The meta-disciplines shape the subject matter of all the traditional disciplines according to the social purpose of the meta-discipline.

Figure 1: The Spectrum of the Traditional Research Disciplines (Bates, 2015)

Bates zooms in on information disciplines twice: first to the general fields contained therein and then to more specific subfields revealed as a list spread across another spectrum:

[1] Again, as the case may be, as she has done so in previous literature helpfully cited in the article.

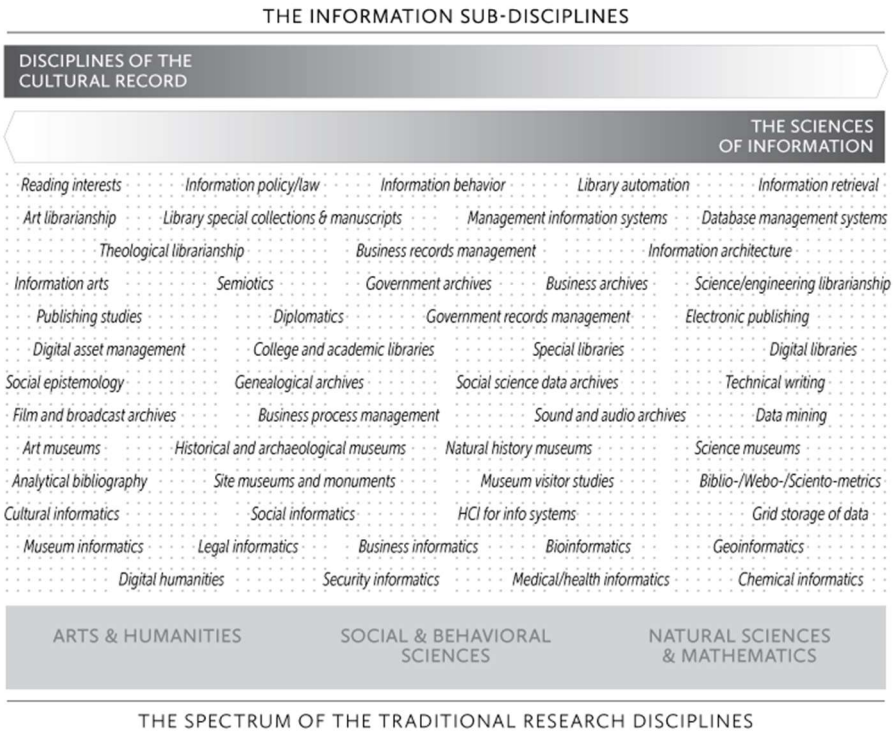

Figure 4: The information sub-disciplines.

Figure 2: Information Sub-Disciplines (Bates, 2015)

This list is not complete, to be sure. One missing component noted by my colleagues at the Graduate School of Library and Information Science at the University of Illinois, Urbana-Champaign was history. Another is my own area of expertise: preservation and, more specifically, digital preservation. Like history, I explain the absence of (digital) preservation as an admission of its own invisible omnipresence: it is imbricated within the work of many of these information science sub-disciplines, ranging from the designated humanistic concerns such as informatics of various flavors and diplomatics to the natural-scientifically- and mathematically-oriented concerns of digital libraries and data mining. Like information work more generally, preservation is a meta-concern that finds itself spread across the LIS spectrum, and therefore, according to Bates' theory of information

professions as meta in and of themselves, across the entire pantheon of traditional research disciplines.

Thus, it should not be surprising that preservation as a field of study and practice should borrow from a range of disciplines and professions: in the digital preservation realm, archival practices blend with systems design and computer sciences. Because preservation borrows from all of these areas, so too does this imbue the field with a rich mix of discourses: those brought from all of the disciplines that converge to form the field of digital preservation; those that are formed within the sociotechnical networks of the broad field comprised of practitioners and theorists; and those that arise in particular locations where digital preservation work happens on a daily basis.

This dissertation constitutes a first step in understanding the discourses at play within the field of digital preservation. It examines the ways in which digital preservation happens in a variety of institutions by investigating the use of an ubiquitous reference standard for digital preservation, the Open Archival Information System (OAIS) reference model. It analyzes this work in order to answer the following research questions:

1. What values are contained within the Open Archival Information System (OAIS) reference model and what methods or practices are prescribed by it?
2. In what ways has the adoption of this model as an organizational system for the preservation of digital content in library, archive, and museum spaces served to challenge or reproduce the hierarchies and discourses of traditional archives and memory institutions?

I examine the implications, practical benefits, and pitfalls of using a tool imbued with functionality and meaning firstly by the creators, the Consultative Committee for Space Data Systems (CCSDS); and secondly by users in memory institutions, whose practice with the model is both shaped by authorial intentions while simultaneously shaping the model for future users[2].

In this way, it becomes necessary to understand the implications of using the OAIS reference model in a variety of spaces, far from its space science home but not so inconceivably far from its lofty aims for broad use, which were a part of its design from the outset. The term "open" in the Open Archival Information System refers to this very element of the creation process: the creators were open to and sought input from a number of collaborators in many fields and disciplines (Ball, 2006; Lee, 2005 & 2009). It is now called the de facto standard for preservation (NSFCC, 2007), and this status is part of why a study of OAIS constitutes a beginning effort to examine discourse within the broader realm of digital preservation.

The Open Archival Information System reference model is purposefully high level, designed to inform the construction of preservation repositories for digital content. Initial drafts were released for review in 1997 and 1999; the Blue Book, an ISO standard, was released in 2002; most recent updates were submitted for ISO consideration in 2012[3] (CCSDS 2002, 2009, 2012; Lavoie, 2004). The significance of the changes depends greatly on the user. At each revision, a fairly minor portion of the text details the standard changes.

[2] "As users engage with technologies… they dynamically co-construct content and the technology itself." (Fuchs, 2008)

[3] As an ISO standard, OAIS undergoes review and revision on a 5-year cycle. The 2012 Magenta Book release will be up for revision again in 2017.

Some interviewees see these changes as important, while for others they are fairly trivial[4]. As a reference model, it does not involve particular recommendations or instruction on how preservation activities should unfold within an organization. Rather, reference standards operate at high levels of abstraction, to "frame future working strategies by identifying high-level issues related to a particular problem" (Seles, 2016, page 46).

OAIS describes entities in a repository, their general roles, and their relationships to one another. OAIS covers both an information model, denoting the types of content to be stored (the files one is saving) and how this information could be arranged and packaged, as well as a functional model that speaks about the positions that entities (various human and departmental actors) need to take, such as who is or should be in charge of general daily and long-term tasks within the repository.

The basic information model for storing content in an OAIS repository involves the construction of information packages: one for submission into the repository, the Submission Information Package (SIP)[5]; one for storage, the Archival Information Package (AIP); and one for sending the information back out again, the Dissemination Information Package (DIP). These packages combine the Data Object itself with Representation Information, which comprises the information needed to render digital data comprehensible to a human being, and Preservation Description Information (PDI), which

[4] One interviewee in particular was keen to begin the discussion of potential revisions for the 2017 version, expressing a hope that the creators would create a public and moderated forum for discussion about potential changes. This desire was partly motivated by her chagrin at the public Twitter "bashing" of OAIS by some of her colleagues. The result of this can be seen at the dpconline.org wiki, which presents an annotate-able version of the OAIS documentation.

[5] Capitalized terms in this dissertation refer to OAIS-specific terms that appear within the OAIS documentation, and more particularly, its glossary.

covers information about the object, including Provenance, Fixity, Context, and Reference data. The PDI serves the additional role of documenting information necessary to determine whether the authenticity and integrity of the digital object have been maintained by the Archive, a function rendered all the more salient as new international standards governing trustworthiness audits are created with OAIS as an underlying foundation. Altogether, these packages should be composed of what is necessary to render the intellectual content comprehensible to a Designated Community of potential users. The idea here is that a particular set of users will have a certain amount of shared background knowledge that allows them to understand particular objects, and that information stored in the OAIS repository should take advantage of this. Some additional pieces of information may need to be constructed and saved in order for an object to be comprehensible; for example, in order for people to play the 1985 version of *Carmen Sandiego*, an institution will need copies of 1985 maps as geographical borders have changed. With a target audience that has shared understandings, there are also some pieces of information that do not need to be saved as that knowledge already exists within the Designated Community. For example, if the Designated Community speaks English, it is not necessary to include an English language tutorial along with a file written in English (CCSDS, 2012; Ball, 2006).

The functional model includes roles for a variety of entities, including ingest, archival storage, data management, administration, and access. The basic responsibilities of an OAIS-compliant repository are distributed among these groups. The functional model is quite complicated, and exact adherence ranges from difficult to impossible depending on whom you ask.

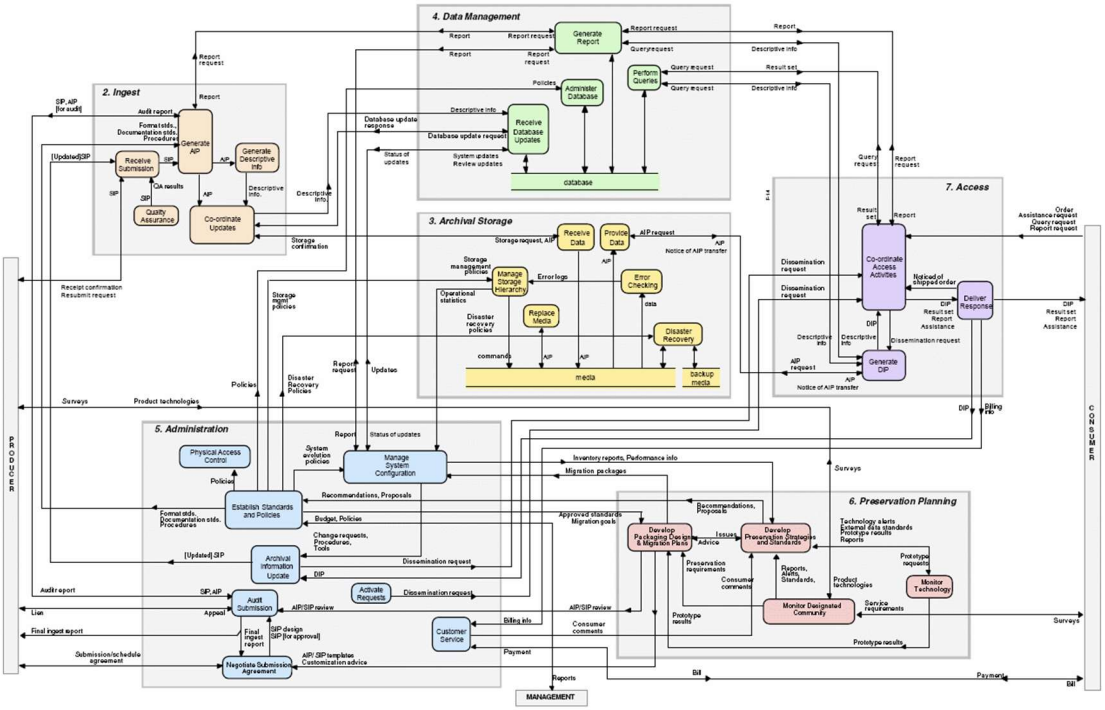

Figure 3: Composite of Functional Entities, CCSDS 2012, page A-2

In fact, in order to be in compliance with this standard, there are considerably fewer requirements than the above figure would suggest. To begin with, OAIS doesn't suggest an individual person to deal with each of the entities laid out above, nor does it suggest that an institution must have separate departments for each of the groups of tasks. Rather, OAIS indicates a set of concerns and the relationships among them: it is possible that in a very small institution, many or even all roles will be filled by a single person. OAIS lays out just six criteria (capitalizations are original to the documentation, denoting terms with special definitions within the standard documentation) for compliance, none of which require adherence to the information or functional models:

- Negotiate for and accept appropriate information from information Producers.

- Obtain sufficient control of the information provided to the level needed to ensure Long-Term Preservation.
- Determine, either by itself or in conjunction with other parties, which communities should become the Designated Community and, therefore, should be able to understand the information provided.
- Ensure that the information to be preserved is Independently Understandable to the Designated Community. In other words, the community should be able to understand the information without needing the assistance of the experts who produced the information.
- Follow documented policies and procedures which ensure that the information is preserved against all reasonable contingencies, and which enable the information to be disseminated as authenticated copies of the original, or as traceable to the original.
- Make the preserved information available to the Designated Community.

(CCSDS, 2012 page 3-1)

Across different disciplines and locations of use, interpretations of what precisely is required by OAIS differ greatly. "Misunderstandings", readily acknowledged to exist by authors of the standard and those who use or study it, cause of a fair amount of tension between groups of stakeholders. In this case, "misunderstandings" refer to two particular but distinct things. First, it is a source of frustration to many of the creators of OAIS, as well as those involved in its continued growth and propagation[6], that there are so many

[6] This includes people who worked in the creation and review process for the generation of the original standards; those who have been party to its on-going revisions per the ISO schedule; and those involved in the creation of related standards, such as TRAC and ISO

misconceptions about what OAIS actually asks of a preservation repository or Archive. Some people claim it is virtually impossible to have a real-life implementation that is compliant with a high-level and theoretical standard or that it is impossible by definition for anything to be compliant with a reference model; simultaneously, others claim that the list above makes it so absurdly simple to have a compliant repository that "even a chicken with its head cut off could be OAIS compliant". This particular quote is oft repeated, particularly by the creators themselves; in fact, it was mentioned during the interviews I conducted. A practitioner stated this belief to some members of the creative team behind OAIS at a conference early in OAIS' life cycle, to perhaps note the perceived generalness or vagueness of the six requirements. This was sufficiently irksome to have made it into preservation texts as a caution: the OAIS creators make it clear that a decapitated chicken cannot, in fact, be OAIS compliant (Giaretta, 2011). This interchange does highlight the tensions that naturally occur when using a model that is purposefully high-level and non-specific to model site-specific practices that are highly detailed and involve potentially many actors.

There is a second way in which differing interpretations of the model's tenets, some of which might legitimately be described as misunderstandings, are a source of frustration between practitioners and creators. Parts of OAIS that are difficult to understand, or those parts for which the rationale is not transparent enough, coincided in interview data with those elements of the model that people complained about the most. Schumann and Recker (2012) contextualize what compliance, which implies some level of implementation that

16363, standards for assessing and auditing the 'trustworthiness' of digital repositories that are explicitly modeled on OAIS.

entails a high degree of fidelity, even means for something that is so purposefully high-level and abstract that it cannot match perfectly to real world examples by definition. This is important because many institutions speak about a desire to be compliant, which is seemingly at odds with a model that claims not to be prescriptive. Daniels succinctly makes the point: "For all its clarity, however, the OAIS is a theoretical model that intentionally avoids any statement on how to go about creating a compliant digital archival repository" (Daniels, 2013, page 12). Certain concepts that are less technical and more social give the most difficulty in both the literature and interview data. Despite detailed work clarifying the role of significant properties within OAIS (Giaretta el al, 2009), interviewees and other literature about the use of OAIS remain dissatisfied with the creators' treatment of significance (Sierman, 2012). Many institutions struggled especially with the Designated Community notion (mentioned three times in the six basic requirements alone), particularly those working in more public spheres, such as libraries and public- and state-sponsored cultural heritage bodies. Explanations about this term offered by creators in interviews differed widely from those expressed by practitioners in interviews.

Why is a study of OAIS fundamentally a part of a larger project to understand discourses within the general field of digital preservation? The reasons for beginning here are manifold. Conditions such as timing are key features of OAIS' success. In terms of timing, OAIS was created and circulated while the iron was hot in a very real way: the entire field of digital preservation has grown around it. This explains its very pervasiveness: it came about at a time when the crush of digital materials was felt acutely by organizations of all types and the need for standard operating procedures became necessary. At a time when many were creating ad-hoc ways of dealing with content, OAIS

arrived and not only changed the discussions, but shaped all future ones. At this point, it has become the lingua franca, as one interview participant put it. OAIS is around even when it is not: two different sets of interview participants at two different institutions, one a private college and the other a major museum, claimed they did not use OAIS, but later admitted to using Archivematica, a digital preservation tool that is explicitly built on OAIS.

I want to address the implications of the ubiquity of the OAIS reference model here and speak to what is meant by implementing a model that is actually a set of very purposefully general guidelines. The benefits and the pitfalls of the OAIS reference model would be a moot concern if it were not as widely adopted as it is; after all, there are so many examples of wonderful tools that failed to gain any traction and many more examples of inferior tools that won the brand war, so to speak. Spence (2006) says:

> When 48 cultural institutions from 13 countries were asked by OCLC's Preservation Metadata Implementation Strategy (PREMIS) Working Group 'How is your preservation repository informed by the Open Archival Information Systems model (OAIS)', 80 per cent considered that their repositories were 'informed' by OAIS or at least 'partly conform to the model' (OCLC/PREMIS Working Group, 2004).

In the same article, Spence also discusses the possibility of adapting OAIS for institutions as small as local gardening clubs in Australia. Ball (2006) lists three pages of major repositories that are OAIS-based across a number of disciplines and institution types. Literature on the topic of implementation and compliance abounds (Ayoung and Tibbo, 2011; Vardigan and Whiteman, 2007; Daniels, 2013). The first page of search results on the topic using the search function at the University of Illinois Library contains results in Portuguese, Slovenian, and Dutch/Flemish as well as implementations in India, the UK, and

Germany[7]. The literature shows no loss of momentum: looking at bibliographic entries that mention OAIS in the title, abstract, and/or keywords reveals new work at a fairly consistent clip through to the present, including Seles' (2016) excellent dissertation on the use of digital repository standards in East Africa. I also argue that the on-going development of standards like ISO 16363 only serve to renew interest in OAIS when it begins to wane. If nothing else, an ISO 16363 audit would require an institution to translate its daily functions into OAIS language, whether the audit is official, as required by many national bodies in Europe, or informally conducted for internal planning, as is the case in most of my US case studies.

This dissertation shows how some people and institutions have used OAIS in the field in a way that also highlights the things they want OAIS to do. It also begins to explore whether or not OAIS actually does some of these things by bringing in data from the Preserving Virtual Worlds II (PVWII) grant to examine how well user-defined significant properties of video games fit into the 2012 version of the OAIS model. There is a tension between the use of OAIS and the thing itself; there is similarly a tension between what its creators designed it to do and what happens when it is used out in the wild. OAIS is supposed to be about standardization, but the data captured from interviews indicates that this does not happen in practice: there is no consistency of deployment from institution to institution, or even of internal understanding of basic terms that are explicitly defined in the OAIS glossary of terms. This is not to suggest some sort of failure on the part of OAIS or its creators. Critical claims arising from this dissertation are not criticism of the creators of OAIS nor the people and institutions who base their practice on it in some measure. Tracing

[7] However, it is perhaps telling that UIUC's JSTOR search returned only one article.

socially informed discourses embedded in systems speaks to larger, historical dominance and oppression and does not reflect personal biases.

Rather, this work serves to raise the questions:

- Where does this variance come from?
- What about OAIS incites this, beyond the changes that obviously come from the variety of institutional and geographical settings?
- What does this say about the practice of digital preservation?
- And, finally, what do these struggles indicate about what practitioners are trying to do when they have to 'do' digital preservation?

The field of digital preservation needs to contend with how these things are realized as it continues to grow and seeks to make machine-actionable the large selection tasks inherent to creating comprehensive archival packages. There are spaces of negotiation and spaces where the model does not give, whether this latter is because the model is somehow unyielding or because practitioners do not ask it to yield. This supports the notion of the tension mentioned above, and spurs this research to find places that yield and places that don't, and to ask what all of this means for the growth of this evolving field.

1.1 Motivations

I submitted a very early version of the proposal for this thesis to a major academic conference that deals in the largely technical realms of digital preservation. On Bates' spectrum above, this professional body is comprised of the outer edge of the social science information professionals, and the leaning is distinctly towards the natural science and mathematical end. In proposing to investigate OAIS and the discourses around it, as opposed to further examinations of its practical deployment in various institutional

settings, I received the following feedback from a peer reviewer: “I am not aware of any significant power differentials among digital preservation practitioners… I'm not aware of the OAIS model being very controversial; in fact its status as an ISO standard point[s] otherwise...” This review inadvertently reveals one of the reasons why this type of research is necessary: this object exists so pervasively and with a relatively low level of controversy for something in such ubiquitous use. As Flanders (2013) says about adding the digital to literary studies, “The introduction of this layer of formally expressed knowledge into the scholarly ecology creates a burden of responsibility to understand how that layer works and what it is saying, or at least to take its existence seriously.” The overwhelming presence of OAIS, not just in many institutions as noted earlier, but also as a constraint on the entire field of digital preservation, has served to render it somewhat invisible. I mean constraint not necessarily in a pejorative sense, but instead in a systems design sense: this is something that needs to be reckoned with in any digital preservation undertaking because the very field of digital preservation and its history as a codified profession is bound together with OAIS[8]. People, particularly younger practitioners in my interview set, do not see OAIS even when it is there. It has become like a scientific tool or technology: like a microscope or a telescope, people see what is on the other side without examining the scope itself very closely (Gitelman, 2006). Just as it is necessary, at some point, to examine the tool for its influence and role in the collection of data by seeing through the scope, it is necessary to understand the work that OAIS is doing, particularly when it is invisible. OAIS

[8] As one interview participant described the constraints of OAIS: “So, I mean when I said constraint, I really mean this in a very broad way… not in a sense of we felt constrained in this individual way, but in the systems design sense that a constraint is a condition that… restricts the solution space to some degree.”

socialized those within the digital preservation field into believing, as the review above appears to, that OAIS is depoliticized and neutral (Halavais, 2009). Resisting the temptation to think of OAIS this way is all the more necessary given OAIS' role in archival and memory environments, spaces which already (problematically) appeal to neutrality as a means of establishing their use-value and trustworthiness (Boyle, 2003).

As a (meta-)discipline, information sciences challenge assumptions of neutrality in a variety of technological spaces: video games, websites, virtual worlds, search engines, code and metadata[9]. Studies about the politics of standards exist within the realm of sociotechnical systems literature, yet there does not yet exist the same kind of critical engagement with OAIS. What is it about the archival and technological context that allows the body of literature to continue to be about the implementation of a neutral object, simultaneously eschewing research about the object itself?

In fact, part of the difficulty here may be that it is not possible to separate what is embedded within OAIS from that which is embedded socially, historically, and politically in the practice of official memory creation. This occurs for a number of reasons, the simple umbrella of which is the fact that OAIS is less a thing in itself than a network, in the actor-network theory or sociotechnical studies sense. Further, it is a network that exists within other networks, some of which are very long-standing. In other words, I cannot, at least through this study, understand the discourses specific solely to OAIS, but rather I can understand situated discourses about practices in which OAIS is imbricated; this is another

[9] Video games (Nakamura, 2009); websites (Nakamura, 2002); virtual worlds (Kendall, 2002); digital media platforms (Chun, 2006; Brock, 2009); search engines, (Noble, 2012; Halavais, 2009); metadata (Presner, 2014; Clark, 2001)

way in which this study of OAIS serves as a very apt microcosm through which to begin to understand discourse within the more general realm of digital preservation. This type of investigation invites precisely the kind of research I have conducted here: in interviewing the subjects I found, I was able to ask them to characterize OAIS and its role in their work. To this question, my subjects provided many answers, both explicit and implicit.

The role of discourse is an important lens to bring to bear on this subject matter, precisely because discourse is bound up with the Foucauldian "regimes of truth" (1972) that are exercised through knowledge structures. Given the function of OAIS and its pervasiveness, I argue that OAIS acts as a knowledge structure through which numerous organizations operate. The social informatics of OAIS, in the words of Bishop and Star (1996), examine the "… social influences, processes, practices, and effects related to how knowledge is structured and communicated…"(page 305) in digital repositories that build their functions according to the OAIS reference model. These discourses speak about how institutions use OAIS in their digital preservation activities, beyond any prior functions they had as analog memory institutions, to structure knowledge internally and to communicate both inside and outside of the organization.

A discourse lens helps enable an examination of the ways in which a model like OAIS reinforces oppressive and exclusionary practices inherent to memory-making models and traditions. "The power of discourse means that it produces the things which it purports to be describing", (Fairclough, 2006, page 210) and this is what is so powerful about an object like OAIS that grows with and around the field of digital preservation. In this instance, it is tricky to know which is the creeper vine that grows on the tree and which is the tree itself. OAIS claims it is neither prescription nor proscription, and this is true in the sense that

OAIS is a reference model and thus contains no information about particular implementation. However, I assert that, through the lens of discourse, OAIS produces the very structures for which it purports to merely provide a framework. These products – the structures – necessarily conceive of various actors and audiences in particular ways, and this dissertation will begin to examine them. More importantly, if OAIS creates the structures within which preservation takes place and digital Archives are constructed, this raises the question: how can OAIS be employed to make an intervention in spaces of memory practice? Can OAIS be used to make better archives? This particular moment in the history of archives, the one that incorporates digital technologies such as open-source aggregators, post-custodial archives, and personal and community archive digital platforms, is often posited as inherently different than previous ones, although many scholars offer a caution to this. Selection issues, both for born digital content and in choosing what gets digitized, as well as a perceived ubiquity and completeness of the Internet of Things and Google create a problematic complacence that overlooks the voices lost in the furor, either because they are not findable in the digital deluge or because, in an act of symbolic annihilation, materials of marginal populations are not actively preserved by memory institutions with the resources to undertake the very active processes necessary for digital preservation[10]. Yet, while I argue that this archival moment is not

[10] "Symbolic annihilation, a concept first developed by feminist media scholars in the 1970s, describes what happens to members of marginalized groups when they are absent, grossly under-represented, maligned, or trivialized by mainstream television programming, news outlets, and magazine coverage." page 27, Gaye Tuchman, ''Introduction: The Symbolic Annihilation of Women by the Mass Media,'' in *Hearth and Home: Images of Women in the Mass Media*, eds. Gaye Tuchman, Arlene Kaplan Daniels, and James Benet (New York: Oxford University Press, 1978), 3-38 in Caswell, 2014.

hugely different than previous moments where the archive was paper- and building-bound, I also firmly believe that the necessary changes that accompany the introduction of digital preservation systems in memory institutions signal a space wherein it is easier to envision paradigm shifts. In recent years, existing employees have seen massive disruptions to workflow and labor practices as they are forced to accommodate the digital elephant in the room. Such changes result in the necessary creation of new roles and departments and the solicitation of new hires. All this disruption renders the politics of memory institutions softer and more flexible: while I reject the techno-deterministic claims that technologies themselves bring democratic change to practices of history and memory making, I simultaneously assert that this is a fertile moment to advocate for institutional and discipline-wide changes in approaches. I seek to ground this discussion in the emerging literature that looks to the politics of technologies by centering and rendering visible the hidden or implicit normative prescriptions therein. I see this work as part of growing discourse that examines ethics in algorithms, following work by Chun (2006), Presner (2014), Drucker (2013), and others. With regard to OAIS, it is necessary to understand what is inscripted in OAIS and the relationship of the model to site-specific implementations as being related to both the political provocations and dictates of the institution in which it is being deployed. At the same time, it is also necessary to understand the relationship of OAIS and its scripts to the type of content being preserved, both in a media archaeological sense wherein we look to the particularities of the media or mediated object itself, and in the sense of the object as boundary object, understood by user communities, Designated Communities, and preservationists (Sherratt, 2015).

While I advocate an examination of OAIS itself in addition to the work it supports, this project seeks to unpack the scripts inscribed into OAIS during its creation. Noble (2012) notes that discourses of technology use the terminology of 'the tool' in a way that focuses attention on the tool in a way that decenters the ways in which creators imbue their tools with human values. No technological invention is value-free or neutral. This language, when brought to bear on OAIS, brings to the fore a question that I ask of OAIS—what is it? Is it a technology or a tool or a machine? This is contested in the sense that interviews revealed a variety of different answers, some of which reflect different understandings of what OAIS is and other which reflect the different uses to which it is put in different organizations[11]. I will refer to OAIS frequently as a technology, and I mean this in a Foucauldian sense; I explain this concept in greater depth in Chapter 2. How does a technological system fit within the literature about platforms, search engines, and algorithms? In conducting the interviews and asking participants to characterize OAIS for me, it is possible to see how differently situated professionals characterize the technology. This type of investigation can lead to a definition of what OAIS is beyond simply calling it a reference model, which given its ubiquitous use in a variety of fashions is somewhat reductive.

If the vast majority of the existing literature about OAIS covers implementation—including how to, why to, why they did, and what happened—then it remains to be addressed what it actually means to use a high-level model. It does not much matter what the model was or was not meant for. The fact that OAIS is loaded with scripts does not

[11] This is similar to the contested definitions of digital libraries found by Bishop and Star (1996)—the kinds of services implied by the term *digital libraries*, the materials DLs give access to, the interplay with traditional libraries and library services, etc.

mean they will play out the way they were intended by the creators when OAIS actually gets used or implemented; users will always come up with their own scripts, and this is true of any technical object (Akrich, 1992). Beyond implementation, OAIS serves a larger goal in providing a common vocabulary about preservation roles and responsibilities, and it has done this so well that it has become a boundary object, used to define preservation terms and tasks across the myriad disciplines that now find themselves charged to preserve the content they create, which at this point in time is almost inevitably digital (Star and Griesemer, 1989). Meghini says "…the basic concepts underlying digital preservation, including its definition, are set by the Open Archive Information System (OAIS) Reference Model…" in an article that speaks to general preservation concerns for anyone doing any kind of digital preservation work. The article is not about OAIS itself, but instead states that it takes its base assumptions, concepts, and definitions from this model (Meghini, 2013, page GRDI15). This speaks to a type of ubiquity beyond its use in repositories themselves: for the general study, whether in practice in a particular location or not, OAIS has become the little black box of preservation. I use this term here to indicate an object whose presence is so pervasive in some ways that it has taken on a privileged space of decentered ubiquity: it is there, even when it isn't (Balsamo, 2011).

Schumann and Recker (2012) point to the fact that its wide use has led to numerous myths and misconceptions about what OAIS is meant to do—like a game of telephone, the more we all talk about OAIS, the easier it is to drift away from what it actually is or means. This drift is akin in some ways to the misconceptions and differing interpretations I spoke about earlier. I would argue that one of the key differences here is that while the aforementioned tensions are inherent to the deployment of any technological object, and

indeed any reference model, this drift is both a product of the very broad use and wide traction OAIS has garnered over the past period of almost two decades as well as a factor that helps enable its spread. The flexibility and all-encompassing nature of OAIS are its draw but, like a game of telephone, its wide and on-going circulation engenders misunderstandings (McDonough, 2008).

OAIS, as a reference model, is not meant to dictate answers to all the questions that arose in the context of my interviews. In fact, I would also say that this, too, may not matter. It is important to say that even if OAIS doesn't tell me how to construct or delimit nebulous objects or huge, indefinable audiences, both of these categories exist and denote content that needs to be preserved. Since this content needs to be preserved, and many preservation repositories are using OAIS to do this work, this kind of information must somehow be made to work with OAIS. Even if this does not end in a modification to OAIS itself, it is something that needs to be modeled in relationship to OAIS so preservationists and archivists working with these complicated collections and with OAIS-informed repositories can do their jobs and do them well.

This project has and will often reference archives and archivists, and thus it is necessary to clarify what I mean by these in this paper. The politics of archives play a large role in motivating this study and trying to understand the ways in which digital practices intervene in or replicate politics found in paper and other analog archival spaces. Most obviously, archive is in the title: it is the A in OAIS. More broadly, however, this project focuses on the history of archives as a model for the conservation of certain kinds of cultural heritage materials: audiovisual materials, media art, and video games. Many types of institutions engage in preservation work with objects from these genres, even though

preservation is not the primary function of a library, museum, or other information institution. This is why the term *archive* and the profession of archivists will come up frequently: they serve here as referents to a particular kind of labor that may form a part of the mission or practices of many different professions and organizations. To be clear, perhaps, this paper more saliently refers to preservation and preservationists in whichever location they find themselves working. To be sure, few of the sites of inquiry call themselves archives; whether or not they are perhaps does not matter. What matters is the role they take in the creation and propagation of particular views of culture and history and how OAIS aids and allows this process, what constraints it may impose, or what subtler influencing factors it might entail. The question I pose here is whether it is possible to understand how application of the OAIS Reference Model affects practice, our understanding of our profession, and what impacts is has on the institution in which it is being applied.

Most of the literature and analysis about OAIS refers to practical implementation. What is missing is an understanding of how the model informs thinking about practice and how it shapes goals for preservation. OAIS is open in its design, and meant for wide use and adoption. Lee's (2005) excellent dissertation examines its origin story carefully, starting with the desire at CCSDS to create an overarching model to inform digital preservation and tracing OAIS through development. In his work, the contributions of outside organizations are made visible: through his chronology, he lists the different bodies and organizations that became involved in shaping OAIS. While this work shows the openness of the

development process[12], there are also ways in which it shows how much this model is a product of and for the space data systems community. This can be seen when looking at the major drivers of the original OAIS standard, as well as in the names that are involved in its on-going revision and the creation of related standards. What does it mean for practice that this tool comes from space science, a field that adheres to largely positivist epistemologies and finds funding in the military-industrial complex? Do it origins matter more or less when OAIS is used not for science data, but for cultural heritage like video games? The epistemological mismatch between positivist and post-structuralist theory and the tensions between quantitative and qualitative methodologies that occur in digital preservation are a mirror of these same tensions, which are manifested on a larger scale within meta-disciplines like LIS that span the traditional research disciplines spectrum. It is perhaps not unsurprising that the bulk of the research within preservation, and on OAIS more specifically, has focused on largely quantitative studies of implementation: this is reflective of a power imbalance in favor of the computer science (positivist) side of this field (Van House and Churchill, 2008). The ways in which this tension manifests itself in practice and what this means for the field of preservation is not well explored although it is often noted. This dissertation takes up this theme and makes a contribution by bringing together methodologies and epistemologies to investigate OAIS as an object and seeks to understand the role it plays in the intersection of disciplines and practices that make up

[12] There are indeed other ways to examine this: simple searches in SCOPUS or Web of Knowledge yield abundant literature on OAIS, and by limiting the date, articles and conference proceedings on OAIS appear in a variety of non-space science disciplines well prior to the first official publication of the standard. The role played by Dutch memory institutions including the Koninklikje Bibliotheek (KB), the royal library in The Hague, strongly influenced my decision to begin interviews in the Netherlands.

digital preservation precisely by investigating discourses and power in conjunction with OAIS primary documentation and interviews about implementation with practitioners. While its implementation in an institution must have social and cultural impacts on the institution and employees, as well as on general practices and politics of preservation, the literature thus far has not examined this. Taking this stock addresses current unknowns in the field of digital preservation such as whether and to what extent the application of the model informs and/or affects labor practices; enables or constrains the possibility to create self-reflexive archives; and enables or constrains the possibility to create less *imperial* forms of archives.

The end goal is something a bit more complicated: rather than a simplified OAIS that strips the model down to very basic components for use by everyday people on simple electronic objects (Spence, 2006), the aim is to understand how OAIS might be exploited to document and preserve digital objects of enormous complexity: think virtual worlds like World of Warcraft, entertainment franchises existing in many media over many years like Carmen Sandiego, or time-based media art like *Hole in Space*. While it may seem that a space science model is not likely to suit such materials, beggars (read: digital preservationists) cannot be choosers. In recognizing the ubiquity of a tool like OAIS, the question becomes: how to use this tool and how do users redefine this tool in some measure for their own purposes without having to go out and reinvent the wheel (read: another system that achieves similar goals to OAIS[13])? This research suggests even more

[13] This might possibly be a different huge, complicated, involved, and all-encompassing model. It could be that it means something else, perhaps an inclined plane rather than a wheel: one could imagine an alternative that is comparatively smaller, simpler, less involved yet sufficiently encompassing. The point is that even if such a utopic philosopher's stone existed or could be created by a team of plucky and tech-savvy preservationists,

specific questions based on cultural and geographic difference in interview responses: what are the European concerns as regards OAIS and how are these similar or different to those in the US and North America? In Europe, the focus seems to be on standardization and its particular relationship to auditing, specifically utilizing the newer ISO 16363 standard that is explicitly based on OAIS. In what ways does this geographic concern for standardization reflect regional concerns about politics of pan-European unity and homogenization of a Euro-identity created in part through governmental techniques enacted through large-scale cultural heritage projects?

This dissertation proceeds in six additional chapters. Chapter 2 examines the existing literature about the OAIS reference model. I utilize literature about archives in the age of the digital to understand what role OAIS could have in the changing or preserving professional practice in archives. I employ literature about sociotechnical systems to provide a way of speaking about the complicated network that is masked by the simplistic term OAIS and explore literature undergirding the theoretical frameworks I use in my analysis.

Chapter 3 details the methodologies that create this dissertation. I engaged in qualitative semi-structured interviews with a variety of digital preservation practitioners and scholars in Europe and North America, and I analyzed the data using content and discourse analysis.

---

development, implantation, adoption, and support take enormous amounts of time and other resources. This is true even for "free" or "open" programs. OAIS, while not trivial to implement or even possible to implement for all institutions, is widely available, well-established, and has ample support in the preservation field and marketplace.

Chapter 4 discusses the findings from these OAIS interviews. It examines the audiences and stakeholders implied by the OAIS model itself as well as those identified from interview data in order to examine the discourses that enable the construction of such subjects. I also investigate how the traditional role of archives is challenged by the current digital moment and the role that technologies like OAIS play in the evolution of how memory practitioners conceive of their work as institutions change. Finally, I explore the ways in which geographical location influences the role that OAIS plays in different institutions and investigate the politics of governmentality as they are manifested in the use of audit standards in European archives.

In Chapter 5 I take a detailed look at the OAIS term *Designated Community*. I examine how this term is approached by practitioners and how the authors of OAIS explain the genesis of this idea. This term takes on different meanings within difference spaces encompassed by the OAIS sociotechnical complex: in some places, it refers to an imaginary while in others it necessarily refers to real people. The tension here is manifested differently with different content types and I ultimately argue that Designated Communities are unworkable in certain circumstances and new modes of thinking are needed to address preservation intended for a particular audience.

In Chapter 6 I look at another particularly difficult term, this time one that purposefully does not appear in OAIS. Combining data from the Preserving Virtual Worlds II (PVWII) grant project with OAIS documentation and OAIS interviews, I examine the role of significant properties and how well these map onto existing entities within OAIS.

In Chapter 7, I conclude by exploring alternatives to OAIS and suggest new avenues of research to create metrics for bounding expanded context of digital objects and on the creation and maintenance of records documenting the Knowledge Base of users.

CHAPTER 2

LITERATURE REVIEW

Introduction

OAIS serves a site of tension for multiple audiences, and these tensions impact work in archives and other memory institutions. This project seeks an analysis of the social and political impacts using OAIS has on particular institutions, memory institutions more generally, and the evolving field of digital preservation. Early articles and theses documented the history and creation of OAIS; most existing and current OAIS-specific literature predominantly examines the implementation of the model. This encompasses actual implementation at various points; projects that examine how possible it is or would be to implement the model; samples or guidelines for creating the SIP, AIP, and DIP elements; and discussions of what implementation of a reference model actually means. Because I look at other issues in addition to implementation, it is necessary to look to other bodies of literature in order to tie together the various concerns at play in examining OAIS in reference to discourses in archives and preservation. Because OAIS is both an ISO standard and a *de facto* standard within preservation practices (Lee, 2005; Meghini, 2013), I continue to extend discussions about the politics of standards and standards adoption to OAIS. I argue that the evolving nature of archival and specifically digital preservation practices and sociotechnical analyses of systems are equally important to understand OAIS and preservation discourse. OAIS calls itself a very particular type of archive, but the disruption of the digital on the preservation work that happens in memory institutions is bound together with OAIS because of its dominance in informing the kind of work that institutions actually undertake when embarking on digital preservation projects of both

born digital and digitized materials. How archives conceive of themselves in this moment and where that intersects with the digital are spaces ripe for investigation: one focus of this work is precisely to understand and also to intervene in the ways in which memory institutions will 'settle' as they develop stable workflows to address digital content in addition to the work they may already be doing with analog content. Related to this is the notion of the sociotechnical system. I borrow from the language of this field to understand what OAIS is when people and institutions speak about it and use it. It is, formally, a reference model, but the various ways in which interview subjects speak about it or even in which authors of literature about OAIS refer to it suggest a variety of other ways to classify it: as a tool, as a framework, as a technology, and equally, as explicitly not any of those things. Instead, understanding OAIS within the context of a sociotechnical system, as part of a network both inter- and intra-institutionally, gets close to describing OAIS as it is referred to in daily vernacular among users. In combination, these different strains of literature provide the grounds for the theoretical and methodological frames I used to analyze the interview data collected for this project.

2.1 *Literature about OAIS*

The general literature about the OAIS reference model covers a period of almost two decades at the time of writing. Primary resources on the subject include the actual standards documentation. Because OAIS is an ISO standard, ISO documentation is available on the subject; however, given the cost barrier, the predominant primary source on the reference model and subsequent ingest-related standards comes in the form of CCSDS's documentation. The text is identical to the ISO documentation; however, CCSDS makes their books available for download free of charge and this availability may be one small

factor that aided in the circulation of this model. This preference for use is borne out in the interview data wherein people who spoke specifically about the model and its documentation referred to the 'color' of the book, or the different color-codes that CCSDS uses to denote the various stages standards move through[14]. The three versions – the Blue Book, the Pink Book, and the Magenta Book – each signify different developmental stages according to CCSDS's internal standards developments processes (CCSDS, 2002, 2009 & 2012). As required by ISO, the standard is reviewed and updated on a five-year cycle. The most recent revisions were made in 2012, resulting in the Magenta Book; salient changes between the 2012 revisions and the previous version are outlined in a blog post by Barbara Sierman, a research and development specialist at the Royal Library of the Netherlands (Koninklijke Bibliotheek or KB) and a member of the Primary Trustworthy Digital Repository Authorization Body (PTAB), a group responsible for creating and training people in ISO 16363 (Sierman, 2012).

Literature about OAIS falls into two broad categories. Firstly, there is general explanatory literature that explains the nature and genesis of the standard. Secondly, there is a large body of work on particular implementations of OAIS. A cornerstone of this is Lee's (2005) dissertation, which covers the chronology of the standard's development. His work documented stakeholders in the evolution of OAIS and their relationships to one another as well as listing early literature on OAIS. These stakeholders include the member institutions and individuals associated with CCSDS as well as others who were major and minor players

[14] See http://public.ccsds.org/about/images/FAQ%20-%20CCSDS%20Colors%20of%20Books.jpg for the summary of CCSDS' color scheme for documentations. Blue indicates a recommended standard; magenta indicates a recommended practice; and pink indicates draft revisions for review.

in the development of OAIS from drafts to standard: data specialists from a variety of scientific fields and representatives from memory organizations such as national or federal libraries and archives in Europe and North America.

Lee's dissertation is a purposeful attempt to understand how and why a non-LIS tool became so prominent in LIS spaces. Lee's work shows when and where the diverse partners involved in the creation of OAIS were enrolled. 'Diversity' in this case refers to the way in which CCSDS normally develops standards: CCSDS normally operates in a closed environment and does not commonly seek outside input. In writing OAIS, members of CCDSD solicited participation from representatives in all these areas. Lee's project traces a timeline of development wherein it is possible to see who was involved at various stages. The stages he describes range from early drafting to the end process of soliciting comments on well-formed and content-complete documents. A second takeaway from Lee's dissertation is that it demonstrates the overwhelming influence of certain bodies in authoring OAIS and the lesser involvement of others. While the development of OAIS was far more open than that of other CCSDS standards, the process was nonetheless dominated by CCSDS and researchers with space science backgrounds. For example, the US was the prominent geographical player in the creation of OAIS, although people and institutions from Europe were also involved and continue to be very heavily involved in updates to OAIS and the creation of related standards like ISO 16363 and ISO 16919, standards for auditing the trustworthiness of digital repositories. However, the standard is referred to as 'American' by some interview participants and this geographical designation is meant to note something of substance about why the model achieved such wide adoption outside space science.Libraries and cultural heritage institutions played a secondary role in the

historical account. I suggest that this is important in understanding the scripts embedded in OAIS and therefore the discourses that unfold when OAIS is deployed in various organizations. The dominant creators came from positivist, global north backgrounds. They also came from archives where the Consumer population is narrower than the populations served by libraries and cultural heritage institutions. This is important not because it means that OAIS cannot work in places like libraries, or organizations with broad and heterogeneous user bases, as some have suggested. The authors of OAIS explicitly designed it with the intention that it would be functional across a variety of disciplinary spaces, including popular culture and cultural heritage. Indeed, in my interactions with the creators of OAIS, they demonstrated an on-going eagerness to work with content and institutional types that are very different to their backgrounds in space science data. Rather, the disciplinary origins of OAIS are salient here because they inform the values the creators scripted into the model. This also informs what kinds of information counts as data within the model which in turn informs how to make decisions about selection and boundary-making for dynamic content; how labor and laborers are conceived of in relation to the model; and what kinds of knowledge and knowledge structures are valuable. This last part is highly situated and particular to locally dominant epistemologies.

As OAIS took off as a standard, a few prominent writers and scholars in the areas of digital preservation and general digital culture studies authored key summaries. These studies play a prominent role in the ways in which OAIS is communicated, circulated, and thus propagated. These are still primary sources for people seeking to understand OAIS, and in some cases are cited as more useful than the primary documentation itself[15], much

[15] Priscilla Caplan's work in this area was noted in interviews.

to the chagrin of OAIS creators and purists who believe that this is part of what leads to the misunderstandings that surround OAIS. This is also noted as a source of OAIS "bashing" that calls for its replacement or radical overhaul. The circulation of these summaries reintroduces, in some ways, the voices and thus the discourses of information and memory professionals as their works become canonical along with CCSDS's and ISO's official documentation of OAIS. Ball's (2006) primer falls under both of the categories of OAIS literature that I have previously described. First, it lays out the major tenets of OAIS. Second, it delves into non-site-specific particulars including listing related standards for metadata, modeling, and packaging as well as listing open source and custom tools and projects built on OAIS. From a meta-disciplinary perspective, it is important to recognize that when consuming and using OAIS, people do not necessarily read the primary documentation and do not necessarily understand it as its creators describe it. Because users turn to secondary sources that describe OAIS – its creation, purpose, contents, and implementations — the ways in which secondary articles are organized impacts the consumption of information about OAIS. It is necessary to look beyond just the manual to see how OAIS is actually taken up. Ball's article notes the information model, enumerates the functional entities in an OAIS repository, and speaks to preservation methods all before getting to the basic six requirements for OAIS compliance. These latter do not appear until section six of the paper. This may be telling: a common complaint of OAIS users in Europe was the fact that people do not read the entirety of the OAIS documentation carefully and that they tend to dwell on the information and functional modeling, particularly the diagrams. Without understanding the model as a whole, there are many common misconceptions about what is actually required for OAIS compliance, leading perhaps to

the two widely held views that compliance is either impossible or possible even for a headless chicken.

Noted digital preservation scholar Lavoie (2000 and 2004) makes the important move of explaining the acronym first and then immediately detailing the six mandatory requirements, as well as laying out the relationship between OAIS and PREMIS, a preservation metadata schema maintained by the Library of Congress[16]. Lavoie notes that 'compliance' is a necessarily vague notion when dealing with a reference model that avoids specificity by design; this notion of compliance came up in many of my interviews as one that particularly preoccupied European practitioners and one that US practitioners were quick to dismiss. It is, however, perhaps telling that they felt the need to dismiss the need to be compliant; it was important enough to mention even when I did not ask about it specifically. Lavoie says that compliance might look like a detailed mapping of existing activities into OAIS terms; it might also look like an organization faithfully implementing significant portions of the model. Written in 2004, the article looks forward to TRAC and ISO 16363, audits which determine trustworthiness through OAIS with much greater detail, although I would argue that achieving compliance with the main six repository requirements does not necessarily require the level of fidelity to the general OAIS documentation implied by Lavoie towards the end of his article. Schumann and Recker (2012) try to contextualize what the idea of compliance even means for something that is

[16] It should be noted, in a move that perhaps continues the trend in the US preservation community of beginning to move beyond OAIS, that the most recent changes to PREMIS include looking beyond the functions of the repository itself to outside actors like depositors. Angela Dappert (2015) describes the move of looking beyond the repository as a move away from the early ties to OAIS, which constrained PREMIS' earlier ability to account for actors outside of the preservation repository.

so purposefully high-level and abstract that it cannot match perfectly onto real world examples by definition, and this is important given that institutions try to comply with a model that claims not to make prescriptions. This means that such attempts must have an inherent locally-determined set of values attached, increasing the heterogeneous nature of the OAIS sociotechnical system. Daniels succinctly makes the point: "For all its clarity, however, the OAIS is a theoretical model that intentionally avoids any statement on how to go about creating a compliant digital archival repository." (Daniels, 2013, page 12).

The second major category of OAIS literature focuses on its deployments. This can be further broken down into two categories: one set of literature that deals with place- or site-specific implementations and a second that deals with general implementation guides or suggestions around a particular issue, term, or discipline. For site-specific work, interviewees, especially those in Europe, frequently mentioned Priscilla Caplan's work on the Dark Archive in the Sunshine State (DAITSS) as a particularly helpful and in-depth model (Caplan, 2004 and 2007). Kara van Malssen wrote a piece about OAIS commissioned by Beeld en Geluid, the national audiovisual archive of the Netherlands; the chief information officer of that same institution details their workflows and implementation of OAIS in a white paper (van Malssen, 2010; de Jong, Delaney, and Steinmeier 2013). Spence (2006) conducted research on the far-flung users of OAIS, noting how many different types of organizations were currently using it. Her analysis showed that it was even of interest to organizations such as gardening clubs in Australia, which demonstrates a remarkable breadth of scale in the types of repositories implementing it. The implication of this work is that OAIS is such a usable tool that it should be made more accessible to small organizations, and as such Spence proposed the creation of an "OAIS *simplex*", a simplified

version of the reference model designed specifically for small programs with severely limited resources. David Giaretta et al (2009) take on the notoriously nebulous term *significant properties*. In so doing, they dismiss the term as too unspecified, noting several sometimes contradictory definitions from various reputable sources in their literature review. Instead, they seek to explore which features already in existence in OAIS might serve the general ethos of preserving those properties deemed most important to the interpretation of an object or, more basically, the rendering of its bits. This latter marks an intentional distinction: as I will explore in Chapter 5, significance can be located outside the bits and the computing environment of an object, and so contextualizing how to render bits will not necessarily account for all the significant properties of a digital object.

Articles focusing on specific issues in relations to OAIS cover topics such as OAIS/FRBR mapping; OAIS and disaggregated, distributed preservation environments[17]; mapping PLANETS preservation planning within OAIS; OAIS as specifically applied to the social sciences; and OAIS' relationship to other standards (McDonough 2011a and 2011b; Giaretta et al 2009; Knight and Hedges, 2007; Spence, 2006; Subotic, Schuldt, and Rosenthaler, 2011; Becker et al, 2009; Ayoung and Tibbo, 2011; Vardigan and Whiteman, 2007). These articles cover a variety of different disciplines, including a significant chunk of

[17] This particular project, SHERPA DP, was like many of its kind completed and appears to have fallen off the radar. A check of the website yielded a long set of links that resulted in 404 error messages; a search of the University of Illinois' library catalog for this project yielded only one set of conference proceedings on the subject, from 2006 prior to the completion of the project (Wilson, 2006). This is only interesting because the notion of distributed compliance is something that OAIS will have to deal with and authors acknowledge this struggle. While this project looked at distributed dark archives for e-prints, the creators of OAIS admitted that linked data, a different form of distributed content, pose real challenges to the model and said, "Be sure not to be audited [on linked data]."

literature from library and information science as well as computer science, health sciences, cultural heritage studies, and a number of domain-specific explorations of OAIS and its relevance to preserving the data of that domain. This variety highlights the breadth of material and disciplinary interests covered in OAIS literature that is grouped under the more general heading of implementation literature: the existing body of literature is rich despite its rather singular focus.

What is missing from the literature about OAIS is precisely the conversation that this dissertation seeks to start: an examination of OAIS itself as a sociotechnical system, from a perspective that seeks to understand the ways in which OAIS functions in power relationships within institutions and more broadly within the realm of memory creation. This is a gap in understanding that is highlighted by the existing work done in related fields: the kind of work that has been done on organizations in STS literature and the body of literature that examines power dynamics and discourses within institutions like archives and standards adoption has not yet been extended to cover the realm of the digital archive or the work that memory institutions do when trying to preserve digital content. I aim to bring these lenses to OAIS, examining the object itself and the role it plays in institutions. The call for this type of research has certainly started to appear: Lindlar (2013) starts her article on OAIS by noting that the largest preservation issues are not technical ones, but largely about the network of stakeholders involved. Research about OAIS has only recently begun to focus beyond the system to an understanding of the larger impacts on practice and the effects on institutions. Schumman and Recker (2012) take on the notion of compliance in relation to OAIS in a way that is being very well received via Twitter. That "Digital preservation is not a pure technical task…" (Lindlar, 2013) is something that can

easily be forgotten in a body of literature about an object that stems from disciplines like space and data science, both of which have decidedly scientifically-informed positivist epistemological leanings and qualitative research preferences.

This dissertation extends the discussion begun in these types of articles. Both acknowledge that sociotechnical networks are at play when working in an OAIS-informed environment[18], one by highlighting the importance of context in determining levels of OAIS compliance and the other by engaging in a discussion of preservation policy that foregrounds the human stakeholders. I add to this discussion by examining the object itself and the role it plays in larger discourses of preservation and memory institutions.

2.2 *Archives and Digital Preservation*

Digital preservation and digital archives are a space of friction (Tsing, 2004). It brings together disciplines and bodies of practice with very long histories such as archives, conservation, and diplomatics. The long-standing traditions of these bodies of practice include strong codes of professional ethics and practice guidelines. Archives and diplomatics also have juridical roots: their work is bound up in the maintenance and production of evidentiary records. Digital preservation sits at the intersection of this branch of information work and computer science and systems design, disciplines that are somewhat newer and have different core values. These disciplines work together on projects of digital preservation, each informing and constraining the other, and there are considerable spaces of tension between them. One such example is precisely struggles with significant properties. They are difficult to describe and even more difficult to render into

[18] However, it is important to note that neither article uses the term *sociotechnical* or explicitly references that literature; the application of the term in this context is mine.

something machine-readable. A reviewer for the iPres conference noted in reference to OAIS and significant properties:

> The use of the term "information properties" in the 2012 version of OAIS instead of "significant properties" reflects an unresolved difference of opinion between domains, computer science and archival science, and there has not been substantial and productive discussion of the issues, though there has been some heat over the issue. Some authors in the computer science portion of the digital preservation community feel uncomfortable with what they perceive as ambiguity in the use of the term significant properties, though the term is informed by longstanding appraisal discussions and practice in the archival community.

The epistemological challenges are another aspect of the friction within digital preservation. As Van House and Churchill (2008) note, "This uneasy alliance of the unreflexively positivist, engineering-oriented computer science with post-structuralist social theory and qualitative methods is often problematic, especially since most of the power in the field resides with computer science."

The *digital archive* as a construct is even more problematic, because this term is so nebulous, or rather very generous in what it currently encompasses. Trevor Owens (forthcoming, 2016) nods to this and offers a list of the some of the things that might be included under this umbrella term:

- Collections of Aggregated Digitized Primary Sources
- Digitized Copies of Entire Archival Collections
- Born Digital Archival Collections
- Web Archives

- Collections of User Generated Born Digital Primary Sources

These are a set of activities more than anything else, or perhaps a description of the materials that an information worker handles, but certainly these are not institutional divisions: many or all of these activities might be carried out, to some degree, by a single institution. The question is: why use the term *archive* to refer to these various activities and material collections? It suggests something about the way in which the term *archive* is perceived. Colloquially, it gets used by popular graphical user interfaces (GUIs) such as email clients or blog sites and tools: archive refers to the collection of content you could down load from Gmail, for example, to save a local copy of your correspondence record or, more to the purpose of the tool that enables this, to copy content over to another email client. In blog spaces, archive refers to what is old, or at least older. These both suggest something about public imaginaries of the archive: they are bound up with discourses about collecting, saving, and age. An archive has a very specific meaning within its own sub-discipline; though archivists played a role in the creation of OAIS, OAIS uses this term to mean something distinct. OAIS contains the word *archival* in its acronym, the Open Archival Information System, and it means something very specific by the term:

> An OAIS is ***an archive, consisting of an organization, which may be part of a larger organization, of people and systems that has accepted the responsibility to preserve information and make it available for a Designated Community***. It meets a set of responsibilities as defined in the standard, and this allows an OAIS archive to be distinguished from other uses of the term "archive". (Giaretta, 2011 page 47)

This definition is specific to an organizational context, and speaks about the general activities involved in preserving content without specifying a content type or even a particular type of preservation activity, and so could be applied to any of the collecting practices listed by Owens—all of them could happen in archival organizations that use OAIS to model or describe their activities and products. Interestingly, this does not say anything about digital content specifically; indeed, the OAIS AIP model has an entity for physical objects. It would, however, be disingenuous to suggest that OAIS was not intentionally designed for digital content: the nod to physical objects is a recognition that physical objects provide structural and semantic support to digital objects. The role of particular hardware as part of a preservation package is a topic of on-going debate: there is not an entity for hardware in the archival information package, although the more recent acknowledgment of the role software plays in an information package may imply or contain references to specific hardware given the ways in which some software elements are necessarily tied to hardware elements.

What is salient about this discussion is what, if anything, the digital archive does differently for the field of memory work than paper and manuscript archives have done previously. Apart from the technological difficulties in managing digital content that is dynamic, unbounded, and interlinked, does the digital moment challenge the largely imperial traditions of archives (Bowker, 2006)? The question is not whether or not these compendia of projects and institutions that can be grouped under the aegis 'digital archive' have more democratic potential: instead, it is whether or not that democratic potential is realized and the methods by which interventions are made to reach this potential. OAIS informs the construction of many of the digital archive types above in the sense that it

undergirds the systems that allow for the preservation of otherwise very ephemeral content. Just as paper archives reflected the practices of those that "had the resources to make their discourses substantial through books and pictures, and these were the materials then put into libraries and archives" (Fairclough, 2006, 221), the issue is less about whether those who existed in archives previously as traces in the margins, marginalized and/or subaltern populations without the power or privilege to leave records of the sort that memory institutions collected, can create digital content and establish a digital presence. It is about whether these populations who are now perhaps better able to make user-generated collections of content online can get this content into a system that is OAIS-compliant so that it persists. I only use this latter term to denote systems that are generally considered appropriate and capable of maintaining material long-term, and long-term here merely refers to the ISO 16363 concept of long enough to safely pass materials on to a responsible and trustworthy successor institution. More content is being created by more people (Rawson, 2014 for example): will more content from those who are not powerful last, or will it be relegated to traces if it cannot find its way into institutions with the resources to engage in its active maintenance over time?

The way in which archives privilege the powerful is built into their very structures at the most granular level. Archives are built on records, and what precisely constitutes a record was historically relegated to documents of various types that were the province of the powerful: letters from the literate with the leisure time and other means to write letters, marriage records from those legally allowed to marry, etc. In order to effectively archive digital content, it is necessary to reconceive both the notion of the record as central and what constitutes a record, opening archival spaces to new, sometimes nebulous classes

of objects. Archives as a practical discipline has begun to move beyond the record as central object. Within the historical tradition of archives, archives consist of predefined collections of records from a particular organization. This admittedly reductive definition is important because it speaks to the functions and practices of archives for thousands of years. First and foremost, the record is central. Another example of the use value of this conception of archives is that it in some ways obviates the need to talk about selection and organization issues, frequent concerns in libraries, as the assumption is that all records from an organization will be collected and organized according to their use function within the institutions where they originated. The vast work of InterPARES speaks to definitions of digital records, covering classes of records that exist in digital realms, how these map onto analog equivalents, the ways in which they do not, and how they are unique to computing environments (Duranti and Thibodeau, 2006; Hackett, Underwood, Eppard, 2008; www.interpares.org). Digital content is so easy to create and so ubiquitously available that we cannot collect every record, a main tenet of traditional archival theory. While the cost of storage of vast quantities of digital data decreases fairly rapidly, the ability to store data does not guarantee access to, or even persistence of, such materials. The persistence of not only the content, but also the ability to view and interact with it, assumes a role of primacy with digital content. What this means is that simply having a record is no longer sufficient to guarantee authenticity; the record is decentered in favor of activities needed to reproduce content and meaning sustainably (Bradley, 2007).

Addressing value and meaning in the OAIS reference model is a purposeful attempt to understand what implications technologies, tools, or systems have for archives. The digital challenge to the primacy of the record is a complex shift. OAIS was created as a

framework to assist institutions in navigating these changes and answer some of the questions about the role of the record and the archive more generally. This type of discussion is important because it enables the critical examination of politics hidden under technologies and historical processes, such as the structures of power existent in traditional archival practice. OAIS has been a success in pervading a field that is a mashup of disciplines. This success speaks to the way in which it is a potentially catalytic part of a network of relatively stable and aligned interests among actors who do not otherwise share interests (Bowker and Star, 1994 and 1998). The sociotechnical lens and the borrowing of terminology about networks and scripts aid in the process of "…inverting our commonsense notion of infrastructure means taking what have often been seen as behind the scenes, boring, background processes to the real work of politics and knowledge production and bringing their contribution to the foreground" (Bowker and Star, 1998, page 234). There are reasons why OAIS has not yet adequately tackled the complex social and political aspects of preservation work like significant properties (Bradley, 2007) and these reasons are parallel to why work has not been done to investigate OAIS as an object itself. Rather than addressing a gap in the literature, this study renders visible processes that are often overlooked, a common occurrence in preservation and information work more generally, such that their politics and knowledge production discourses can be brought to the fore and examined for what they contribute, or could or should contribute, to the practical discipline.

There are ways in which digital preservation as a professional practice maps very well onto analog archival principles, and other ways in which it does not. Daniels (2013) provides one interpretive mapping of traditional archival principles onto OAIS terms from

the perspective of a cultural heritage institution, although like many of the previously cited OAIS papers, her work focuses mainly on a local interpretation and less on a deep, theoretical understanding of the ways in which OAIS and archival science do and do not overlap. I argue that trends in archival science, such as calls for more pluralistic forms (PACG, 2011) and towards sociohistorical understanding (Douglas, 2010) as opposed to constructs and assumptions of neutral history, need to be brought to bear on OAIS.

The push for pluralism in archives from a single cultural paradigm to a multiverse (PACG, 2011; Caswell, 2013) offers particular promise for some of the difficulties facing digital archives in general and those implementing OAIS in particular. A common complaint from librarians about the restrictiveness of the Designated Community entity in OAIS reflects the fact that digital preservation has been conceived of in ways fundamentally different from analog conservation. If analog conservation is about a particular object and what is done to it to keep it safe and sound—think acid free folders and climate controlled rooms—digital preservation is less about the object itself as well as being a more active endeavor. It is not about where the item is put and left, but more about the ongoing activities around an object to ensure access to its content for a particular set of projected users: digital preservation is fundamentally aimed at a particular population (the Designated Community). It does not happen in a user-less vacuum. While analog conservation rarely happens in such a vacuum, the major difference here is that it can. Recently, OAIS creators working for the APARSEN project have begun to acknowledge and codify the value of preserving for multiple communities (APARSEN, 2015).

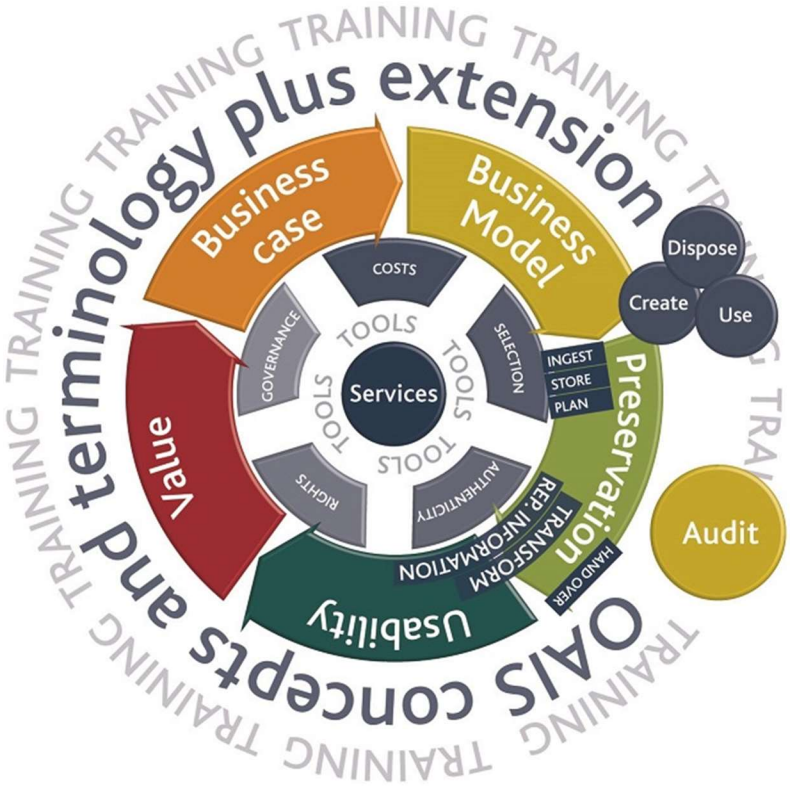

Figure 4: APARSEN Common Vision Model

Expanding access to content to multiple communities is part of a business model based on a digital object lifecycle where preservation is based on OAIS. The pitch for curation of data to serve multiple user communities is a financial one that happens after primary preservation activities for a Designated Community have occurred. I argue that digital preservation needs to look at multiple audiences earlier in the process and for reasons beyond the admittedly important ones of establishing financial stability in the long run.

Caswell (2013) describes the basic principles of archival pluralism as being informed by the four basic principles of religious pluralism: “energetic engagement, understanding, strengthened commitment, and dialog” (page 274). Of equal importance in this call is that archival pluralism simultaneously seeks to avoid four major perils: “claims

of universality, inattention to power, silencing dissent and collapsing difference". This latter lens can be applied to archives generally, as Caswell has done. It can also be applied to a critical reading of digital preservation and of OAIS itself. This is in keeping with recent trends in archival practice that challenge traditionally assumed Jenkinsonian neutrality of archivists and archives (Nesmith 2002; Cook 2001; Ketelaar 2001). As laudably open as the OAIS process was, end users of the content do not represent a significant part of the design process. The audience for the open design process was a collection of users of the model itself, and so there are many examples of communities that could be considered marginal in relation to the financial and reputational power of data science present in negotiations and renegotiations of OAIS as it moves through drafts and through periodic revisions. Next generations of OAIS revisions, or perhaps next generations of digital repository standards, should consider real end users who represent Designated Community members, in much the same way museums and archives have begun to solicit participation from both audiences as well as from marginalized groups who contribute content (Caswell, 2013; Boast, 2010; Schwartz and Cook, 2002). There is a possibility for OAIS or models like it to act as recuperative technologies (Coombe and Wiess, 2015), to provide spaces wherein formal organizations of record and memory like archives include voices that have been traditionally overlooked; this is not built into the current model given its origins and design process. This dissertation highlights the spaces where this is a possibility, to inform and argue for the need to include these voices in future iterations. Such recognition requires a more qualitative understanding of the design practice.

2.3 Standards and reference models

It is important to view OAIS through a sociotechnical lens as a complement to viewing it through the lens of standards development/adoption and the role of reference models. Reference models are generally high-level, aimed at shaping future practice and discussions: subsequent standards about actual practice will be built around them. They are not guidelines to practical implementation, and therefore 'compliance' can be difficult to determine: it is hard to comply with something that denies that it is prescriptive, and perhaps impossible according to this definition of reference models. One question that arises is who gets to decide what is minimally critical in the realm of compliance? How are disputes about this mediated? OAIS is in use in a number of institutions: how do these organizations deal with concerns that still exist around the concepts of Designated Community or significant properties, for example? Herbst's (1974) concept speaks to the ethos behind the creation of OAIS- it seeks to reduce barriers and allow for seamless and continual work within institutions that were/are already coping with the realities of digital preservation. Rather than prescribe adequate systems or methods, OAIS sought instead to codify activities, to give a shared language and thus prescribe the boundaries of good preservation work for all the denial about prescriptions within OAIS.

It is, importantly, a standard and a particular kind of standard, a reference model as opposed to other kinds of standards that dictate more specific rules and guidelines, like ISO standards on food safety temperatures or even trustworthiness audits like the 9000 series, or ISO 16363 and ISO 16919, which are about digital repositories. Previous theses about OAIS, prominently those by Lee (2005) and Seles (2016), delve into standards and standards adoption literature since both theses are concerned with adoption, albeit in

different ways and settings. In some ways, an investigation into OAIS is inherently an investigation that looks at a successful instance of standards adoption. I call OAIS successful given the overwhelming reference to OAIS by practitioners in a large number of developed world memory institutions; Seles labels it less so in an investigation with both developed and developing world institutions wherein participants from both types of region listed their difficulties in implementing digital preservation programs that match well with the ISO 16363 audit standard. OAIS is a standard. It is a de facto standard by way of its adoption even prior to it becoming an official standard. It is also a de jure standard in the sense that it gained ISO status and that certain governmental agencies began to require audit processes and certification for state-mandated and -operated digital repositories.

Lee's historical investigation into the creation of OAIS examines literature regarding the creation of standards in order to note the varying motivations different actors have in participating in such processes; this discussion helps ground his examination of the way in which the creation process unfolded, particularly how different agents representing different interest groups and disciplines came to be involved. This is the main impetus of the project: to understand how this object, not created within the realm of LIS, came to be so widely discussed in that discipline in the early 2000s. OAIS is also a particular kind of standard: it is a model, and more particularly, a reference model. Lee identifies 22 characteristics of reference models, and the single major consensus in the literature about standards, as summarized by Lee, is that reference models in particular exist at a higher level of abstraction than other standards models and that they are useful tools in the structuration processes within a field. OAIS came into being at such a time that it was able to play a large role in the structuration of digital preservation as a sub-discipline by

directing the development of the field and by giving a language to the type of work that people within the field were doing. In this way, this particular standard is tightly imbricated with the entire field a way that other standards, such as ISO 15489, an Australian-based standard for record management, are not. The role that Lee describes for OAIS as a standard is a role that I still find it to play. In interview data, participants noted the fact that OAIS is a highly used ISO standard as part of its allure to their organization, and European participants did indeed use it to describe and in some ways legitimate their work.

Lee's dissertation forms the germinal cornerstone of in-depth research about OAIS. Seles references it heavily; my interview participants mention it as informing their practice, particularly those in Europe; and it was the starting point for my own research on this topic. My dissertation, like Seles' and Lee's, takes as a basic assumption or hypothesis that OAIS is packed full of biases (Lee), and what I would also call scripts (Akrich, 1992) and discourses. Fundamentally, standards are social constructions that need to be understood within a particular milieu. Importantly, they have the power to both anticipate change and sow "seeds" to enable change to occur, while simultaneously having the power to reinforce existing social structures (Lee). This recognition is important because it explicitly rules out a techno-deterministic reading of OAIS wherein it has the inherent power to do one thing versus another or where it is a neutral object that does not carry the weight of human intentions shaped by social circumstances. Seles notes that means that existing digital preservation standards have a developed-world perspective that may impede their wide adoption in the global south, and this outcome is yet another manifestation of the way in which I argue that OAIS serves to replicate power structures that have existed in archives

for millennia. Seles refers to ISO processes for standard making more specifically, noting the following four criteria that drive the creation of a new standard:

(1) Standards must answer a market need.

(2) Standards must be developed by a team of international experts…

(3) who must be drawn from various industries and sectors to ensure a balanced representation of interests.

(4) Standards must be based on consensus, and the comments and opinions of all stakeholders must be given equal weight and consideration.

(ISO in Seles, pages 51-52)

Seles' overarching critique of OAIS (ISO 14721) and RAC (ISO 16363) is that the representatives from a variety of agencies involved in creation were not representative enough: the developing world was poorly represented (or not represented at all), and this resulted in standards that contained assumptions that render it non-transferable to developing world situations. Such assumptions include, according to Seles, the presumed access to stable infrastructure for electric power and Internet and access to technologies such as computers or HVAC equipment. Her findings, based on research in East Africa, are that these assumptions make OAIS virtually impossible to implement as it is currently written in that geographical context, and her recommendations include having more representative groups involved in future standards creation and revisions. This last is important because ISO standards are required to be revised on an on-going set timetable. While Seles explores the various reasons why East Africa was not included in the development of the dominant digital preservation standards, she notes that issues related to technology access that are endemic to the area do not mean that there is no need for

digital preservation and digital preservation standards in the developing world. The exclusion of East African and other developing world areas can be partly explained by the same assumptions which render OAIS un-useful: OAIS creators interviewed by Seles thought that "anyone" could participate in open meetings regarding standards creation because all that was required to participate was an internet connection. I would argue that part of this disconnect is explained in the very wording of ISO's key criteria for standards development. It is telling that the first criterion is a "market need" for a standard. ISO's website says:

> ISO does not decide when to develop a new standard, but responds to a request from industry or other stakeholders such as consumer groups. Typically, an industry sector or group communicates the need for a standard to its national member who then contacts ISO. Contact details for national members can be found in the list of members.

ISO lists its members, which are predominantly professional organizations and industry specific groups in the developed world. The "market need" can be read in relationship to the members of this group: in developing a standard, there must be a demonstrable market need according to the markets of the ISO members. ISO standards are not universal. Perhaps they are not even global: they address the needs of predominantly developed countries. So while it is true that East African nations have a demonstrated need for digital preservation of electronic records (the primary subject of Seles' work), it is not a market need according to ISO or its members. Seles' own demonstration of the general lack of digital preservation professionals in East Africa in some ways points to a "market" that is not yet developed enough to express a market's needs, or was not when OAIS in particular

was developed. This type of language firmly situates the projects of standards bodies within the realm of capital, and thus it is unsurprising that standards are developed to be pitched at those with the capital to purchase them, and not just in the sense of being able to afford the fee associated with purchasing the actual documentation for a standard. Instead, standards are a form of commodity (Schiller, 2006), and those without purchasing power do not have a say in their construction. The move of capital to exclude the margins and continually centralize power is not unintentional; exclusion from standards excludes marginal actors from the very language and infrastructure that would allow participation in the market. ISO standards like OAIS do not transfer well to developing nations because they are not meant to, and I argue that they will not be designed to accommodate the needs of non-developed world actors until the electronic records of East Africa and other developing world regions fill a market need for the developed world. In this way, standards seek to homogenize processes and facilitate exchange and transfer of goods and knowledge in ways that privilege already privileged classes while purposefully excluding the margins. These discourses are scripted into OAIS because OAIS was developed within this framework of understanding about the role of standards. Making standards functional in different ways, for non-dominant populations or projects, requires a fairly radical rethinking of such projects, and this dissertation seeks to lay the groundwork necessary to do just that.

2.4 Sociotechnical Systems

The focus here is to look at OAIS through a sociotechnical lens. One question I posed to interviewees is how they would characterize "OAIS." This is a very particular request in that it elicits not only people's automatic responses based on the common discourses

around OAIS with which they are familiar, but also because it often elicits personal analysis on the part of the interview subject about how they classify OAIS. Many people say it is a standard and/or a reference model, because on paper it is. In practice, this may also be the way most people perceive OAIS. It is certainly the way they most often write about it. However, there are other ways of categorizing it. Characterizing OAIS as a technology in the critical theory sense opens the possibility of examining the ethics and politics of design as well as possibilities of counter-hegemonic potential when technologies are used in ways counter to their original intention to empower oppressed classes of actors. Mirroring Trist's (1981) findings about the organization of workers and technologies in mines, the deployment of OAIS in institutions similarly follows many of the principles he identifies. OAIS foregrounds functions over particular jobs or individuals; it presumes group work processes in the sense that it divides up responsibilities among different entities that comprise a whole functional model[19]. The overarching processes take primacy over particular workers and their responsibilities because it is a reference model that looks globally across an institution to examine the confluence of processes necessary to make digital preservation function. In another sense, the primary six requirements for OAIS compliance are often taken up by what might be termed a primary work system in STS parlance: a group of people and digital architectures that comprise the functional group

[19] However, it is important to note that this group of entities may actually be filled by very few individuals. Particularly in very small organizations where all technological infrastructure is handled by one person, the vast majority of tasks might be filled by the same persons. The full model speaks to functions outside the preservation tasks themselves, including administrative and funding tasks, meaning even in very small institutions, it's likely that there are at least a few people involved. It is conceivable, however unlikely, that someone creating an OAIS-informed system of preservation for personal materials could fill all roles on the expanded diagram in Chapter 1 by themselves.

responsible for digital preservation. This is particularly true in larger organizations wherein preservation is not the primary focus: think, for example, of the department involved in digital preservation at a library within a larger college or university. In organizations for which the primary responsibility is preservation, archives in particular, the whole of the OAIS functional model might still comprise only a work system[20]. Of the institutions I visited, there was still a marked divide between the entities responsible for analog conservation of paper, film, and art in several of them. In others, however, there were moves towards integration. This includes enrolling existing employees in digital workflows and integrating new employees building a new sociotechnical network by "disentangling (and possibly disassembling) humans and non-humans from old sets of ties within a large sociotechnical network and reframing…" them within a new space that not only has to deal with digital concerns, but which is also prepared to foreground them (Callon, 1999 in Kaghan and Bowker, 2001, page 263). In effect, organizations have to create and manage boundary conditions to give employees not directly involved with digital work a sense of autonomy. In such cases, OAIS is a macro-organizing tool that touches on many elements within a broader work system. It is not just a reference model, but also an actor of sorts within such institutions.

Part of the process for the manager, and for OAIS in general, is the construction and constant reconstruction of networks that keep OAIS central and essential. Law (1992) defines the concept of durability within a sociotechnical system:

The first has to do with the fact that some materials are more durable than others

[20] "The *work system*, which comprised a set of activities that made up a functioning whole, now became the basic unit rather than the single jobs into which it was decomposable" (Trist, 1981, page 9).

> and so maintain their relational patterns for longer. Imagine a continuum. Thoughts are cheap but they don't last long, and speech lasts very little longer. But when we start to perform relations – and in particular when we embody them in inanimate materials such as texts and buildings – they may last longer. (page 387)

OAIS, taken as a representation of a network of stakeholders, constructs durability for itself. That is, it maintains its position of power in systems by constantly embodying itself in materials and practices, by perpetuating its role through the creation of new tools and standards: it is always part of the digital preservation discourse, even when it is an unnamed background assumption, decentered and invisible yet still present. Part of the way in which OAIS constructed its durability was fortuitous, which is not to say deterministic: numerous interviewees related that it came about at just the right moment to gain sweeping adoption. Many institutions were in an episode of having to cope with digital materials but not yet having formalized systems for preservation and long-term access when OAIS first began development in the 1990s. OAIS was widely adopted because so many institutions needed a framework like the one it provided and there were not major alternatives. It grew with the field of digital preservation, such that it is an obligatory part of the field. It has maintained its position as an obligatory passage point (Latour, 1987) through the creation of additional standards and tools that bring it back to the forefront of the digital preservation discourses when it might otherwise have faded away. Examples of this include the creation of other standards based on it, like PAIMAS and PREMIS; tools designed with it in mind like DSpace and Archivematica; and most recently, the creation of ISO 16363, an ISO standard for auditing trustworthiness built on OAIS. With durability, effects change and things become located in new networks of relations. This engages

materials in networks more thoroughly and makes them more stable over time, making it all the more difficult to alter the ways in which the network constructs power: durability itself is relational, not inherent. As OAIS solidifies its presence as central to the digital preservation discourse, it becomes more difficult to conceive of the alternatives that the successful punctualization (Law, 1992), or representation of OAIS as a single, static entity, masks. The question of alternatives is an important one, and one that I asked interviewees about. No one presented a concrete answer to this question, other than the handful of respondents who said simply that there was (and/or is) no alternative to OAIS.

While the admitted project of this dissertation was to understand more about OAIS so as to put it to more democratic and imaginative uses than it has been, thinking about the alternatives calls to mind two particular theorists, Akrich (1992) and Drucker (2013). My work with OAIS springs from a desire not to waste already limited resources reinventing the wheel: in seeking the perfect model for reflexive digital archives, I thought it best to begin with the models in existence and see if they could be made to work for my longer-term research interests in the preservation of dynamic digital popular culture. Channeling Akrich (1992), my advisor carried this metaphor forward and suggested that the problem might not be the wheel at all: perhaps the real issue here is that what is needed is an inclined plane. In other words, how can the digital preservation community consider alternatives that perhaps require something radically different rather than a rebuilding or recasting of what is already existent? Akrich's work on the deployment of technologies touches on this in speaking to the ways in which what is actually needed by the group who will be primarily working with the technology is not always what is created by the designers. She talks about this through the language of scripts, examining what is

inscripted in new technologies and how these scripts play out or fail to do so upon deployment. Akrich's work comes from the sociological arena of actor-network theorists, and it is easy to cast OAIS' deployment in the language of this school of thought. In fact, OAIS might be something of a black box—something that is quite present, used often, but whose inner workings as a sociotechnical system are poorly understood or largely ignored. Black boxes are a process of closing off parts of a sociotechnical network, but this process is always a negotiation and never complete, like Law's concept of durability (Kaghan and Bowker, 2001). This type of on-going negotiation is typified in the continual redeployment of the OAIS standards through new standards and industry tools. For example, the ISO 16363 standard for auditing repositories based on trustworthiness is built explicitly on OAIS; so too is the popular, off-the-shelf tool Archivematica. The original creators of OAIS have explicit interests in these, and in the case of ISO 16363, authors can be seen working on both projects. However, in a bid to be cautious about the over-use and deterministic employment of the term *black box* (Hamilton, 2015), perhaps another term one might use from this same body of literature is punctualization. That is, OAIS masks itself as a single entity, a non-human object, simply a reference model, when in fact it is a heterogeneous network comprised of a number of actors which can be human, organizational, and technological. To present itself this way is a purposeful strategy to increase its own power in relationship to other actors, other institutions, objects, architectures, peoples, and schools of thought about the general subject of preservation. In its resistive struggle to maintain a stable position, strategic moves like the creation of attendant standards for ingest and audits or partnerships with (relatively) easy-to-use open-source asset management tools like Archivematica introduce OAIS into new networks while helping to

solidify its position and maintain a considerable amount of power as an utterly obligatory passage point in the shifting network of the world of preservation, and the other macrosocial networks of which this is a part. Like Foucauldian discourses, these moves both recreate the object and help to propagate it. These moves happen in both temporal and spatial ways. OAIS came to prominence at a time when many organizations were seeking just such a system for purposes of regulation, standardization, and control. Its emergence at this particular point in time makes it hard to have conversations about digital preservation where it does not creep in: it constitutes the very lingua franca of the field. Its spatial mobility can be seen in its creation at the global center—the US and Washington, DC specifically, and its subsequent spread to global peripheries, laid out in the literature in the numerous case studies about its deployment in Asia, Eastern Europe, Australia, Latin America, and Africa. As is the case when working within a very pervasive system, it is hard to envision alternatives; in this case, an alterative to OAIS would mark a very real paradigm shift. What would alternatives look like? Could they be tools to eschew some of the need for control in ways that encourage more open and transparent processes? This is a fairly radical consideration to make, because the inclined plane, as opposed to the wheel, might require practitioners to reconsider the commonly held standard goals of preservation in favor of less straightforward understandings of what needs to be preserved and how. Rather than making a new standard that advocates control and normalization, the shift might be in changes to expectations in the direction of something purposefully messier and seamful. This calls to mind the work of Drucker (2013), who specifically notes that the difficulty in changing how we think about interface design is that the very dominant vocabulary comes from the engineering community; in much the same way, we can see

how the dominant vocabulary in preservation given to us by OAIS and its creators at CCSDS make it difficult to understand what a humanist version of preservation might look like. It is perhaps interesting to note that the interviewee who had to most to say on this subject decried the lack of alternatives and had many possible preservation futures to suggest was an engineer, not a humanist. Drucker's work on performative materiality suggests a limitation to OAIS that has yet to be explored: how to document and thus preserve performativity. While there is a solid body of literature about performance preservation[21], preserving the performativity of digital materials (or interfaces) is a challenge, something that I would argue requires the documentation of data that archives in general are not accustomed to collecting. It is not something that OAIS treats directly, and indeed is something that I am convinced is not adequately dealt with by the alternatives to significant properties proposed by some of the OAIS creators (Giaretta et al 2009).

It is necessary to keep the humanist focus at the forefront of this investigation. Theories like actor network theory really focus on scientific communities (Bowker and Star, 1998). While I do not make the claims that digital preservation is a science, this type of discourse is popular in the general meta-disciplines where I locate the sub-field digital preservation: library science, information science, and archival science. Even if this is not language that is commonly used, "preservation science" being uncommon, I suggest that this is a background assumption in the sub-field. As is the case with digital preservation

[21] In fact, this is something the creators of OAIS are quite keen on. In literature and in interviews, they showed an eagerness to engage with interactive dance/music/art performances, such as dances that had digital technology elements and accompanying music elements. There is a ready admission to not fully understanding such work on the part of the authors: this serves their point even more, which is to demonstrate the flexibility or transferability of OAIS to contexts as far flung from science data as possible.

more broadly, OAIS itself comes from scientists: data and natural scientists as well as library, information, and archival science specialists. This assumption can be seen in the appeals to notions of neutrality and naturalness inherent in discourse around what makes a digital repository or archive trustworthy. Digital preservation seeks to place itself, as a discipline or set of practices, outside of discourse, as demonstrated by a reviewer who stated that digital curation audiences would be interested in the institutional findings of a project about OAIS and not critical studies that explore Foucauldian discourse and constructions of power. This explains, in part, the quantitative bent of the existing OAIS research.

2.5 Theoretical Frames of Analysis

Under the aegis of locating discourses, I asked a number of questions about my data throughout the coding process and after I had coded and grouped nodes and responses. Some of these questions were mentioned previously in Chapter 1:

- What about OAIS incites variance of use and approach, beyond the changes that obviously come from the variety of institutional settings?
- What does this say about the practice of digital preservation?
- And, finally, what do these struggles indicate about what practitioners are trying to do when they have to "do" digital preservation?

Additionally, I borrow from the work of Dean (1999, page 32) in asking of my data:

- What forms of person, self, and identity are presupposed by different practices and what sorts of transformations do these practices seek?
- What statuses, capacities, attributes, and orientations are assumed of those who exercise authority and those who are to be governed?

- What forms of conduct are expected of them? What duties and rights do they have? How are these capacities and attributes to be fostered? How are these duties enforced and rights ensured? How are certain aspects of conduct problematized? How are they then to be reformed?
- How are certain individuals and populations made to identify with certain groups, to become virtuous and active citizens, and so on?

These questions are aimed at eliciting the discourses that exist within the realm of digital preservation work, and how these vary among people with different statuses and projects. Through these questions, I aim to locate and investigate the role of existing discourses that come from the various fields of expertise that combine to create a digital preservation profession; I also look for new discourses that arise when traditional archival and memory work go digital.

The type of discourse analysis I use is inspired by the work of Michel Foucault and focuses on how power works and how it does what it does (Barrett, 1991). This type of discourse analysis examines the material practices of institutions via the collected verbal and written texts (Rose, 2011). I look to find discourses that are not formed by individuals, but are rather socially constructed by interactions with individuals, institutions, and technologies. "Socially constituted forms of discursive power include constructions of difference and authority, blame, accountability, how it categorizes particularities" (Fairclough, 2010), and I investigate the organizing impulses behind OAIS, the politics of particular institutional deployments, and the role of authority it invests through this lens. This type of discourse analysis focuses on the production processes and uses of objects like

OAIS, and thus this method acts as a complement to the investigation of scripts that focuses on the object itself via an understanding of the processes and uses.

My data collection approach, multi-site interviews, provides a broad survey across institution types and geographical locations. Because preservation work is informed by so many different disciplines and areas of professional practice, I examine the role of existing discourses of archives and the interplay of these with the introduction of digital technologies in memory institutions with a focus on the work that OAIS does in this space. Discourse is "a coherent pattern of statements across a range of archives and sites" (Green, 1990). Part of the power discourses have in addition to rhetorical strategies and claims to truth comes from their institutional location (Fairclough, 2010). This might refer to the institutional location of OAIS within CCSDS and ISO; to its use in particular institutions; and further to particular locations or segments within institutions. This allows me to locate trends that enable the construction of an understanding of discourse as it is constructed around digital preservation as a practice or meta-sub-discipline. The broad survey of locations and institutions helps paint a better picture of the discourses as manifested within a number of Archives.

In the Introduction, I noted that I often refer to OAIS as a technology. It is from the Foucauldian analytic framework that I borrow the notion of technology to refer to the OAIS reference model. While OAIS is a reference model and therefore not a concrete piece of technology, as is pointed out by users and authors, I refer to it using this term throughout the dissertation. This term refers to particular use where technologies refer to techniques of power/knowledge that are "diffuse, rarely formulated in systematic discourse, made of bits and pieces" (Foucault in Rose, 2011). This echoes the piecemeal construction of digital

preservation as a profession and discipline as well as the heterogeneous creation and use of OAIS more specifically. Further, it centers the idea that memory institutions especially provide a setting for "works of culture [that] also functions as a technological environment which allowed cultural artefacts to be refashioned in ways that would facilitate their deployment for new purposes as part of governmental programmes aimed at reshaping general norms of social behavior" (Bennet, 1995, page 6). This ties in part to Foucault's later work on governmentality and the ways in which techniques of government can be enacted through cultural programs; this lens become useful in understanding the difference between the adoption of OAIS in Europe in contrast with the United States. I argue that politics and programs to foster a centralized pan-European identity that encompasses the heterogeneity of local populations plays into the literal adoption of OAIS and the push for de jure audit practices.

3.4.2.2 Akrich and Scripts

The use of discourse analysis enables the exploration of the material institutional politics of OAIS' deployment in a variety of settings. I use the work of Madeleine Akrich to examine what values are contained with OAIS itself so that I can examine the role these play in these institutional politics. Inscription is the process by which the authors of OAIS, for example, imbued their technology with assumptions and values. Its subsequent spread and adoption in heterogeneous locations provide spaces of tension that expose the seams in the object itself, rendering these scripts visible and allowing me as a scholar to de-scribe what has been packed into OAIS. Akrich and Latour define scripts and the various modes in which they are created, read, and otherwise discerned as follows:

> ***Script, description, inscription, or transcription:*** The aim of the academic written analysis of a setting is to put on paper the text of what the various actors in the setting are doing to one another; the de-scription, usually by the analyst, is the opposite movement of the in-scription by the engineer, inventor, manufacturer, or designer (or scribe, or scripter to use Barthes's neologism); for instance, the heavy keys of hotels are de-scribed by the following text DO NOT FORGET TO BRING THE KEYS BACK TO THE FRONT DESK, the in-scription being: TRANSLATE the message above by HEAVY WEIGHTS ATTACHED TO KEYS TO FORCE CLIENTS TO BE REMINDED TO BRING BACK THE KEYS TO THE FRONT DESK. The de-scription is possible only if some extraordinary event – a crisis—modifies the direction of the translation from things back to words and allows the analyst to trace the movement from words to things. These events are usually the following: the exotic or the pedagogic positions (we are faced with a new or foreign setup); the breakdown situation (there is a failure that reveals the inner working of the setup); the historical situation (either reconstructed by the analyst through archives, observed in real time by the sociologist, or imagined through a thought experiment by the philosopher); and finally the deliberative experimental breaching (either at the individual or collective level). No description of a setting is possible or even thinkable without the mediation of a trial; without a trial and a crisis we cannot even decide if there is a setting or not and still less how many parts it contains (Akrich and Latour, 1992, pages 259-260).

I find this method of analysis to be an appropriate complement to Foucauldian discourse analysis and generally applicable to this research for a number of reasons. Discourse

analysis aims to trace exchanges of power. Within the realm of memory in particular, power is what enables allocation of resources and values: quite simply, what is preserved is what is valued by those in power. Akrich's language of scripts is precisely a method by which analysts and academic researchers examine what is in-scribed into a new technology; in other words, what values are written into OAIS by its authors:

> ...when technologists define the characteristics of their objects, they necessarily make hypotheses about the entities that make up the world into which the object is to be inserted. Designers thus define actors with specific tastes, competences, motives, aspirations, political prejudices, and the rest, and they assume that morality, technology, science, and economy will evolve in particular ways. A large part of the work of innovators is "*inscribing*" this vision of (or prediction about) the work in the technical content of the new object. (Akrich, 1992, pages 207-208)

Akrich later names part of this process the I-methodology, to indicate that designers often base their assumptions about values, capabilities, status and general subjectivities of end users on themselves (Akrich, 1995)[22]. They constitute a Designated Community for a new technology that is filled with variants of themselves, to use language I will discuss in depth later. As noted in the definition of scripts, a de-scription exercise can only follow a trial and a crisis: that is to say, I as an academic am only able to look for scripts in OAIS if there is a mismatch at some juncture that renders an object seamful (Sherratt, 2015). If no such crisis occurs, then the technology remains seamless and such analysis cannot take place. The

[22] To be clear, as this is often misunderstood, the I-methodology is not a method that designers purposefully apply nor is it something Akrich advocates for. It is to describe an often unconscious practice that explains the frequent mismatch between development and reception of 'technologies'.

ability to look for scripts presupposes the tension that I hypothesize as the crux of this dissertation: as stated earlier, mainly that OAIS is a source of tension in the spaces where it is adopted locally and generally within digital preservation as a growing profession and meta-sub-discipline. It is this crisis that results from the multiple readings of OAIS as an object and technology that results in its heterogeneous adoption and the loss of the univocal authority on the part of the authors of OAIS. Per Derrida, these multiple readings allow for deconstruction which "…counters… the authority of the author by focusing on the material aspect of signs inscribed on pages… such inscription leaves language open to multiple meaning, that spacing of traces differs and displaces meaning away from the author, that the linear form of the book… are open to close reading…works by exclusions, supplements, and marginalizations which may be reintroduced in a subversive reading… Deconstruction attempts to destabilize the march of univocal meaning in written texts by unlocking the logic of difference that it hides." (Poster, 2004, page 406). These are precisely the scripts I seek to find in the ways in which practitioners and scholars use and characterize OAIS.

2.6 Summary

In this chapter, I have reviewed literature about OAIS; literature about archives in the era of the digital; and the relation of literature about sociotechnical systems to the project of this dissertation. The primary value in these bodies of literature is their intersection. This work takes OAIS as a punctualization of a network, and this concept is important to avoid statements that seem to ascribe thought processes and sentience to a document while acknowledging that there is intent behind OAIS and that it yields a considerable amount of power even as a paper or digital document. Recent work in the

field of archives to address long-standing politics of exclusion ground an understanding of the work that OAIS does, as it both supports existing archival discourses while challenging others.

Literature about OAIS reveals the pervasiveness of the standard and its importance for the growth of digital preservation as a sub-discipline of information work. It simultaneously demonstrates the need for research applying a critical lens to its mass adoption in ways that have already happened for other standards. It is within this context that I analyze the interview data to understand what OAIS means to the practitioners and scholars I interviewed for this project.

CHAPTER 3

METHODS

Introduction

In this chapter, I describe my research approach and justify its appropriateness in reference to the research questions and my own epistemological leanings. I offer three justifications of my process in the purpose statement. Firstly, I explain what can be gained by semi-structured interviews for answering the research questions. Secondly, in light of the literature in the previous chapter, I speak about why I chose this line of inquiry. Thirdly, I offer a justification for my focus on practitioners in cultural heritage.

I also describe the data collection process as well as information about my interview subjects and the places where I encountered them. Next, I outline my analytic methodology and the theoretical frameworks I used to understand my findings. I used content analysis to examine the trends within the codes I applied to the interview transcripts and additional documents. Two complementary theoretical frameworks enable the exploration of two aspects of the data in relation to my research questions: I employ Foucauldian discourse analysis to examine institutional and professional politics and exchanges of power within the realm of digital memory preservation; and I use Akrich's language of scripts to understand the values written into OAIS and how they play out in these institutional contexts when OAIS gets unpacked. Finally, I discuss some limitations to both the methods and the project more generally.

3.1 Research Questions

The two central questions underpinning this research are:

(1) What values are contained within the Open Archival Information System (OAIS) reference model and what methods or practices are prescribed by it?

(2) In what ways has the adoption of this model as an organizational system for the preservation of digital content in library, archive, and museum spaces served to challenge or reproduce the hierarchies and discourses of traditional archives and memory institutions?

I argue in this dissertation that OAIS serves as a site of tension for the multiple audiences who play a role in the creation of digital archives and the general sub-discipline of digital preservation. One site of tension arises precisely because the authors of OAIS inscribed the model with values inherent to their own professional practices. While its creation was an open process compared to other standards developed by CCSDS, it was still driven by the space science community and took place in a very elite, first world and predominantly North American context. Yet, it is simultaneously deployed in a variety of locations where the local values and epistemologies may come into conflict with those inscribed in the OAIS reference model.

Because of the hybrid and meta nature of digital preservation as a sub-discipline of information sciences, it is informed by the dominant discourse in all of the fields from which it borrows. Additionally, this participation is a space for constituting new discourses about the role of technology in memory practices and the role of the memory institution in the digital era. In this case, discourse refers to a coherent pattern of statements across a range of sites (Green, 1990) that has the power to produce the things which it purports to be describing (Rose, 2011). As such, this research examines how the actual use of OAIS in

the field speaks to existing discourses and the role it plays in the creation of new ones particular to the evolving profession of digital preservation.

3.2 Collection Methodology

These research questions presume two spaces of analysis: understanding the use of OAIS in the field and understanding the values and discourses implicated by this use. In order to investigate these questions, this research project involved two lines of inquiry: an analysis of semi-structured interviews with a variety of practitioners and scholars who have come into contact with OAIS and of the OAIS reference model itself as a boundary object and sociotechnical system. The method of data collection, which I describe in detail in section 3.2.2, focused primarily on semi-structured interviews with people who encounter OAIS in their professional activities. Through the interviews with practitioners, I can assess the role that OAIS plays in their profession; examining the crises identified by interview participants in concert with a close reading of the documentation about the object itself allows me to examine the values contained within the model itself as well as the values attached to it by users when it gets deployed.

3.2.1 Purpose Statement

In this section, I offer a three-part justification of my research methods: why semi-structured; why this kind of analysis; and why the focus on cultural heritage. I argue that there is a tension between the network of technologies, standards, and human/institutional actors that make up the sociotechnical network that is indicated by the colloquial use of the term OAIS; the intentionality or scripts of the design; and its real-world uses. By comparing interviews about how OAIS has been implemented and literature about OAIS, digital archival practice, and sociotechnical systems, I can understand the values contained within

the object itself and how these values play out when the object is used in memory practice. In the previous chapter, this dissertation provided a meta-analysis of the literature that suggested the tension incurred by OAIS and its various implementations. This tension also manifests in the ways in which the reference model is formally described in documentation; by practitioners who have learned about its use for work; and the ways in which it actually functions in practice. This is occasionally described explicitly in literature, where people find that OAIS does not neatly serve an organization's functions. However, these tensions are not often legible even to the authors of the literature. Instead, they often describe the formal terms of OAIS, and in describing their deployment reveal the spaces of friction that come from applying the model. To that end, a meta-analysis of the literature allows us to see tensions in what otherwise appear to be purely descriptive explainers of local implementations.

Second, semi-structured interviews confirm the existence of these tensions between what authors intend the model to do, what practitioners feel it should do, and what they actually do with it. This is present, albeit indirectly, in the existing literature. This dissertation aims to make these tensions explicit. In cultural heritage and libraries, these differences are particularly apparent. Describing objects with very large and diverse user communities or complex media objects that reference outside materials are difficult within the current iteration of the model[23]. Articles with titles like "Beyond OAIS" (Nicholson and

[23] I use this phrasing for a number of reasons. Firstly, because OAIS must continually undergo revisions as long as it is an ISO standard; it was last revised in 2012 and is currently undergoing revisions to be formalized in 2017. Second, because I do not believe that there is something fundamental to the core of the model that makes it incommensurable with solutions for preserving things like distributed content. There are two aspects to this: firstly, it is possible that the model itself can be updated to handle this content; secondly, it is also possible to build new standards or industry protocols that sit

Dobreva, 2009) suggest the struggles that cultural heritage institutions face. This research project continues this line of inquiry by seeking out the situated stories of users. In asking interview participants to characterize OAIS, the tensions between the formal model and its application come to light. In order to understand the implications of its ubiquitous use in preservation, I use interviews as a data collection method in order to find out how people characterize OAIS and their relationship to it in order to understand what OAIS does within the networks constructed around it.

I focused the semi-structured interviews on institutions and researchers working in the realm of cultural heritage. This choice stems in part from my own previous work and research experience in the area of film, art, and video games. My work on the Preserving Virtual Worlds II (PVWII) grant served as an inspiration for this dissertation: while exploring significant properties of video games, the situatedness of significance became very apparent (Hedstrom and Lee, 2002). One of the challenges in documenting significance for preservation is that the social construction of significance means that what is most important about a game is as likely to be something external to the code of the digital object as it is to be imbricated in the object itself. This is where the difficulty for preservation arises: what needs to be preserved is not a particular instantiation or a particular object, but a larger *work*[24] (IFLA, 1998; Tillett, 2003). This *work* must be

---

atop OAIS that might address issues like vast Designated Communities and distributed content. The particular tensions around these two issues will be discussed in chapters 4 and 5 respectively.

[24] I use this term in the sense it is employed by Functional Requirements for Bibliographic Records (FRBR). The FRBR group 1 entities are, hierarchically:

- *Work*: a "distinct intellectual or artistic creation."
- *Expression*: "the specific intellectual or artistic form that a work takes each time it is 'realized.'"

represented within a digital archive as a digital object or collection of objects, which in turn must be artificially bounded because what fits within the broad notion *of* work is fluid and changing. With boundaries, it becomes possible to pick an array of discrete objects that serve as a representation and preserve these, a process and view of what digital preservation means that is not currently standard. The field still struggles with some forms of basic bit-level preservation; the notion of preserving a collection of items to represent a *work* is on an entirely new scale of difficulty. This type of preservation activity would include the preservation of digital objects like the code as well as documenting other items, such as information about the social and technical context of a specific time period embodied in existing documents, or in purposefully created documentation about the Knowledge Base of a particular set of users. Cultural heritage institutions are thus an exemplar of the tensions between OAIS standards and implementation and the larger cultural context in which they operate given the complexity of their audiences and digital objects.

An investigation of the role OAIS does or could play in the growth of preservation practice to adequately address these issues does not lend itself to quantitative methods. Digital preservation, as a field and in its literature, acknowledges the need to account for the significant properties of an object as part of preservation practice, while simultaneously acknowledging that it has not yet found a way to address these issues

- *Manifestation:* "the physical embodiment of an expression of a work. As an entity, manifestation represents all the physical objects that bear the same characteristics, in respect to both intellectual content and physical form."
- *Item:* "a single exemplar of a manifestation. The entity defined as item is a concrete entity."

(McDonough et al, 2010; Webb, Pearson, and Koerbin, 2013). Key here is that, within the profession, there are no agreed upon manual methods for selecting and bounding the content to comprise complicated Information Packages that ensure on-going manifestations of the most significant properties. Because this kind of consensus has not yet been agreed upon, there are likewise no machine-actionable methods for creating this content as of the writing of this dissertation.

Rather, given the nascent stage of this part of the sub-discipline, it is necessary to gain a clearer picture of the situation from within which each participant expressed these experiences and views. Since cultural heritage objects are inherently socially constructed and a product of a particular place and time, so must their preservation be.

3.2.2 Semi-Structured Interviews

Semi-structured interviews enable me to garner information about the situation and context within which the participants labor. A qualitative approach allows for the mutual construction of meaning between the interview and interviewee. Because it was my intention at the outset to interview participants with varied backgrounds and experiences, a rigid interview schedule did not make sense for this project. Rather, the levels of experience differed from participant to participant such that while I was able to ask questions along similar topics in most cases, specific questions and phrasing needed to be adapted to each interview. The interview process was very iterative: if I noticed a trend in previous interviews, I incorporated questions about these in future interviews in addition to the general topics common to all interviews.

This variance among participants has it benefits, and there were also ways in which this entailed more work on my part as a researcher. Madison (2005) points out that the

more precise the questions in the interview process, the more concise the data will be and that codes for later analysis will flow in a more obvious manner. For this research project, it was important to balance a desire for simplicity on one end with the recognition that the varied experiences and roles fulfilled by the subjects meant that the questions could not be as detailed and specific at the outset as they might be when interviewing a group of subjects with more homogenous experiences and thus, perhaps, more homogenous responses.

The topical interviews I engaged in with subjects encompassed elements of both an oral history, or the recounting of the emergence of OAIS and what that meant for participants' practice/profession/industry, as well as personal narratives, or their individual expressions and perspective on OAIS, their experiences with it, and points of view. The oral history component is important because I am looking for information that is at the intersection of personal subjectivity and a particular place and moment in time (Madison, 2005); this is important for locating OAIS within the general landscape of preservation discourse, particularly when looking at how interview participants characterize OAIS itself. This work effectively classifies OAIS, a task which necessarily involves spatial and temporal elements (Bowker and Star, 1998). Within the interviews, my focus was not about what happened or happens, and indeed such information is largely captured by the existing literature about OAIS. Instead, the focus was on the how and why: not "what are you doing", but rather "why do you do it that way", for example.

3.3 Specifics of the Data Collection Activities

I conducted semi-structured interviews, the purpose of which was to get participants to examine how implementation of OAIS in their particular institution is

connected to the way they practice. In total, I conducted 19 interviews and spoke with 28 different individuals, 20 based in the U.S. and seven in Europe. The discrepancy between the number of interviews and interview subjects occurs because I interviewed some subjects together. Indeed, where possible I tried to conduct interviews with multiple people at the same time as I found the information garnered from these types of interviews to be the most enlightening: when people spoke to their colleagues or appended my questions with questions of their own, it yielded rich data.

3.3.1 Participants

In choosing interview participants, I looked at a number of criteria. My prior experiences speaking with practitioners who, anecdotally, expressed frustrations with OAIS that they identified as stemming from its origins, combined with findings from the literature and the nature of my previous work led me to concentrate my interviews on those working in the cultural heritage and library sectors. While I began with a cultural heritage focus, I also aimed my net broadly to capture a variety of different perspectives. I wanted to speak with users in both Western Europe and the US, because these are the two geographical sites that figured most prominently in its creation and that feature most prominently in the literature about OAIS deployment[25]. I also hoped to capture the experiences of people across the sociotechnical networks that form around OAIS. This meant talking to people who work specifically with the model itself such as technology and research personnel in large institutions; founding digital project managers in smaller

[25] I overlook here the obvious privileging that allows European and US institutions to engage in pricey preservation projects and the production of academic literature about these endeavors.

organizations; and digital project managers in paper organizations making the gradual shift to digital work. I also intentionally spoke with people who were more tangentially aware of OAIS, those working in technology and systems support roles who perhaps have heard of OAIS but are not trained in library or archival science and thus are aware only of the functions of the model for which they are responsible.

The final group of interview participants is the worker who perhaps is not aware of OAIS at all but is required to engage in digital preservation work in an institution that is using OAIS. In my field work, these tended to be traditional archivists and catalogers at large, long-standing institutions who were required to adapt to the digital changes occurring in their organizations. Their familiarity with OAIS would vary: some people I spoke with had never even heard of OAIS, despite their organizations being leaders in its deployment and the fact that they headed major departments. Their impressions of the ways in which "the digital" has impacted their profession is key to understanding the sociotechnical implications of the ubiquity of OAIS, even if they cannot directly characterize the object and even if these interviews required a very different set of questions than those posed to the other interview participants. In all phases of the research, particularly the last phase, I had to consider that the roles of participants were sufficiently varied such that OAIS did not function as a known boundary object.

The interviews took place in three phases, and the participants in each are described in brief below.

Phase 1: This was the first round of interviews conducted. It followed the most scripted pattern in comparison to other interviews as the OAIS experiences among these participants were fairly homogenous. This phase was time-limited based on travel

constraints. The interviews took place in a Western European country that has been actively involved in the construction of OAIS and its attendant standards like ISO 16363 since the late 1990s. This participation was reported by the participants themselves and confirmed by a review of OAIS literature. I began with participants in an audiovisual archive working in the office of information and research and development. Through recommendations, I spoke with participants at a national library and a center that deals primarily with performance art. All three institutions are tasked with its preservation by the state. In the national library, the interview subject was the head of the research department. Due to the small size of the performing arts organization, the participant there was in an administrative role and wore many hats. The three organizations were located in three different cities, but had considerable communication between them. All three sites were aware of the work of the others, and all participated in state-wide and European Union-wide preservation initiatives among other shared projects. Of the five interviews conducted for this dataset, four were audio recorded. The fifth was not recorded as the interview subject opted not to sign the consent form until after our conversation had concluded; instead I took notes.

Phase 2: This phase follows the ISO accreditation of a new standard related to OAIS in 2014, the ISO 16919 standard that certifies auditors to conduct ISO 16363 audits for trustworthy repository status. As this standard was built on the OAIS model very intentionally, many of those involved in its creation were also involved in the creation and revisions of OAIS. I attended a course related to ISO 16919 and ISO 16363, and over the course of one week attended sessions about digital preservation standards from about 8am to 5pm each day. During the course of the training, I was able to speak with the instructors

about the standards and ask them questions about OAIS and their experiences with it. The instructors were predominantly from Europe; one was based in the US. Three came from space science and one from state archives. This perspectival difference was noticeable in the data collected. Part of the data stemming from this phase of the research includes the instructional materials distributed to participants in the course—the content analysis of this dataset includes these materials as much as the conversations. Unlike the other two datasets, these conversations were not audio recorded. Instead, I took extensive notes.

Phase 3: The data garnered during this phase is the most varied as it involved participants across North America over the space of approximately six months according to the availability of the interviewees. The initial site of investigation was a large private university. I chose to begin here because of its large audiovisual collection and because it has archives that are dedicated to the representation of political and non-mainstream causes. Because this research looks at the ways in which digital archives can resist the hegemonic discourses in traditional memory institutions through new technologies and policies of preservation, finding "reflexive archives" was a priority. In this one institution, I interviewed six different people involved at different levels of preservation. This sub-sample included three traditional archivists who have had to adapt to working with digital materials and systems; these archivists were trained in paper and even papyri. One was the head of a sub-unit at the university, while two others were archivists within different sub-units. The three other interview participants worked in the department that handles digital technologies for the university's collections in libraries and archives. Two worked in more macro policy and research roles, and a third in much more technical spaces. Five of the participants had long-standing ties to the organization; one was newly hired.

From here, I followed contacts and suggestions to a variety of different institutions. I can roughly divide these interviews into practitioners and researchers. I spoke with practitioners in a public library, a private archive, three different private universities and colleges, a museum, and a consulting firm. I also spoke with science curation researchers as well as two researchers who work in digital preservation more generally.

After conducting the interviews, I was surprised in practice by the total lack of awareness of OAIS on the part of some participants. This is not pejorative or a judgment to say that participants lack knowledge they should have, but rather an interesting finding that points to a fragmentation and siloing of labor roles involved in preservation work: institutions known for their adoption of OAIS and their role in digital preservation have numerous preservation employees who have never even heard the acronym OAIS. For those working directly with OAIS as part of their daily tasks, many pointed to the common language stemming from OAIS as part of its appeal. However, there were participants, particularly those whose work began in and still focused more on analog materials, who used very different vocabularies and knowledges to describe their relationships to the technological object in question. The work tying the datasets together into common groups occurred during the coding and analysis processes (Madison, 2005; Carspecken, 1996; Lofland and Lofland, 1984).

I sought different classes of interview participants based on their experiences with OAIS and the types of labor they perform in the preservation environment. Unintentionally, the types of roles occupied by interview subjects also tended to follow patterns based on geographical location. Participants in Europe working in preservation-oriented institutions filled roles such as a research and development (R&D) manager and chief information

officers responsible for overseeing OAIS and translating it for the institution and technology teams. In other words, these are people who are tasked with really understanding what OAIS is and interpreting it for others. In the US, the interviews covered a significantly larger number of institution types, but in those instances where people were working in organizations with a preservation focus, the participants were more likely to be what I call practitioners, those working generally in the area of digital preservation and handling a number of day-to-day preservation tasks. I spoke with participants who work with the model without knowing much about archival or other preservation theories/sciences.

3.3.1.1 Sampling

This research relied heavily on snowball sampling: I began with just three specific people or institutions that I wanted to speak to, with the intention that they would tell me who they thought was doing the most interesting work in the areas of OAIS and cultural heritage. In each instance, interviewees mentioned additional institutions or individuals they thought could contribute to the project and I followed each recommendation given to me. I was able to contact all but three of the people recommended by interview subjects.

It was not my intention at the outset to interview people in the realm of science data. However, as the interviews progressed, it became clear that many of the participants assumed that the application of OAIS within science disciplines was easier given that OAIS originated in space science. They assumed that the creation of a Designated Community in particular would be a simpler process in science, while describing this process as challenging if not impossible for themselves. As such, I decided to add interview participants who work with science data. While these are not meant to be representative—

I do not assume that two people speak for all of science—I felt their perspectives were necessary to round out the data. I especially wanted to ask them the question: is this easier for you because you are in the sciences? They also brought an interesting curation lens to the project, as both work in data curation.

3.3.2 Locations

Here, I detail the geographical locations for the interviews. I aimed to capture perspectives from a variety of locations, and I make these locations explicit here because of the large role geographical difference played in perspectives about OAIS and the kind of work being done with it currently. The initial points of contact for interviews varied in their geographical location, because I purposefully chose one institution in the US and one institution in Europe. The third planned inquiry took place as part of a training course about a standard related to OAIS. This training took place in a third country. I conducted these interviews at different points during the research project; snowball sampling and geographical separation meant that I conducted a number of additional interviews in between each of these preliminary contact points.

I set up the first set of interviews via a professional contact at an audiovisual state archive in Western Europe. I chose both the person and institution for a number of reasons. The institution is key here, because the nature of cultural heritage preservation in Europe in particular means that practitioners often work both in a particular institutions and for many broader preservation projects funded centrally by the European Commission (EC). The country was selected based on a review of existing OAIS literature and looking at the early non-science adopters of the model; this one was a particularly early and active developer and adopter of OAIS. I had only one contact here thanks to a colleague, and

through him was put in touch with two other participants working in the same national institution.

The second set of interviews were conducted with professionals who were involved in the creation and deployment of OAIS. The opportunity to speak with some of them arose through a training program for a different digital preservation standard. This opportunity also occurred in Western Europe and yielded a distinct yet informative body of data from the OAIS author perspective.

The third planned inquiry took place in the United States. I chose to conduct interviews in a large East Coast city given the concentration of entities engaged in digital preservation located in or close to the metropolitan area. I began here with another set of professional contacts established through my earlier coursework in preservation at a large research university that engages in digital preservation through its library. In this case, I started with the institution (the university) and reached out to several contacts spread throughout branches and departments there.

Where possible, I conducted the interviews at the institution where the interviewee worked. These site visits are of interest because discourses are occasioned by their rhetorical strategies and claims to truth, but also by their institutional location; these institutions act as apparatuses for discourse formation and so are worth studying in and of themselves (Rose, 2011). While the study of the institutions themselves is not the focus here, these undoubtedly play a role in the variance of responses from different participants. I did conduct a number of additional North American interviews via Blackboard Collaborate in order to speak with participants who were located in places I was unable to visit.

3.3.3 Interviews

Interviews lasted between 30-150 minutes, with the average length being around an hour. In all cases, I began by asking for consent according the IRB exempt classification of my research project. In most cases, I followed this by a request to record the conversation. Audio recordings were subsequently transcribed for analysis.

3.3.3.1 Interview Questions

Given the different levels of familiarity with OAIS on the part of the interview participants, I did not use a specific interview schedule. I also added questions iteratively: as themes arose in analysis of audio recordings and transcripts, I specifically asked about these in subsequent interviews. While the interviews were semi-structured so as to allow changes that are informed both by subjects' experiences and the variety of different kinds of work the subjects do, the questions centered on a similar set of themes. Common topics of conversation included:

(1) Characterizations of OAIS

(2) Experience working with OAIS and related standards like PAIMAS and TRAC/ISO 16363

(3) Alternatives to OAIS

(4) What is helpful about OAIS

(5) What is constraining about OAIS

(6) Awareness of others using OAIS and the relationship of their institution to others engaged in digital preservation

(7) I concluded all interviews by asking the participant(s) if there were any questions or themes they thought I would ask about but did not or if there were additional things they wanted to share with me.

Where it was relevant, I asked practitioners to share copies of internal policy and workflow documents. As themes about Designated Communities, future iterations of OAIS, and differences in adoption between the US and Europe emerged, I began asking subsequent subjects about these themes. In particular, I asked participants to conjecture about causes of geographical differences in adoption and was explicit that I was asking for conjecture. I was similarly careful about asking about the future of OAIS and upcoming revisions: I was explicit that I was inviting participants to speculate and that I wanted their informed speculation.

As the project progressed, I came into later interviews where subjects knew that I had been working in this area for some time, and I was solicited to explain certain aspects of the model for participants. In particular, subjects would say things like "you probably know more about this than I do" or "I am probably getting this wrong". Additionally, there were cases where people spoke erroneously about requirements in OAIS. I did not in any of these instances "correct", although I did offer clarification if subjects asked for specific information.

3.4 Frame of Analysis

The method of analysis involves examining the provided texts, interview transcripts, OAIS documents, and materials given to me by interviewees. The framework for analyzing these texts is qualitative and reflects my own epistemological leanings.

I use the term "texts" to refer to the objects of analysis in this project. While much of what I examine are written documents, there are also a number of "verbal texts" in the form of interviews, which are then rendered into fixed text via transcript. There are many things that are conveyed in a verbal exchange that cannot be conveyed in a transcript, and it is a limitation that I predominantly coded transcripts over the audio recordings themselves. The use of text over speech shifts power away from the speaker to both the author (me, in this case) and readers:

> Compared with speech, writing is a way of storing language, fixing it so that it can be read by those not directly intended by the author. Writing thus promotes the transmission of culture from generation to generation, the transformation of cultural works into monuments and the elevation of authors in authorities. Writing also fosters the development of critical thinking on the part of the reader: by stabilizing the words on the page, the reader can reflect upon them, go back to earlier passages and re-examine links of argument, and accomplish all of this in isolation without the presence of the author or community exerting any pressure on the act of interpretation. (Poster, 2004, page 405)

This is important to note because it is yet another incidence of the role my own biases play in the construction of the work: I had the compounded privilege of selecting and conducting interviews; overseeing the transcription process; and providing my own interpretation of these materials through this dissertation.

3.4.1 Coding

I used qualitative, iterative coding (Charmaz, 1983) to begin the analysis of texts. I coded for specific terms within OAIS, particularly where they appeared frequently in

interviews. These included: SIP, AIP, and DIP; Designated Community; Knowledge Base; and Representation Information. I also coded nodes where people spoke specifically about their use of and experiences with OAIS, including nodes about characterizations of OAIS; descriptions of institutional deployment; and explanations about why practitioners chose to use it. I used a number of *in vivo* codes where I noticed trends, such as:

- THAT OAIS diagram[26]
- Alternatives to OAIS
- Services
- 2012 Revisions
- We are/aren't doing a good job
- Bit level preservation is hard enough without extras
- Audits (a set that included formal uses of audits, informal uses of audits, TRAC and ISO 16363, DIN, Nestor, and Data Seal of Approval)
- Tools (a set of child nodes that included general mentions of tools used for implemented preservation work, as well as frequently mentioned specific tools including, most predominantly, Archivists Toolkit, Archivematica, Islandora, and FTK)

I also employed codes aimed to collect data for the identification of discourse and scripts, including themes about cost realities, decision making protocols, constraints, and values. I

[26] I add capitals here to note the emphasis with which numerous participants spoke about THAT diagram. I do not actually think this always refers to a single canonical diagram within OAIS: OAIS contains many and there are a few that are used with regularity in papers or presentations about OAIS and preservation more generally. Many people complained of seeing "THAT diagram" at too many conferences or asked me to give them a minute to find "THAT diagram" online to refresh their memory.

used NVivo coding software to make codes, code sets, and memos on transcripts, OAIS documents, and e-texts submitted by interview participants.

3.4.2 Analysis

While my method is not ethnography, the types of interviews I conducted were influenced by the commitments of this type of work. Ethnographic work speaks to my own epistemological positionality in how I ask questions and how I describe and interpret data. This positionality is exemplified by the way in which Madison (2005) describes the work of ethnographers:

> As ethnographers, we employ theory at several levels in our analysis: to articular and identify hidden forces and ambiguities that operate beneath appearances; to guide judgments and evaluations emanating from our discontent; to direct our attention to the critical expression within different interpretive communities relative to their unique symbol systems, customs, and codes; to provide insight and inspire acts of justice; and to name and analyze what is intuitively felt.

These 'critical expressions' are precisely what I seek to elucidate from the interview data, and these are the types of questions I ask of my data to conduct the analysis. As per Madison, when I go to these locations and engage with participants within their own professional spaces and situations, the work is ethnography whether or not I choose to label it in this way. Through my interpretive standpoint I represent a place and people to others, and it is worth describing the interview exchanges in this way because it renders visible how much power in I hold in being allowed to make a public interpretation of the situations and experiences of others.

I include this to highlight the hybrid nature of both the investigative as well as analytic processes. I combine methods and frameworks that are grounded in similar epistemological concerns that echo my own: I aim to make meaning in concert with participants, to understand the power implications of the mass adoption of OAIS and the role that the values inscripted therein play on the discourses of memory practices in the digital era. My data collection method was to conduct semi-structured interviews as well as to collect documents: official documents about OAIS as well as those submitted to me by participants. This work is qualitative and ethnographic in nature.

In order to analyze these data, I used a single method of analysis informed by two theoretical frameworks. I iteratively and qualitatively coded texts. This is the primary analytic method. I engaged with the coded content by asking questions of my data. These questions are inspired by and grounded in the guiding theoretical frameworks of this project: Foucauldian discourse analysis and Akrich's work on scripts in technology. While I explain these in greater detail below, let me preface this by explaining why I use these two. My data collection yielded information about OAIS itself as well as the institutional space within which it operates. As such, I sought complementary methods that allowed me to investigate these two related but separate concerns. Akrich's work on scripts provides a language for speaking about the values put into OAIS by its authors. Discourse analysis focuses on the institutional materiality of power; in other words, this gives the language to speak about what happens when those values get unpacked in a variety of institutional contexts.

3.5 Limitations of the Study

OAIS is a very large object, if it can even be seen as a single object. Furthermore, it is widely adopted, as I have discussed above. As such, any project on the OAIS Reference Model is, by virtue of taking on such a topic, limited in its scope. I am choosing on purpose to focus on small parts of it that seem important to the populations I am interviewing.

Additionally, my sample is small, consisting of 28 participants. While I cover a wide variety of institutions, the data I gathered is not representative of all of the types of institutions using OAIS. The variety of institutions may also constitute a drawback: because I looked at so many institution types, I gathered between one and six perspectives from each, meaning that these results cannot speak generally to the experiences of most practitioners in each of the areas. Additionally, I spoke to people who are practitioners and scholars about OAIS. My sample does not include users of OAIS Archives or content producers, although I introduce the latter in Chapter 5, where I include interviews conducted with video game programmers. These producers were not interviewed within the context of OAIS, but nonetheless provide a different perspective. Additional research is needed to examine the ways in which OAIS transmits discourses to Consumers and the ways in which it constructs subjects through its activities.

There is also the potential language barrier to consider: seven participants are non-native English speakers, and while all work frequently in English-speaking institutions and consortia, there is the possibility of lost meaning. I chose to work with organizations in first world countries, many of which were large and very well-funded. As such, it is important to understand that despite the differences between the various interview responses and the anticipated heterogeneity of responses from subjects, most institutions were constructed within similarly powerful and mainstream discourses.

These experiences may not speak very much to people in other types of institutions and do not take into account the kinds of discourses found in and around smaller, alternative organizations; likewise, the preservation practices that may grow from this research may not be applicable at all types of institutions.

Rose (2005) describes Foucauldian discourse analysis as not inherently reflexive. She also describes the way Foucault and practitioners of discourse analysis have lacked elements of reflexivity in the ways in which they separate themselves as authors from their works and the subjects, and in the manner in which they make claims to truth. She also asserts how claims to modesty in the analysis process can have an effect of reflexivity; it will be important to this research to be mindful of any claims to truth or truth making that I might engage in while describing findings in the interviews and other primary materials on OAIS. Discourses are also, by their nature, persuasive and self-producing. In engaging in this space and making the call for reflexive archival projects that recognize their own discourses, I am of course creating one of my own. This is like the unavoidable catch of archives themselves: they can never be made neutral, nor would it be beneficial for them to be. At best, I can continually recognize and identify my own discourses and document them, rendering my own power centered and visible rather than de-centered and invisible.

Akrich (1992) names some of her own limitations in her work on the de-scription of technical objects. The size and complexity of an object like OAIS makes an analysis of all aspects of the entire object all but impossible. Here, I recognize that I am looking at a very limited piece of the object and that my analysis of the scripts related to it will also be partial. The analysis here is purposefully limited and does not seek to make sweeping statements about the entirety of OAIS and all of the institutions that use it.

Finally, my own epistemological proclivities constitute a bias that has a strong bearing on the project. My own post-structuralist leanings are at odds with some of the more positivist authors of OAIS, and indeed there were times where my questions did not make sense to them. Because my focus was to trace discourses, power, and scripts, I looked for these and asked about the themes and terms within OAIS that seemed most promising for research in this area. For example, while a large portion of the body of this dissertation examines the term Designated Community and while most of this is drawn from incredibly rich data on this theme provided by interview subjects, it is important to note that I solicited this information and it did not often come up organically in a conversation that was generally about characterizations of and experiences with OAIS. I recognize that this dissertation is about a particular slice of issues within digital preservation broadly, while also suggesting that these findings provide a useful heuristic for thinking about the relationship between the digital and analog eras in archives and futures for preservation practice.

CHAPTER 4

AUDIENCES, ARCHIVES, AND GEOGRAPHICAL DIFFERENCE

Introduction

This dissertation examines the hypothesis that the Open Archival Information System Reference Model (OAIS) serves as a site of conflict and tension for multiple audiences in ways that impact preservation work in memory institutions. Implicit in this hypothesis is the notion that OAIS performs myriad roles in a variety of spaces: within specific institutions; within the developing practice and professional boundaries of preservation as a sub-discipline of information science; and within larger spaces of discourse about preservation, archives, and the practice of collecting and keeping things over time. The hypothesis also assumes a variety of audiences, including content producers and users as well as digital preservation practitioners and other workers in institutions charged with doing preservation. This chapter uses the interview data described in the previous chapter to investigate these roles and audiences. I discuss three of the themes that arose during the coding of the interviews and other texts: audiences and OAIS; the traditional and changing roles of archives; and adoption differences in Europe and the US.

4.1 Discourse, Audience and OAIS

As stated above, the deployment, use, and continual revision of OAIS presuppose a number of audiences. Audiences identified within the OAIS documentation itself are laid out in general categories (CCSDS, 2012, page 2-2):

The simple model shown in figure 2-1 depicts the environment surrounding an OAIS.

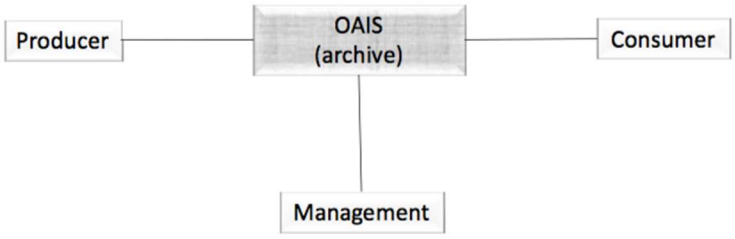

Figure 2-1: Environment Model of an OAIS

Figure 5: OAIS Audiences

According to the OAIS documentation, a Producer is not necessarily a person, but rather a "role played by those persons or client systems that provide the information to be preserved. This can include other OAISes or internal OAIS persons or systems" (CCSDS, 2012, page 1-14). The language of "role" here serves to clarify not simply that the actor could be nonhuman—it could be sensors feeding data into an Archive, for example, or an institution depositing content—but also that the Archive itself might be the source of the content. Consumer likewise refers to the "role played by those persons, or client systems, who interact with OAIS services to find preserved information of interest and to access that information in detail. This can include other OAISes, as well as internal OAIS persons or systems" (CCSDS, 2012, page 1-10). Consumers can be client systems rather than human actors as well.[27]

These terms have particularly defined meanings within OAIS that differ in connotation from their popular use, though they aren't wholly incommensurate. In interviews, subjects spoke about Producers in ways that connoted external bodies,

[27] In all my research I have never heard anyone describe a Designated Community of nonhuman actors, although this is explicitly accounted for in the definition of Consumer.

something akin to donors in the traditional archival sense. Interview subjects frequently cited donor and user agreements in reference to Producers. As a researcher I am guilty of this too: when coding interviews, I noted my own tendency to label any discussions of donors as automatically falling under the node of "Producers." The interview subjects made no explicit references to the term as defined in the OAIS glossary. This may be in part because use of this terminology predates institutional adoption of OAIS.

In interview data, a common theme emerged regarding Producers and producers qua donors: the people and institutions who donate the bulk of the materials were described as not caring about preservation. This was common across a number of institution types, and interviewees levelled this description as a complaint about broadcasters, twentieth-century artists, scientists, and humanists. In the case of researchers, the primary concerns as identified by practitioners in preservation and data curation were the production of scholarship and later access for additional scholarship. Access was a constant theme: interview participants described framing their preservation activities as a means to provide access in order to "sell" services to content producers. This distinction between preservation and access suggests a division that is not necessarily present in OAIS: while the model is designed to guide preservation activities, virtually immediate[28] access to content within an Archive is consistent with the informational model. When someone or something can query the system, find material, and get a response (the DIP), which is based on some kind of content within the Archive (the AIP), it does not matter whether this exchange happens hours or hundreds of years after ingest;

[28] That is to say, access that seems immediate to a human actor, as opposed to ability of a computer to keep track of the infinitesimal but real lengths of time it takes to process queries and kick out a material response.

time frames are not specified by OAIS. What constitutes "long" is entirely situationally dependent. One interview participant, a manager of digital preservation at a European national library, noted that her department voted on a definition for the concept of "long-term": "So I tried to, I advised the organization let's start with policies, then at least we have starting point what we want to achieve, if we talk about long term preservation, what do we mean when we call long term, is it five years, 10 years, 100 years? Nobody asked that when they were collecting paper collections but for digital collections you need to say something about it, it's now 100 years, we voted for it, it's 100 years, at least we have a figure." The glossary within OAIS defines long-term as, "A period of time long enough for there to be concern about the impacts of changing technologies, including support for new media and data formats, and of a changing Designated Community, on the information being held in an OAIS. This period extends into the indefinite future" (CCSDS, 2012, page 1-12). In other words, not a particular time period but subsequent to a series of defined events, similar to the temporal aspect of the move from SIP to AIP to DIP. The "future" in this case is a moving target. It is, of necessity, an imaginary unto itself. In the case of an OAIS and within the context of this thesis, it refers more to a place within the visual of the model than a time. It is the output or what comes out on the other side. While the Dissemination Information Package (DIP) must temporally followed the Submission and Archival Information Packages (SIPs and AIPs), this does not necessarily connote long length of time: it can happen in a matter of miniscule computer-calculated pieces of seconds that seem instantaneous to a human interacting with a system. It could also refer to information being requested from the system at some point in the colloquially defined future: tomorrow or 100 years from now. During a training course for ISO 16363 and ISO

16919 (standards for trustworthiness of digital archives and the standard to certify auditors to make such designations), an instructor, a European OAIS author with origins in space science who has also been heavily involved in the periodic updates to OAIS, noted that "long-term" means long enough for any aspect of an AIP to change. Within the context of an audit refers, the length of time something must be preserved is concrete: it is the time it takes an Archive to hand off its content to successor Archive(s). "'Future" is very situationally dependent.

Because people use OAIS to think about long time periods when in fact OAIS describes a set of relational processes that are not time-dependent, it creates a tension. While many practitioners recognize that OAIS functions do not necessarily refer to long periods of time, discourses of archives and preservation connote durability and longevity. Discourse refers to groups of statements that structure the way something is thought about, which in turn influences the actions taken within the spaces of the discourse. Popular discourse about the longevity of archives inspires technology companies to use the "archive" label for older web posts or emails, for example. The social spaces occupied by popular imaginaries about the role of archives produce the rules and conventions that govern how we understand the work of archives: practitioners are as much subject to these discourses as the general public. They understand archive models to connote periods of time, which overrides their understanding of process and order as described in OAIS documentation.

Language from interview participants suggest that practitioners within archives are equally subject to such discourses in the way it becomes difficult to understand the temporal component (or lack thereof) of processes described by the information model

within OAIS. It is hard to think about immediacy within the context of OAIS because it is employed precisely to guide thinking about keeping things for long periods of time. This creates a mismatch where one need not necessarily exist: an OAIS-informed system can be about providing ongoing access to content.

What was common across these discussions was that the Producers are less an audience for OAIS and more stakeholders in some larger network of which OAIS is a part. Participants denied using OAIS-specific terminology with donors and content producers. A technology officer at a European national archive said, for example: "No, I don't think we [are] bothering [external Producers] with SIPs and DIPs and [AIPs]… no, it's [that] we try to describe in the contracts what's required of [external Producers], in terms of which metadata they should deliver and which format they want us to preserve…" Yet while interview subjects do not use OAIS language with Producers, they simultaneously appeal to OAIS as a method to "sell" the services of the Archive to Producers. The head of digital programs at a US private archive explained: "...there is something about [OAIS] being a standard, and I think it actually helps that it is… science-based… that lends a credibility to it for our donors." This latter is not specific to OAIS, of course, but rather is part of larger discourses about standards adoption or what is purchased by an organization by investing resources in adopting and implementing a standard, even if only in part. While OAIS is not "implementable" as a reference model, the process of mapping workflows to it or employing tools and technologies built atop OAIS purchases a legitimacy with audiences of the Archive who are not necessarily audiences for the standard itself.

The role of Producers is intimately tied to the ingest process and the Ingest Functional Entity as described by OAIS.[29] In terms of post–OAIS developments, there has been a strong focus on the ingest function. As per Lee (2005) and Cargill (1997), there have been strides in the creation of industry standards by information technology providers. CCSDS created the Producer–Archive Interface–Methodology Abstract Standard (PAIMAS) to detail relationships between content producers and the deposit archive (CCSDS, 2004). This standard, ISO 20652, has not seen the same level of adoption as the OAIS reference model. The technology officer at a European national archive summed up the general ethos of most of my participants when I asked about the use of PAIMAS or whether any of CCSDS's other standards or models were useful to their daily work. The response: "not on [the] level of [OAIS]." Others were not even aware of the other standards.[30] Yet despite the fact that ingest is perhaps the best-defined part of OAIS, not because it is necessarily treated differently within the model itself but because it was the subject of subsequent developments, interview participants still expressed a desire for more specificity and guidance about the role of the Producer/producer because these roles are not adequately described by OAIS. This is true both (and separately) in the sense of dealing with external content donors *and* in the sense of acknowledging the fact that sometimes the Archive is the Producer. Multiple participants spoke about "pre-ingest" work, which they felt was not

[29] "**Ingest Functional Entity**: The OAIS functional entity that contains the services and functions that accept Submission Information Packages from Producers, prepares Archival Information Packages for storage, and ensures that Archival Information Packages and their supporting Descriptive Information become established within the OAIS." (CCSDS, 2012, page 1-12)

[30] Even among the already narrowed pool of interview participants who were well aware of the OAIS standard. I did not ask this question of people who self-identified as knowing little or nothing about OAIS.

well-defined by OAIS; at one institution, workers developed and named additional Information Packages specifically to deal the space between existing IPs pre-ingest and the staging space between SIP and AIP. The technology officer at a European national archive said: "Yeah, I think the whole part on pre-ingest part, of course, in OAIS you just have ingest, but we have for [our subject domain] the negotiations with the [donors/producers], we have sort of defined a pre-ingest phase where you do all the negotiations with your producers." This space, so comparably well-defined by CCSDS, was still considered inadequate by users of OAIS.

The audience for OAIS itself as conceived by practitioners plays an important role in power negotiations and the status of the Archive and archives more generally. Alan Sekula (1986) says, "…archives are not neutral; they embody the power inherent in accumulation, collection and hoarding as well as that power inherent in the command of the lexicon and rules of a language…" (page 155). The ability of some within an organization to use the fluent language of digital preservation connotes additional power for the Archive over Producers and Consumers. Moreover, this exists in relationships between the digital elite and those who work in related analog spaces within the same organization. Within the fragmented labor of the organization, for those who do not speak OAIS-informed language but need to conform their daily activities to fit within digital workflows, there is tension engendered by non-digital laborers' need to speak the lingua franca for communication purposes, their need for status and to be taken seriously by the digital elite, and their desire to continue with the daily work they have done previously. This desire is not a reticence to take on new tasks or arbitrary resistance to change. It is also motivated by the commitment of existing employees to the professional expertise they embodied in an era before "we

[were] all digital archivists…" (Owens, 2014).[31] The balance between enrolling existing participants without deskilling them or stripping away existing professional expertise can be encapsulated in two quotes from different participants within the same institution. Both are "traditional" memory practitioners who now have to shape some of their work to cope with the digital initiatives at the major private US university where they both work. One, a special collections curator, said: "I have watched my job go from being a working humanist to basically a soulless bureaucrat, technocrat… It's hard to do, we're expected to become not just cognizant. We're expected to do it all ourselves…." The other, a senior special collections archivist, said:

> So, I would say I'm trying to educate myself on those tasks that are really relevant to me as a supervisor, and a reference archivist, and working with donors. So, I'm only trying to acquire the knowledge that affects my daily duties, you know what I mean? … [There is] the often bandied motto: every archivist should be a digital archivist, which I totally understand, the point of that statement, but when you're managing the number of people and collections that I am, I cannot have the level of technical knowledge that [a member of the technology services staff] does.

Both participants express different levels of tension, with varying approaches to the changes wrought on daily work. But the limited knowledge of both partners about OAIS was telling: it speaks to whom an organization considers worthy of including within the

[31] Two participants mentioned this interview to me as explaining how they felt about their role within an organization, and none of them did this in a way that expressed explicit discomfort with the change in their role and day-to-day work. This speaks to the incredible diplomatic skill of Sybil Schaefer to enroll participants in digital projects without alienating them, something that is all the more notable given the stories I heard about alienation from participants at more than one institution in more than one country.

official parlance and whom it does not. This inclusion/exclusion places power in the hands of those who possess fluency of language. It is not just those outside the Archive who are disempowered by this: the introduction of the digital repository into organizations like archives and university libraries serves to fragment workers within these institutions, and the power, status, and funding shifts away from long-standing domain experts such as subject-specialist documentalists, archivists, and curators to digital services.

Questions of audience are imbricated in a discussion of discourses within the sub-discipline of digital preservation because of the ways in which discourses produce subjects (Rose, 2011). The audiences explicitly delineated in the documentation of the model include the Producers and Consumers, as well as staff within the archives, connoted by the management block on the diagram at the outset of this chapter. The writing above suggests where the model places more power and influence. This raises a follow-up question: what kinds of subjects does OAIS produce, according to the interview data? There are donors, who are variously described in the interviews as: outside the archive; enthusiastic participants; bread and butter of organizational activities; frictional partners in the creation of memory; uneducated about the role of archives and particularly the digital functions therein; and uninterested in preservation. There are the practitioners themselves: those who wrote OAIS and continue to evolve it through updates and new, related standards and tools as well as those who read and write about OAIS. Within an Archive itself, there are divisions among the engineers and designers and memory practitioners, particularly those with an analog focus. This manifested in interviews and exchanges; there is a particular tension even among authors of OAIS from different sides of this divide as articulated by participants.

While more than one participant explained that different approaches and different difficulties in implementation were not domain-specific—that is to say, there was not a noticeable difference in working with cultural heritage data over science data within the realm of digital preservation projects informed by OAIS—some participants noted an overreliance of memory institution professionals on the OAIS reference model. One researcher chalked this up to the lack of experience by less-technology- and systems-educated partners in digital preservation: their lack of knowledge in this area made them cleave unhelpfully to OAIS. This participant, a Canada-based digital preservation scholar, said:

> In terms of the attitude, I think I find the cultural heritage sector being much more stuck in the OAIS model than others. And so that might be because it's always worked so well for cultural heritage, but I actually doubt that. I think it's more because there is less of an engineering understanding. And the funny thing is—{the funny} thing maybe about that is that it's not about the know-how to actually design it.

OAIS conceives of a number of audiences, both for itself as a technology and for the products its technology has helped wrought. The ability to participate in the OAIS network through use of its language and an understanding of how its processes were derived fragments labor into the audiences "in the know" versus those who are relatively outside, even when some of these outsiders are actually working within the record-keeping spaces of archives. In the next chapter, I will delve more deeply into yet another conception of audience: the all-important Designated Community, which is the stand-in for users within OAIS.

4.2 Traditional Role of Archives

OAIS borrows heavily from archival theory, just as the field of digital preservation does. In trying to cope with the challenges of digital preservation, practitioners and scholars alike have sought analogies with existing practices of conservation, record-keeping, and diplomatics: sometimes these analogies are helpful, particularly as archives have begun to move in more inclusive directions. In other cases, they constrain what practitioners are able to do with digital technologies. In all cases, borrowing from archival science means the terminology and foundational concepts come laden with discourses and scripts of their own which are then incorporated and changed within the digital preservation environment.

> An OAIS is an Archive, albeit in a very particular sense:
>
> An OAIS is ***an archive, consisting of an organization, which may be part of a larger organization, of people and systems that has accepted the responsibility to preserve information and make it available for a Designated Community***. It meets a set of responsibilities as defined in the standard, and this allows an OAIS archive to be distinguished from other uses of the term "archive." (Giaretta, 2011, page 47, emphasis original)

While it may not be an archive or archives in a traditional sense, the use of this language is purposeful. The authors of OAIS are careful to stake a particular claim such that they are able to employ this term without having to engage with the field of archival science about definitions of the profession. The tenets borrowed from general archival theory here are the notion of a preserving information over time and doing something with that content such that it remains accessible in some way. There are also explicit assumptions in here,

bound to imaginaries about the role of archives. As mentioned earlier, an appeal to archival terminology connotes longevity of the information, an important move given that the creation of OAIS comes out of space science which began developing technologies like OAIS precisely because they became increasingly aware of the ephemeral nature of their digital content. The European OAIS author with origins in space science mentioned the importance of trying to hold onto digital intellectual capital. This language also purchases a reputation associated with imaginaries about archives: archives are supposed to be neutral, authentic guardians of historical evidence. This discourse dovetails with the largely positivistic epistemologies of science data. Indeed, one author of OAIS noted that the authors did not worry very much about authenticity in early versions of the model when the only content under investigation was science data, as though this type of information contains inherent authenticity that is conveyed simply by preserving readable content. The implicit assumption is that this kind of content is also neutral and needs only guarding, not interpretation. OAIS also appeals to archival language about Provenance as yet another way to guarantee authenticity. OAIS employs the term *Context Information*, which means something akin to "archival bond." So while OAISes are Archives, they are not necessarily all archives. By borrowing from this field and history of practice, so too does OAIS borrow from the dominant discourses of archives, which becomes apparent when OAIS is deployed in memory institutions in a move which is often concurrent with said institutions having to make changes to long-existing practices to cope with digital materials.

Though digital technologies may not necessarily result in major changes to memory practices, new technologies have caused a disruption in existing workflows in many organizations. Digital content behaves in such a way that it requires archivists to think

differently about their daily work. For example, value within an archival tradition is often based on notions of originality or scarcity: when processing a donated collection, institutions will deaccession commercial items, particularly those which are readily available in an open market and those whose intellectual property rights belong to a third party. The focus is on the rare and the unique. The volume and ubiquity and very functionality of digital content challenge foundations of value as assessed in this way (Smith, 2007). There is so much digital stuff, and with search engines like Google finding random digital detritus with such ease, digital stuff seems ubiquitous. Concepts of originality are hard to translate to a digital environment: there is no one single canonical version of a digital document, because every time a file is reopened or sent via email, a new copy is made. In fact, the work done in digital preservation is fundamentally different from analog work in the sense that in analog spaces, a discrete object is preserved, while for digital content, what is preserved is a set of instructions that allow for a recreation of the digital file. In this way, coping with the digital has resulted in changes in archives where the labor of official memory making and keeping take place.

The media-archaeological approach aims to deal with both discursive and nondiscursive elements in constituting nondeterministic narratives of the interactions between humans and technologies as a way to understand more about the technologies themselves. In this way, it is a promising method for understanding more about OAIS as a technology in the Foucauldian sense. The role of archives in constructing historical objects is bound together with the influence that media and memory technologies like microfilm and library catalogs have on the content and the ability to understand that content within an archive (Ernst, 2013). Ernst uses the phrase “expanded archive” to refer to this type of

analysis. I suggest that this notion of the expanded archive is something that becomes possible and even desirable with tools like OAIS, which allow for the documentation of contextual and interpretive information that has not been the province of archival science in previous historical moments. This is a disruptive shift in traditional archive practice. For example, Duranti (2010) explicitly separates medium and message in a presentation about coping with the "Digital Wild Frontier" as part of the InterPARES project. She defines *document* as "recorded information (i.e., information affixed to a medium in an objectified and syntactic form)" (page 10) as part of what makes a record; she defines *medium* as a "necessary part of the technological context, not of the record" (page 11) as part of the characteristics of a digital record. This language suggests that technologies, systems, and historical objects are somehow separate concerns, or perhaps concerns at separate spaces of labor within an archive: it is someone's job to deal with technology issues, which are a different set of concerns from those who deal with the actual content.

This type of thinking echoes my earlier findings about the fragmentation of labor in digital archives. What allows these conceptions of archive to live side by side with the digital archive is precisely these declinations in terminology: Ernst deals in expanded archives, OAIS deals in Archives, and Duranti works in a space that might be called traditional archives. These definitional differences also allow the three things to be put in conversation with each other, because it creates a space of sufficient difference that allows for engagement, as opposed to someone simply dismissing an OAIS as not being a "real" archive. Yet the work of an OAIS-based repository that considers media and message jointly is a challenge to existing notions of an archive. Representation Information within an AIP takes the inherently generative practices of archives and dials them up: preserving

content in an OAIS archive is an autocatalytic process that continually generates more content and more work for itself in a way that pales in comparison to the accumulation of benign neglect that happens in analog settings.

An additional challenge to more traditional archival practice is related to the desire for sufficient control over archival materials, something that is complicated by the external dependencies of digital content. At the outset, the ability to gather widely dispersed and nontraditional documents challenges not only standard collecting practices in many institutions but even their technological capabilities. In one of her East African case studies, Seles (2016) describes the way official government business is often transacted over mobile networks given the relatively ubiquitous access, stability, and intuitiveness of mobile technologies and networks in places where the availability of computers, consistent internet, and consistent electricity are not as present. Archives in this region were aware of such transactions but had no remit or method to obtain this data. A suggestion from the author was that archives be given the ability to set such a remit, to influence departments in the way they create and disperse documents such that they are better able to capture and preserve government records.

This notion of catching material earlier in its lifecycle, of pushing practice upstream, is a newer one for preservation practice. The example I use with my students is this: in the past, collectors did not warn sculptors not to carve statues in marble because their weighty extremities might fall off. Yet this is precisely what happens in many digital preservation exchanges. While some institutions and some interviewees made it very clear that they do not make format proscriptions or prescriptions with donors, particularly museums, in other places administrators were actively exploring solutions to extend the influence of the

archive back into production for future ease of ingest, access, and preservation. Entire projects, like the Preserving Digital Public Television initiative funded by NDIIPP have operated in this way (Rubin, 2010). Research and technology workers at a national archive in Europe described the process by which documentalists were being retrained under the OAIS-informed institutional reorganization as media managers with the job of working within production spaces as liaisons to aid in ingest processes.

Archives also traditionally exercise considerable control over the reception of the information that it disseminates. Cvetkovich (2003) says of archives, "The audience for the story is crucial to its effect…" (page 93), and in the space of the digital archive this takes on additional weight. In paper archives, access to content is limited: often, to access content, researchers have to travel to a particular institution and seek permission to work with materials which are then brought out by an archivist who may monitor the use of materials. The process is highly controlled so that archives can control the story via their control of the audience. However, through digital dissemination methods, these situational constraints can be removed. While requiring authentication of remote users is possible and even assumed or suggested (E-Ark, 2015), viewing and using materials happens in a much less-contextualized space when users query Archives from a distance and engage with the materials from their own homes or offices. This kind of distant interaction lends itself to more voyeuristic kinds of engagement with archived content. Cvetkovich promotes feminist interventions that aim to create opacity by encouraging witnessing instead of voyeurism, where witnessing "… requires a kind of participation on the part of the listener that is not merely voyeuristic" (page 93). This raises a question: can and should digital archives police the use of records? This is a question that archives need to answer in the

digital era, as the control they have previously exercised over records looks different now. What makes digital archives different is, in part, the wider or simply more removed audience.

4.2.1 Knowledge Bases and Traditional Archives

A knowledge base is defined in the OAIS glossary as "[a] set of information, incorporated by a person or system, that allows that person or system to understand received information" (CCSDS, 2012, page 1-12). It speaks to the shared understandings of a particular Designated Community and lightens the preservation load in terms of what needs to be documented in Information Packages. A common example is language: if all members of a Designated Community speak English, then a digital text in English need not come with English grammar, lexicon, and other language tools. If at some point in the future, the Designated Community comes to speak Spanish largely or entirely, then the Information Package would need to be changed to contain English-to-Spanish language assistance, as the Knowledge Base of the Designated Community has changed. OAIS requires Information Packages to change with time as Knowledge Bases change or at least require Archives to monitor for such changes, though few of the many implementation studies available on the application of OAIS in particular organizations touch on the methods or strategies to document this Knowledge Base or understand when and how it changes. Additionally, OAIS and OAIS-related audits explicitly do not require formal documentation of a Designated Community's Knowledge Base. As such, this is a particularly difficult concept within OAIS, one that was described as being unexpectedly difficult to account for during the creation of ISO 16363 and the test audits that were done to vet the standard. I argue that part of the difficulty here is bound up in the role discourses about the

traditional role of archives play in institutions where OAIS is deployed. Elitist and imperial notions about what work a memory organization does and for whom prohibit the process of exploring Knowledge Bases and, more pertinently, of expanding the type of work an archive does to supplement any deficiency between an Information Package and a user, given what is or is not a part of their Knowledge Base.

The Knowledge Base is supposed to contribute to the understandability of the digital objects preserved within an Archive. Digital objects are meant to be independently understandable to members of the Designated Community; what exactly this means is nebulous and part and parcel of some larger concerns that arise with the use of OAIS. I begin with an example. While working as a researcher on the Preserving Virtual Worlds II (PVWII) grant, my primary focus during the game-analysis phase of the work[32] was an analysis of the video game franchise Carmen Sandiego. Throughout the course of this work, I was tasked with documenting the content of the game play experience in addition to technical metadata about a variety of the games within the series in order to investigate significant properties: I was tasked with understanding what makes Carmen, Carmen. Among many findings, what made her a special video game heroine was her detailed backstory, a pioneering idea that came about because one of her creators was a former Disney animator who brought with him the Disney practice of creating detailed "character books" that described history and personality traits for main-story characters.

What Carmen would and would not do was largely conceived of during this early creation process, and this came to have enormous meaning throughout Carmen's transmedia journey from 5-¼ floppy disk to television character. She is a product of a

[32] PVWII, its phases and deliverables will be detailed further in Chapter 6.

particular time and place, and carries signifiers that indicate these things immediately for those in the know, that is to say a particular Designated Community with a common Knowledge Base of basic twentieth-century Western history. Among other things, the character Carmen wears a red coat: as a figure of mayhem and espionage in the 1980s, this color connotes her background as a vaguely Soviet bad guy *à la* Russian villainesses from the Bond franchise and similar popular media sources. This red coat is one of the few ways in which this affiliation is presented in relationship to the character: her other traits, such as her name and hair, code her as vaguely Latina, in itself another trope to indicate how dangerous she is. But the connotation of the red coat is lost when she is transported to a post–Cold War era; undergraduates, as of the writing of this paper, are too young for this connotation to have any meaning. Is Carmen's red coat significant enough that an Archive should invest in preserving a record of it? For example, should the Context entity within a Carmen AIP link to an article or an interview with creator Gene Portwood, where these links between red and Russia are made explicit? In a digital era, the addition of this kind contextual information is not outside the realm of possibility; it is not even that difficult or resource-intensive to do. McDonough et al (2010) suggested the following model for an AIP for video games:

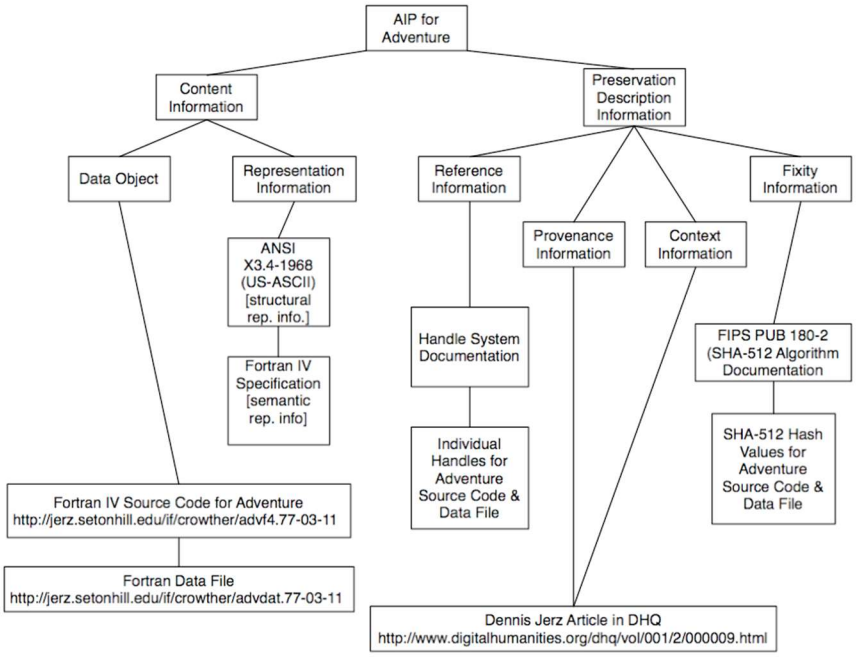

Figure 6: Adventure AIP

This mock-up of an OAIS package for the video game *Adventure* suggests utilizing a couple outside resources to fulfill requirements for three entities: Data Object, Provenance Information, and Context Information. Using a model like this would make the red coat of Carmen Sandiego understandable to a Designated Community of North Americans born after 1985.

So how could the preservation of Carmen's red coat work within an OAIS-informed archive in actual practice? During a session on sustainability auditing, a European OAIS author who is an archivist at a state archive said:

> We have so many things in the history of mankind, [it is okay to] lose some bits.[33]… We have preserved a lot, and we lose a lot because otherwise it is impossible. Preservation is always losing something. In the digital environment you lose less. You add things! [It is the] opposite issue: we preserve too much! [We] have to have a balance some way.

This quote suggests that this notion of adding information in Information Packages is a feature that is made newly easier by digital and networked technologies. Yet this author also questions whether or not archivists should be preserving this additional contextual information. This is echoed in a comment made by a manager of digital preservation at a European national library about the introduction of the PREMIS metadata standard:

> …then came PREMIS 2007 and the pressure became bigger to add more metadata and to think about metadata and in a new system, we can do it now. And we are thinking about it, but then you have the everything is open now, so people say "oh oh, we can add lots of things" and other people think "why should we do that," so we have lots of discussion. And then, well, it in my opinion you should have the discussion why are you using it, can you explain why you put that information there and if you can't find a reason, why should you add it?

On the one hand, I reject the deterministic notion that just because it is possible to do something, it is desirable to do something. On the other hand, I argue that there is tension between what it is possible, even desirable, to do with digital technologies and what people in certain professions are accustomed to doing. Librarians and archivists, for example, are

[33] This statement occasioned looks of consternation and/or reprobation from her fellow OAIS authors in the room at the time; it is notable that she was the sole archivist and the other three authors were all in space data.

in highly codified professions with rich histories that support standards of practice. These exist in tension with what can be done with a technology like OAIS. This same library professional above said, in response to my question about Carmen's red coat:

> …yeah well, I did eighteenth-century literature for my studies and went to archives and libraries and read books and records in the archives and what you explain about the red coat wasn't there, I need to find that for myself, so how far do you go by giving information to your public or what do you expect them to know themselves, and that's, well we haven't discussed this here…. We have some old material, from the Middle Ages, we have pieces of books, not the book but just fragments of it, we don't give an explanation. The people who are doing research on that already know it's a fragment, so I'm not so sure how far you need to go.

A research and development officer at a European national archive described an ongoing project with Wikipedia that involves the use of the archive's data and allows for members of the default Designated Community to "crowdsource" additional descriptive metadata. He mentions a similar ethos:

> They add categories to [articles] so it's better. It's really enriching on both sides. We don't have a way… [to] bring back that sort of knowledge. Yes. [Capturing and preserving this crowd sourced data] is something that we could… do. There are some tools being developed but it's still a new thing and also it's a consideration. Like, do we really need all that knowledge? Is this knowledge that's beyond the scope of our media historical expertise?

A digital preservation and curation scholar at a US public university expressed a need for balance, like the OAIS author/archivist, saying: "And so, I don't know. I think the real design

challenge there is figuring out where the sweet spot is for some, you know, sort of initial knowledge that you can convey without too much additional work that will go a long way, right?" This suggests that there is the desire to do more with technology, but real concerns about resources have to be factored into decisions about what to preserve.

Yet this last interview subject also said, tellingly, during the course of the same interview: "I mean, if somebody is writing a history dissertation in the future, you can't expect every interpretive element to be handed to them. That's part of what they are trying to figure out, right?" There are two things at work here. First, there is the lack of standards or even conventions governing how to go about making boundaries around expanded context to fill in a Knowledge Base precisely because this is not something archives have done in the past. The digital preservation and curation scholar at a US public university said: "So, it's like you look at the conventions…, right? You have a scope and content, you have a biographical node and things. But it has to have an end to it, right? Otherwise, you would be explaining every item which is not something that we do, right?" There is the real anxiety over trying to do too much that is informed by real constraints on resources within most memory institutions. Yet the second anxiety is bound up with discourses about what it is exactly that archives are meant to do.

There are two distinct sets of discourses at play here. First, there are clear discourses about the traditional role of an archive or memory-preserving institution. The idea is that objects are preserved so that they can be examined and studied by users, but this is the limit of the role of the archive. The archive, and those who do the labor and record keeping, are not supposed to engage in what is considered to be the interpretive work of seeking out and defining the boundaries of contextualizing information. I suggest

that this is related to dominant discourses about the neutrality of archives and libraries, because rendering decision-making processes about which articles to link to within an AIP inherently means that a practitioner has made the subjective, or perhaps situated, decision of deciding where the edges of sufficient context lie. The second discourse at play has to do with "doing all the work" or "giving it all away." What is interesting in the quotes above is that both the US university professor and the European state library administrator assume that those at risk of having too much "handed to them" are researchers in pursuit of the creation of some imaginary brand-new scholastic production. Such researchers are a rarified and elite Designated Community: if the basic Information Package at a state library does not make information independently understandable to a doctoral candidate, for example, what does this say of the ability of nonacademic audiences to access and, separately, use this same content? Such an attitude inadvertently perpetuates elitist practices within memory institutions that limit usability under the umbrella of providing general access.

In asking the OAIS authors about this particular scenario with Carmen's coat, the European OAIS author with origins in space science suggested that there is a difference between an institution making content independently understandable to all of its users and presenting the content of the Archive in a way that these users can do "everything with it." That is to say, it is sufficient for users to be able to simply view content and know what they are looking at, even if they cannot participate in it, analyze it, remix it, etc. The example is easy to see within a science context, befitting this author's experience: it is sufficient for researchers to see the resulting tables if this is what the Archive has preserved; the onus is not on the Archive to provide the user with everything needed to perform the experiment

afresh. Another OAIS author, a consultant from the US, gave the analogy of a religious text: an archive can preserve the textual content of a religious treatise, and when it is presented in a readable format in the right language to a user, they should be able to read it. He said, "… [We] need to know ASCII, for example, to be able to read the content. But who *really* understands [a religious text]?" The notion of using versus understanding versus "being able to do everything" with an object is one that is not explicitly defined within OAIS. Because it is not, it is a space wherein other dominant discourses will come into play, in this case those from archival and library professions about the traditional role of practitioners in these professions.

Another discourse bound together with popular conceptions of archives is about control. Control is an archival discourse as well as part of the discourse about standards and the purpose of standards. Standardization represents a desire to exert control over a body of heterogeneous actors—in fact, participants identify this as part of the reason why they adopt OAIS. OAIS Archives want to participate in this control: adoption is purchase on a measure of control over whatever industry activities are covered by a standard. The need for control was expressed explicitly by European interviewees who were aiming, eventually, for official audits as an internal measure and as required by national and pan-European formal and informal requirements. This topic was tempered in US environments: while practitioners there do not mention the word *control* so explicitly or often, they simultaneously look to standards like OAIS and audit standards like ISO 16363 as tools for internal improvement or loosely benchmarking internal processes.[34]

[34] Although it should be noted, as pointed out by an interview subject who is a Canada-based digital preservation scholar, that there is nothing specific enough in any of the standards to act as benchmarks. The authors themselves do not speak in this way, as they

OAIS borrows some terms and standards of practice from archival science. This imbues this newer discipline with some of the power vested in archives: simply using this word connotes longevity, neutrality, control, and trustworthiness. Yet while digital archives build on the foundations of traditional archives, the technologies that undergird them simultaneously challenge longstanding practices and beliefs. Apart from the disruption caused by the changes in workflow wrought by the volume and difficulty of digital materials, there exist tensions between what archives can do with technologies and self-conceived notions of the neutral role of the archivist. These discourses, as much as what is scripted into OAIS itself, are part of what dissuade the use of this technology in any radical or recuperative fashion.

4.3 European Union and United States

Over the course of this research project, data emerged to suggest that practitioners in the Europe and the US have different attitudes towards OAIS that are generalizable based on geographic location. This section examines some of the differences that became apparent between the two regions; how the participants themselves explain these differences; and the ways in which the socio-political environment in each place helps shape responses to OAIS.

My interviews took place largely in the US and Europe. They included participants from the UK, Italy, the Netherlands, Canada, New Zealand, and the US. Because of travel requirements, the European interviews happened in a short space of time at the outset of the project; the North American interviews spanned a period of six months and happened

---

come from similar backgrounds to the interview participant, who described the reticence of engineers and systems designers to speak about a specificity which they cannot claim.

after the conclusion of the European interviews. I decided to get perspectives from both of these geographic locations during the design phase of this project. I chose the locations because the authors of OAIS hail largely from these places and, because of this, institutions in these places have the longest experience with OAIS.[35] The country where I started my research in Europe is one of the earliest adopters of OAIS in its national memory institutions; this was noted by interview participants there and confirmed by a bibliometric study of existing literature on OAIS, dates of publications, and countries of origin. This country also has institutions dealing with preservation of a number of different format and content areas: audiovisual materials, library materials, art materials, web content, and scholarly data. I did not begin the interview part of this research project with any preconceived hypothesis about geographical differences in attitudes towards OAIS or policies/politics of adoption. I also did not pick the sites to function explicitly as geographic comparisons: instead, I chose interview subjects and locations to cover a variety of roles played within institutions working on preservation and a variety of levels of maturity in terms of preservation work and policies.

Yet a noticeable finding in this research was the stark difference between adoption and attitudes in the US and Europe. These became apparent as soon as I began my Phase 3 interviews in the US. In European interviews, OAIS was either stated to be or could be assumed to be very central to preservation planning. Institutions built explicit workflows around it and used custom commercial software that was build atop the OAIS information model. One interview participant, a manager of digital preservation at a European national

[35] I acknowledge that I have largely left out the Australian contingency here and also acknowledge the excellent and mature preservation programs they have in place and the contributions they have made to standards that complement OAIS.

library, noted that the program her institution uses was used by another national library in another European country. Three other interview participants at a single institution, a national archive, explained that their institution decided to reorganize their departments and processes to map directly to OAIS, so that all departments would now fall under OAIS functional entities. One of them, technology officer at a European national archive, explained this process in part:

> [For] our reorganization… we have built the new digital archive structure on OAIS so at the ingest process you have the storage process and data management and then you have the access process. … [T]he functions of the people working here, they are really embedded into these processes so but still it's theory. This organization will start … and people [will] really change their way of working because now they are part of the ingest and now they are part of the storage…. I think also a very important aspect of introducing OAIS in an organization like this that has been working and archiving for several decennia you need a lot of communication you really need a lot of, well, knowledge transfer. Otherwise people will not understand where they are part of. So what we have done the last couple of years we did a project to sort of formulate all the requirements we need on all these levels in order to implement OAIS. So that means we make instructions for setting up submission agreements and owner agreements we set up quality criteria, you describe your designated communities…

While this level of literal adoption was surprising even to some European authors of OAIS,[36] European practitioners seemed to take OAIS much more seriously than their US counterparts do now.

I say "now" to suggest that there is a temporal element to this: perhaps it was the case that when OAIS was newer, institutions in the US hewed to the standard more literally. This was partly because it provided guidance for institutions that were really looking for something to help them grapple with digital content in the era before major standards and projects in this area. One participant, a digital development director in a US public library who has worked in a number of US preservation-oriented institutions and projects, noted:

> So… [when] OAIS was still new enough that it merited significant explanatory work on our part as managers to staff and executives, and I would assume it would be safe to say, did not quite have the ubiquity of a place in the technical service layers that it does now, like of— of it being at least something that one could take for granted as an element of discussion. I think at the time…. There was quite a bit of excitement at the model. The effort was to incorporate it, because it was seen as the cool kid on the block.

Many interviewees mentioned Priscilla Caplan and her work on the Dark Archive in the Sunshine State (DAITSS), the Florida digital dark archive. One US participant, a digital preservation and curation scholar at a public university, described this as among the more literal interpretations of OAIS saying, "So, DAITSS, for example, right? When Priscilla

[36] Of the OAIS authors I spoke to, three were European and one was American. This was perhaps not surprising given that these interviews took place in Europe and that many of these authors are involved in consortial digital preservation projects that span Europe and take OAIS as a founding assumption.

Caplan was doing her work… in Florida, I think she took a much more direct—like let's map the OAIS to our architecture approach—than most initiatives that you see in the US." As time has passed, the atmosphere that engendered these literal deployments has faded in the US. More than one US participant told me they do not take OAIS too seriously, but rather consider it background noise, use it for the common terminology, and, as three of my younger interview subjects noted, reference it because it is what is taught in library or archive school. In the US interviews, there were considerably more references to off-the-shelf tools that were built based on OAIS, such as Archivematica. One archive had been using Archivematica for a fair amount of time relative to its development, while a couple of other institutions were just rolling it out for new initiatives. In one case, a major private grant funded the roll-out project; in the other, it was the start of that institution's first systematized digital preservation efforts. One participant said, "…it just so happened that Archivematica, it is built to very much like reflect that OAIS model…." The relationship between OAIS and Archivematica is not accidental and this quote connotes the role that OAIS has begun to take in US contexts: it has become not only black-boxed, referring to something that is accepted without people peeking too much under the hood,[37] but also has assumed a place of even greater privilege: dominant invisibility.

Dominant invisibility refers to traits that come to be assumed because of their dominant roles: they do not have to be stated because they are always assumed to be in play. If a gender is not specifically ascribed to someone in positions of relative power, they

[37] More than one person complained that people do not really understand OAIS, and one OAIS author and one OAIS practitioner stated that this is largely because people do not read past "the first 80 pages" of the model and rely on a simplistic understandings of "THAT diagram."

are assumed male. If an ethnicity is not described, the assumption is Whiteness. OAIS has begun to move into this kind of space within digital preservation: it is becoming an invisible discourse. Of course it did not "just so happen" that Archivematica maps to OAIS: in fact, it was designed specifically to do that. And to the credit of this particular interview subject, he knew this. But his choice of language is telling about the role that OAIS takes in US contexts.

As this variance between EU and US perceptions became apparent in interviews, I asked subsequent participants to speculate about its cause. I received a number of answers, and will discuss them below according to the following themes, which are not mutually exclusive: compliance, cost realities, and audits.

4.3.1 Compliance

The notion of OAIS compliance is inherently complicated by the fact that OAIS is a reference model. OAIS authors therefore claim that it is without prescription: if it does not tell institutions what to do (or what not to do), then it is hard to see how institutions could be said to comply, if compliance is defined as adequately adhering to a set of prescriptions. Yet compliance with OAIS was a topic that arose frequently in interviews, despite the fact that I never asked any participant whether they were compliant or wanted to be.

In a European setting, compliance did not arise as a separate topic of conversation. Rather, the discussions focused more on successfully meeting the requirements for a full audit. These discussions were very technical in nature, and addressed processes, workflows, and research that the institutions were doing to ready themselves for an audit. At one institution, the technology officer kindly gave me copies of a white paper detailing workflows of their OAIS implementation: such documentation is an essential step for an

audit. She also described other internal documentation efforts that were underway, such as the creation of a controlled vocabulary specific to the domain of the institution.

In the US, however, compliance with OAIS was often mentioned as something separate from auditing. Instead, it was related to general imaginaries about doing what OAIS requires: I say imaginaries because notions of compliance were very institutionally specific and sometimes people suggested they were doing things to comply with OAIS that OAIS does not actually stipulate. As mentioned previously, OAIS only "requires" six things, and much of its functional model is entirely voluntary. A manager of digital preservation at a European national library and a European OAIS author with origins in space science both suggested that most people do not read OAIS closely enough or all the way through. This lack of application or attention may be partly to blame for the anecdotes about even headless barnyard animals being compliant, much to the annoyance of the OAIS author and the preservation manager, the latter of whom was troubled by people "bashing" OAIS on Twitter when it appeared that they did not adequately understand the standard.

Additionally, compliance in US interviews echoed the more casual attitude US practitioners espouse towards OAIS. A US-based data curation specialist suggested, "So, OAIS does have this concept of packaging your data, and… we may not be using an OAIS compliant packaging formats approach, whatever that means…." A digital archivist at a small liberal arts college in the US said, "I feel like I've kind of done that in my mind, like casually I've been like oh yeah, I like use OAIS, but I probably am not conforming to the actual standard like of the way it's been written, like the big blocks in my head, I'm like those are useful." The director of digital services in the library of a private US university said, "So, we're pretty—the repository where we do OAIS is pretty compliant, and so we do

it. But, again, we're not—I don't think we're that obsessed about it, and [rethinking our next repository iteration], we're not thinking that hard about it." These quotations sum up the casual attitude towards OAIS. They also confirm, to some extent, the concern of the European practitioners who suggested that many institutions and professionals are not reading the standard carefully or referring to the actual documentation of the standard. Instead, they build on the general ideas they have of OAIS, of the black-boxed popular image of what they think OAIS represents. This is not a criticism, but rather speaks to an evolution of the role that OAIS plays within preservation networks in the US: it has become a foundational assumption that, because decentered, is not often closely examined.

4.3.2 Cost Realities

When asking participants to speculate about the causes of the different attitudes between the US and Europe, the most recurrent theme among these explanations boils down to the cost realities of developments in digital preservation and the relationship this has to institution-specific projects. A Canada-based digital preservation scholar who works frequently in Europe as well said, "And one guess might be the dynamics of the large scale research project that happened. In the last 10 years, a lot of the research in Europe, that was highly visible, was done in those large integrated projects." A digital preservation and curation scholar at a US public university said:

> This [centralization of work and reliance on standards] also percolates through research funding, right?... So, many countries have depended a lot on European Commission funding that has these high level frameworks that everybody has to fit into. And even at their national level, then they often have a relatively centralized research council that determines funding for research within that country. The US,

> on every one of those fronts is just more diverse, right? We have a huge number of university based archives, and libraries that are in many ways setting a lot of the trends for the research and development. You have a much larger diversity of funding sources, you have no one central national library archive that's basically setting the expectations for everybody else to follow.

This lack of centralization in the US was seen as another cause for the more relaxed attitudes towards OAIS. Many US participants described the Ford, Mellon, Grammy Foundation, and federal grants that they received as individual institutions to work on institution-specific projects. In the US, there is no centralized requirement to follow certain digital preservation standards and there is less centralized funding to informally encourage standard participation. This de-centralization was often tied back to the difference in the standings of national memory institutions in the US versus Europe: participants from both Europe and the US reminded me frequently that the US has no national library. Interviewees described institutional partnerships and related funding opportunities as much more centralized in Europe. A digital preservation manager at a private US university said:

> So, basically everyone… seems to be in research over here, where [in Europe researchers comes from] a national organization, or a public sector organization. [They have to] say that this is an obvious need that's going to be as a society, you have to make some sort of argument [to] someone who wants to give you money.

Work in the US is dispersed to academic research institutions, where projects are often specific to a particular collection or set of collections at the institutions receiving the grant.

In Europe, national memory institutions and other professional organizations are involved in large consortia of projects for the development of digital preservation standards. A US-based data curation specialist similarly said,

But they have the EU, they have places where a large number, or a big percentage, a majority of the funders can get together and basically say, "We're going to go about doing it this way." So, that's an area that has national trends, national viewpoint, that simply doesn't exist in the United States.

Another data scientist put it more starkly: not only is the money more centralized, but this imbues the funding agencies with power to dictate the paths of future preservation projects. Further, the funding opportunities are greater in Europe than in the US, which leads to more institutional development and adherence which are not possible given the smaller scales of US funding:

> Yeah, certainly from—at the government level, for sure, and even within the funding agencies, the research councils in the UK for example. Some of them actually have very strict policies that allow them to shut down funding for an entire institution…. I would say that, you know, the amount of funding that Europeans are putting aside for managing data is substantially greater, of what the US is doing.

These findings are mirrored in language from European interview participants. A manager of digital preservation at a European national library noted their institution's participation in a "European project about costs." A research and development officer at a European national archive described a number of projects wherein his institution worked with all the national memory organizations in his country as well as numerous other preservation hubs.

The centralization of funding and thus power to dictate procedures for preservation has other effects too. The next section touches on the different geographical attitudes towards official audit processes. Audits are expensive: an OAIS author with a background in space science estimated around $10,000 per year to maintain the certification and a higher layout for the initial, in-depth audit. A digital preservation manager at a private US university shared that his institution was prepared to budget $50,000 over five years for auditing. The cost is prohibitive to many organizations. One interview participant, a Canada-based digital preservation scholar, suggested that no small institution has the resources to undergo official audits. US practitioners similarly pointed out the cost and other resources entailed in an official audit as part of why they were not going to engage in one. More than one participant noted that European funding agencies are moving towards requiring certification for institutions seeking funding, meaning that European institutions are more likely to budget for costly auditing.

4.3.3 Audits

In the realm of digital preservation, there exist a number of audit and risk management tools designed to help institutions assess strengths and weaknesses as well as put forward public documentation that certifies their trustworthiness. In terms of internal assessment, there are tools like the "largely defunct"[38] Digital Repository Audit Method Based On Risk Assessment (DRAMBORA), which is based on risk assessment modeling and provides a framework via a web tool or paper checklists for institutions to identify and categorize risks. For public certification, there are protocols like the NESTOR seal of

[38] So described by interviewees.

approval based on DIN, the German standard for long-term preservation, and its Dutch counterpart, the Data Seal of Approval. The most commonly referenced audit standards in my interviews were TRAC (OCLC/CLR, 2007) and its successor, ISO 16363 (CCSDS, 2011). Audit standards like the Data Seal of Approval, TRAC, and ISO 16363 are all explicitly based on OAIS. This means that undergoing an ISO 16363 audit requires a mapping of institutional functions into OAIS language even if systems were not designed this way. I argue that ISO 16363 has served to reinforce and recreate OAIS's role within a larger sociotechnical network for digital preservation. Where OAIS started to fade into the background in some settings, ISO 16363 serves to thrust it forward again. This move reinforces the dominance of the scripts within OAIS as they now require levels of adherence to gain certification. While OAIS as a reference model is not prescriptive, ISO 16363 is entirely so: it details the myriad things an Archive has to do in order to be certified as trustworthy. This does not boil down to specifics of workflow descriptions or best format recommendations, but rather serves to keep the higher-level notions about the relationships among the parts of an Information Package and the roles of various actors within an institution intact. This is especially true in a European context, where many institutions want certification and where larger, pan-European institutions are pushing for certification as well. In speaking to professionals at three state memory institutions in Europe, all spoke of working towards eventual ISO 16363 certification. None felt they were ready as of fall 2014.

The US attitude towards audits, similar to US attitudes about OAIS more generally, is less formal. While one participant described his institution's financial commitment to submit to an ISO 16363 audit by an outside body after having spent some time with a

consulting firm preparing materials, most US participants expressed their use of ISO 16363 and even its predecessor TRAC as a tool for internal use rather than public certification. A digital repository manager in a US museum summed up this ethos, saying, “we need to know is are we doing our job right and that doesn’t require official certification but that standard was very useful for understanding where we’re falling short.”

In these cases, the use value is described as something to help the institution measure the status of its programs in order to identify strengths and weaknesses. This kind of work plays the dual role of providing a platform on which to build conversations with non–preservation people inside institutions: it helps explain and in some cases justify the work the preservation people are doing. These types of conversations are, not exclusively, but often strongly linked to discussions about funding. In terms of strengths and weaknesses, the need to document policy was a common struggle, as identified by pilot projects for ISO 16363 itself and the work of Sierman (2014) which noted the disconnect between policy and practice, and vice-versa, in Archives. A digital repository manager in a US museum noted sagely that audits would aid the creation of essential documentation saying, “So, if I get hit by a bus tomorrow and any colleagues need to get X done, they can just look at this and know who they need to talk to.”

A US-based data curation specialist also pointed to limitations in what is purchased by the audit process, saying:

> I think we're probably landing in a similar place to other kind of certifications, which is— and I don't mean this to be disrespectful in any way. There's almost a mechanical aspect of something like 16363 that it's a checklist, it really is a checklist. And if somebody were to say, "Do you have secure access to your data center?" And

> you say "Yes," you know, you could potentially get that checked off your list. So, what does that really mean? You know, does that mean that there is a card swipe system to your datacenter? Does it mean that there's a key? How many people have that key? How secure is the card swipe system, so on, and so on, and so on…. So, I think something like the TRAC audit checklist addresses those mechanical pieces, sort of the necessary conditions, but not necessarily sufficient conditions.

At the end of the day, most US practitioners eschewed the need to undergo official audits, even if some of them rather wistfully explained that they would like to be auditable at some point; no one felt that they were ready to be audited currently. Instead, as a digital library director at a private US university explained, "If you can get your content together and disseminate it, that's proof that you had [laughter] good archival practices."

These examples highlight the differences in dominant concerns among preservationists in Europe and in the US. Processes in Europe are much more centralized and much more standardized. This can be seen as part and parcel of politics of pan-European identity espoused by the EU. These politics of governmentality trickle down into the individual memory institutions as they are swept up into a project of creating a unified European identity.

US concerns revolved around moving beyond OAIS and the future of digital preservation. The general assessment is that OAIS has been useful, but that perhaps its heyday has past. Radical changes are needed, if not new standards entirely and for now, practitioners adhere to surface elements of OAIS as opposed to engaging in literal deployments of its functional or information models.

4.4 Conclusions

This chapter has examined many findings from the interview data under three overarching themes: audiences, the role of traditional archives, and geographical difference. This chapter identified scripts embedded into OAIS that stem from its positivist origins as well some of the discourses in which it participates: for example, discourses of archival neutrality or of unified pan-European identity. Following these general themes, the next chapter will examine a single term in close detail: the *Designated Community*. By examining this particular aspect of OAIS more closely, I will argue that there are ways in which OAIS fundamentally does not work, but simultaneously contains the potential to act as a recuperative and resistive technology, with the power to democratize archival practices.

CHAPTER 5

DESIGNATED COMMUNITIES

Introduction

The previous chapter looked at some general themes that arose across multiple codes in the interview data. This chapter provides an in-depth look at one particular term. One site that exemplifies the tensions brought about by the deployment and general reliance on OAIS is the conversation around the term *Designated Community*. Designated Communities are manifestations of a particular kind of audience imagined by OAIS, not for the model itself but the content OAIS will be used to preserve. Designated Communities are constructed end users that are more often simulacra (Baudrillard and Glaser, 1994) or personas (Miaskiewicz, Sumner, and Kozar, 2008; Lage, Losoff, and Maness, 2011) rather than real people. This term is a source of frustration to some in digital preservation, particularly librarians who find the need to specify a particular community to be at odds with their professional, and sometimes legal, mandate to serve broad populations. The authors of OAIS, in response, insist that the requirements that are built around Designated Communities are meant specifically for use by Archives that serve diverse publics or, in some cases, "the world." The authors' conception of the term's use value to memory practitioners who deal with large audiences and popular material is based on a set of assumptions stemming from an epistemological space that results in a mismatch between what OAIS recommends and requires and how some memory professionals see their work-related duties.

In this chapter, I examine the definition of Designated Communities and the ways in which this term arose in interviews with OAIS practitioners, scholars, and authors. The first

section serves to expose the characterizations and concerns users express with the requirements of OAIS as dictated through this very central term. The section that follows uses two theoretical frameworks, scripts and discourse analysis, to examine this subset of data. I argue that Designated Communities are fundamentally at odds with the professional ethos of certain memory professions. I also contend that this term is the source of considerable power for the Archive, in a way that privileges the institutions more so than the publics and audiences they serve; simultaneously, I discuss how this term is also a site where OAIS can be used to make interventions in traditional memory practices that allow the co-construction of more inclusive and pluralistic histories.

5.1 Defining Designated Communities

Designated Communities are defined in the OAIS documentation 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"[39] (CCSDS, 2012, page 1-11). This definition is packed with a considerable amount of tacit information. Firstly, the word *potential* denotes the fact that these Consumers are not necessarily real. Instead, "potential" immediately indicates that Designated Communities refer to a particular imaginary about the kind of Consumer who will use the content stored

[39] For clarity, I include here the definitions for the 2 specific terms mentioned within this definition:

- Consumer: The role played by those persons, or client systems, who interact with OAIS services to find preserved information of interest and to access that information in detail. This can include other OAISes, as well as internal OAIS persons or systems. (Ibid page 1-10)
- Archive: An organization that intends to preserve information for access and use by a Designated Community. (Ibid page 1-9)

by an Archive in the future. In other words, Designated Communities are often entirely conjectural. They could refer to something as concretely constructed as personas. Such a formal construction proves tedious and frustrating to some users. A participant who is a digital preservation and curation scholar at a US public university said: "…if you take it so literally that you think that it's almost like doing, you know, like personas when you're doing interface design, where you have to characterize exactly who these people are, and like what they eat for breakfast and everything, you know. It might be frustrating, because you're thinking, 'I can't specify my users that clearly.'" On the other hand, the use value to such a process is helpful when projecting future activities and needs. Later in the same conversation, the same participant continued, "…and I do think that that's why a designated community notion can be useful. And again—I mean, maybe it actually—I was dismissive of it before, but maybe it even is something kind of like personas and human factors work, right, where it's like you just think—as a heuristic, you think, 'Okay, 20 years from now, what are the main things that if somebody didn't know them, they would be in a complete loss, you know.'"

Designated Communities might, on the other hand, be a simulacrum derived from qualities imagined by the managers of an Archive, some hybrid constructed of an understanding of Producers and Consumers in addition to assumptions held by the Archive managers themselves (Akrich, 1995). The important thing is that Designated Communities are rarely, if ever, a reference to real people in real time. A single Designated Community could encompass a number of different user groups, as Kari Kraus, a Preserving Virtual Worlds II (PVWII) investigator and digital preservation and English scholar at a public US university, noted. As business cases, audits, and maturity models are derived from the OAIS

reference model, it is possible for an Archive to have multiple Designated Communities for discrete collections. Findings from the APARSEN project mention additional user groups in relation to expanded services and sustainable business models: Archives can seek to promote their relevance and increase or ensure revenue by catering services curatorially to different groups of users by expanding the Representation Information[40] of objects (APARSEN, 2015). In this APARSEN deliverable, the authors move away from Designated Communities to users. This terminology is significant, given the presence of OAIS specialists within the APARSEN project: the curation work that expands the usability of content to more Consumers, now finally termed *users*, is thus indicated to be a separate or additional process to core preservation activities because the authors mention users, who are purposely not part of the OAIS model.

Two additional things about the definition of Designated Communities: first, the definitions of Designated Community and Archive within OAIS are circuitous. The definition of the former notes that Designated Communities are defined by the Archive; the definition of the Archive is "an organization that preserves and makes accessible and usable content to a Designated Community" (CCSDS, 2012 1-9). In other words, the Archive defines the Designated Community and is similarly defined by it. This places a considerable amount of control in the hands of the Archive, who is able and required to make this

[40] "**Representation Information**: The information that maps a Data Object into more meaningful concepts. An example of Representation Information for a bit sequence which is a FITS file might consist of the FITS standard which defines the format plus a dictionary which defines the meaning in the file of keywords which are not part of the standard. Another example is JPEG software which is used to render a JPEG file; rendering the JPEG file as bits is not very meaningful to humans but the software, which embodies an understanding of the JPEG standard, maps the bits into pixels which can then be rendered as an image for human viewing." (CCSDS, 2012, page 1-14 - 1-15)

designation, and leaves relatively little control for actual end users as there are no recommendations within OAIS that suggest consulting with user groups to make the designations: the third of the six requirement mentioned earlier says only that an Archive must "determine, either by itself or in conjunction with other [unspecified] parties" who the Designated Community is (CCSDS, 2012 page 3-1, insertion my own). This is important in light of the second item of note: that Designated Communities can change over time. "Change over time" gives an organization attempting to build an OAIS-based repository or to use OAIS-based audit models to obtain certification of trustworthiness latitude to modify its Designated Community in a way that matches the work it is already doing. This is not a pejorative statement about activities or practices of the Archive: while this could refer to organizations that change their documented Designated Community to avoid workflow changes, for instance,[41] the intent as expressed by OAIS authors matches concerns that also find expression in the first of the core OAIS requirements: that archives "negotiate for and accept appropriate information from information Producers" (CCSDS, 2012, page 3-1). Digital preservation can be complicated. Several interview participants note that they do not have criteria about which formats to accept. Some say their concerns with this are occupying unnecessary thought and time on the part of practitioners and others express that they are not allowed by their organization to discriminate against content in that way.

[41] This possibility was acknowledged frequently during conversations with OAIS authors within the specific context of talking about documentation of policy and practice by an institution for audit purposes: authors I spoke with frequently noted that Archives "can do whatever they want" and still be technically certified, but that this probably was not in their best interest. There was an explicit recognition that merely being certified as a trustworthy digital repository according to a scheme like TRAC or the more recent ISO 16363 standard does not guarantee that the Designated Community or other stakeholder groups actually trust and would chose to work within said organization.

On the subject of formats, one participant, a digital preservation manager at a private US university, made an apt analogy: paper could be considered a single format, and conservators have devoted copious resources and time to the methods and tools that ensure its long-term persistence. Spending similar amounts of time on each digital format is exponentially taxing. The authors of OAIS recognized this, and continue to be keenly aware of this difficulty as they make periodic updates to OAIS and write additional standards like ISO 16363. Thus this type of language occurs throughout OAIS explicitly and implicitly, conceived to give organizations codified license to reject certain types of content or formats, or to treat different content and formats differently within daily workflows. The ability to change a Designed Community can be seen as part of a larger discourse within archives relating to an expressed need for control to manage a gargantuan task; control and the inherent complexity of digital preservation work are active discourses themselves within the field that drive institutional decisions as well as ongoing choices OAIS's authors make about how it evolves.

5.2 Designated Communities in Interviews

The topic of Designated Communities fascinated me before the outset of this project in relation to two issues: first, the difficulties posed by serving very broad audiences and second, the fundamental assumption within digital preservation that preservation is for someone, rather than simply a set of activities performed on a thing. In this chapter, I use terms like *complicated digital objects* and *very broad user bases*. I do not signify a dichotomy with this language, nor do I assume that OAIS was only designed for comparably simpler cases. As OAIS was written and continues to be reimagined by its authors and primary proponents, it is a tool that is applicable in any discipline with [almost] any kind of content.

I say “almost” because of an admission from an OAIS author I spoke with that linked data is something that they are still working to include. That is the only content type I have heard OAIS authors admit might not fit with OAIS. Otherwise, they argue strenuously for its flexibility and usefulness in all other situations with all other content types. The terms I employ here are inherently relational, however, and used commonly by practitioners in everyday speech and it is worthwhile examining what they mean. The development and maintenance of OAIS was spearheaded by scientists dealing with space data, and the assumption that arose in conversation, both formal interviews for this dissertation and more casually with colleagues, is that OAIS works best for science data because it was designed by scientists. In this case, the Designated Communities are not broad or heterogeneous because the materials are not popular. Rather, a Designated Community is a specific set of scientists who share a Knowledge Base of domain-specific knowledge and in this case the Designated Community can be seen to generally align with a large portion of the projected Consumers: the people most likely to want access to the digital objects are fellow scientists within the same discipline. This contrasts in some ways with popular content like video games wherein there are many end Consumers who interact with the content for a large variety of reasons. A Designated Community for video games as conceived of by an Archive may include only a small subset of actual Consumers or users of the original content. Likewise, when dealing with science data, the example often given is of data tables: in this case, the object consists largely of text or perhaps still images and is largely static. This contrasts with the myriad components that make up a virtual world in a video game, including particular peripherals and hardware required for interaction; specialized copyright-protected media; dynamic input by multiple users, etc. While there

are certainly kinds of science data for which the Designated Community would be broader than a specific set of domain colleagues and science data that is more complicated than text, I argue that video games can be safely said to constitute an outlier, difficult case by comparison.

What happens when an Archive has to preserve popular culture content for which the audience might conceivably be "the world," as two interview participants said? A manager of digital preservation at a European national library and I had this exchange:

RB: For a video game series like that, the Designated Community is…

Participant: The world [*laughs*] and that's the problem for libraries.

A US-based data scientist also used this same language in reference to the communities served by university consortia data repositories. My previous work with the preservation of video games, which have incredibly large and heterogeneous user bases,[42] highlighted this challenge and brought to the fore the way in which this notion is bound together with the assumption of preservation for someone rather than preservation of something. One interview participant, a US-based data curation specialist, summed up this central tenet of digital preservation: "the concept of a designated community is also an important one…who you're actually trying to preserve your content for is equally important in some sense to what you're trying to do with the content." This quote serves to show how tightly the core terms of OAIS are bound to the foundational principles of digital preservation as a professional practice and a sub-discipline.

[42] See, for example, the number of copies of a new video game title that sell in week one for a popular series like Halo or Grand Theft Auto, and compare this with movie box office numbers for the same week.

The idea of preservation for someone plays out in a number of ways within OAIS-informed archives and here I will speak to how it is inextricably linked to the notion of Designated Communities. Henry Lowood, one of the investigators from the PVWII research project who is a game preservation scholar from a private US university, said during the final grant meeting: "I need to speak up for my designated user community. There are people who are more interested in the source code than they are in the game. In my environment, technological historians are a pretty important stakeholder." Examining significance of video games for preservation necessarily invokes discussions of Designated Communities because significance is situational: it depends on the audience and their relationship to the digital object. This in turn determines what aspects of the digital object are given precedence for resources: with very complicated digital objects like video games, the digital preservation community acknowledges that some concessions will have to be made over the long term and some aspects may not be preservable indefinitely. The quote above from Lowood highlights the tensions that arise from this. While game producers interviewed for PVWII expressed that the affective play experience was more important that the underlying code, Lowood notes that for a different Designated Community, the original and executable code remained of primary interest. The point is simply that different people want to save different things, or, in OAIS speak, different Designated Communities require different Archives and different Archives create different Designated Communities through their daily actions, despite what they may or may not put on paper about who is in their Designated Community.

It is also interesting to note the language used by this game preservation scholar: he uses the term "*designated user community*." First, he is employing terminology that stems

from the common language of OAIS, despite the fact that OAIS was not an explicit part of the conversation in which this comment was made, or indeed a focus of the multi-year grant project at all. This is yet another demonstration of the ubiquity of the OAIS reference model and how it is often invoked unconsciously. Yet this phrase also demonstrates a misconception among more casual users of OAIS.[43] The investigator did not say "Designated Community" but rather "designated user community." I confess to speaking in this same way myself prior to commencing serious research and study on this topic: as a casual user of OAIS, a practitioner in digital preservation who did not deal with OAIS structures daily, I often conflated the term Designated Community with the concept of users and in fact used this same linguistic hybrid of the two until I was sharply admonished by a European OAIS author with origins in space science. Designated Communities are importantly not users, and the OAIS documentation stays away from language about users in reference to Consumers. This is a purposeful move as described by the authors of OAIS. I asked OAIS authors to explain the concept of Designated Communities further in particular reference to repeated comments by library and national archive professionals about the ways in which the Designated Community requirement felt restrictive compared to their remit. The authors of OAIS employ the term *Designated Community* instead of *users* because they do not want to speak about real users for two main reasons. First, OAIS authors do this to avoid the pretense of a level of precision that is impossible to have about what is essentially, as one interview subject who is a digital preservation and curation scholar at a

[43] Although I note at the outset of this argument that, because this quote is taken from a conversation that is about digital preservation generally and not necessarily OAIS specifically, this language is not conclusively indicative of a conflation between users and Designated Communities. I think, however, that is statement is more telling than it appears at first glance and perhaps even more so than the person who said it realized.

US public university put it, a “heuristic” to help plan for dissemination needs in the future. Designated Communities are purposefully vague, therefore, much to the annoyance of some interview subjects. One subject, a Canadian digital preservation scholar, said: “[Designated Community is] also extremely weakly defined in the OAIS. There is almost nothing about it. It's there. It's there and it's supposed to work.”

The second reason Designated Communities are different from users, according to the OAIS authors, is precisely to serve institutions that deal with popular materials and broad, heterogeneous user populations. It is when dealing with complex materials and large user bases, as opposed to specific research data with smaller user bases, that it becomes necessary to make selection decisions about content and how it is treated over the long term. The example given by the European OAIS author with origins in space science was that of a national library. In that case, the institution is required to serve users who encapsulate the entire population of a country. This means from birth to death, across all levels of education, language, ability, etc. To make all digital content usable and independently understandable to all these users, the OAIS author argued, would mean that the Archive would have to include an entire education system with each object. The implication: it is not possible to do this, so for policy and planning purposes, institutions like libraries should construct a representative Designated Community that has common capabilities and understandings and create Information Packages around this, acknowledging that while such a constructed Designated Community might encompass some of the institution’s actual users, it by no means pretends to be a stand-in for all users more generally.

I found this explanation to be enlightening; that is to say, the content of it was not something that I had garnered independently from all the reading I had done about the standard. I asked some of the interview participants about this, and in fact two interview participants, a digital preservation manager at a private US university and a digital preservation consultant at a boutique US firm, were present at the exchange I had with the OAIS author. They both expressed that this explanation about why Designated Communities are not users was similarly enlightening to them. In sharing this anecdote with other colleagues and interview subjects, many expressed surprise and one colleague, a digital preservation and STS scholar at a public US university, even described this explanation as "disingenuous," perhaps suggesting that this story was constructed post facto to answer the frequent criticisms about this term. However, the same interview participant who was quoted above describing the definition of Designated Community as "weak" also said the following in response to my recounting of this story:

> So, I mean that's an obvious argument to someone with my perspective, for example. So, maybe it's so obvious that we all [in the realm of systems design] fail to communicate [the real purpose behind Designated Communities] in writing properly….There's a curious divide, sometimes, in communication between the designers, engineers, scientists, and the others, where the scientists and engineers are much more aware of the limits of what they mean when they say those things, models and systems and processors. But to those who are not scientists or engineers, it sounds—like it sounds a little scary and it sounds too strict.

The notion of a Designated Community might also be part of the moment in which OAIS was developed: early realizations about the difficulty and resources entailed in doing good

preservation work inform the fundamental notion that digital preservation is not just preservation of a thing but rather preservation of a thing for someone and, further, for some purpose. The purpose is something that gets taken up in greater detail in spaces like the APARSEN Common Vision report that suggests an expansion of dissemination services as a model for encouraging funding. It is also what current data curation practices are based on: the idea of adding value to existing content through its arrangement and delivery in a way that allows for better use according to someone and some purpose. The idea that digital preservation is a set of ongoing activities that happen with an intentional outcome beyond the analog practice of keeping things as much the same as possible is concurrent with the development of OAIS. The concept of preservation for someone is inscripted into OAIS authors in the way they describe the term *Designated Communities* as well as the antiprograms[44] constructed by practitioners engaging with the model in the real world.

Yet part of what may make the aforementioned European OAIS author's explanation seem so eye-opening to some practitioners is that, despite the reassurance contained within this origin story for the term *Designated Community*, it is a concept that is fundamentally at odds with the purposes of certain types of organizations. I had two participants describe their Designated Communities as "the world," one from a European national library and one from a consortium of US universities. When it becomes necessary to cut some users from the designation, it results in a system that may allow public access without allowing usability of digital materials. The US-based data scientist said:

---

[44] This language similarly comes from Akrich and Latour (1992): "***Antiprograms:*** All the programs of actions of actants that are in conflict with the programs chosen as the point of departure of the analysis; what is a program and what is an antiprogram is relative to the chosen observer" (page 267).

> So, if you decide that you are going to designate your community, that you are going to say that, "I'm only targeting this set of people that I'm—you know, so even though this information is available to the world, the things that I do to that information to make it usable are only targeted at this set of people, then the problem is easier. And I think, you know, in the end that's what you do, no matter what. So, yes, this is publicly accessible information, but it's not publicly usable information. And I think curators have, you know, have a hard time with that. I mean I would think curators hate that.

I suggest that most librarians are not too fond of this idea either, and this explains their consistent complaints about the Designated Community requirement in OAIS. It is complicated to align the Designated Community to the Archive's preservation tasks so that there is consistency, a fit between the overall tasks and objectives and how those tasks are carried out by an Archive. The issues of fit and Designated Community are made clearer when talking about institutions like libraries which have heterogeneous user bases and trade in popular materials. The inherent claim here, however, is not just that it is difficult to designate a community in a library, but that OAIS is misaligned with these institutions: it does not fit their social reality or the objectives of the organizations, and it does not fit the work practices of staff. Video games, as I will discuss in the next chapter, are another excellent digital preservation heuristic because of their complexity but also because of their popular appeal. Cultural productions like games enable new practices and, with their wide reach, new publics (Cvetkovich, 2003, page 10). These creations serve as a challenge to the notion of the Designated Community, because they suggest an ongoing creation of publics through the cultural productions spurred by the consumption of popular materials like

games. This represents an inherent tension and breakdown of the system. On the one hand, OAIS and those who speak for it, like the authors I interviewed, would say that the notion of the Designated Community is designed precisely to deal with this type of situation: a recognition of the ongoing forming and reforming of consuming groups that places untenable demands on the resources of digital archives that are themselves still evolving, like the cyclical and episodic nature of memory itself (Bradley, 2007). Yet this is an insufficient explanation when confronting the practical realities of implementation: it is yet another acknowledgment that OAIS as currently scripted is not functional for the needs of institutions with large user bases like public libraries or archives of culture, and it is disingenuous to claim that it is. This term *Designated Community* shows, from a sociotechnical standpoint, one of the real problems of OAIS being applied rigidly: to require Designated Communities to be separate from users is a fundamental misalignment with some organizations.

Despite the importance of this term, many interview subjects did not have the Designated Communities of their Archives formalized. One participant, the digital preservation manager at a private US university, noted that he meant to put it into the overarching policy he developed with his organization but had not yet, although he felt there was an obvious place that it would fit. One of the OAIS authors I spoke with, an archivist from Europe, described how difficult it was for many organizations to identify their Designated Community when she and her team were doing a tour of test audits ahead of the standardization of ISO 16363. In explaining what constituted sufficient proof of a Designated Community and what certainly did not ("everyone" was explicitly forbidden, as were "implicit" definitions), she said, "We didn't know that [this aspect of the audit

process] would be so hard." Another subject, a manager of digital preservation at a European national library, said of her peers in the field of digital preservation:

> And sometimes people say, well there is no need to trouble about it, because everyone is your designated community so you don't need to do extra things, compared with what you need to do with researchers, because you need to explain the formulas and et cetera, but I am not so sure…. No, I always in the TRAC discussions, in the TDR discussions, I always added my web archive as an example where the exception was because we don't have a designated community, but don't have a user either.

Designated Communities came up in relationship to audits precisely because they are required documents for audits and because, while many institutions think and assume a Designated Community, many have not yet defined one on paper. This lack of policy and documentation, as well as the mismatch between documentation and practice in archives is something that has been covered in recent research as a fairly pervasive phenomenon (Sierman, 2014), and this was recognized by the authors of ISO 16363 and ISO 16919. These authors repeatedly pointed out the need to independently assess not only what Archives say they are doing but to verify and document (on the part of auditors) what Archives are actually doing: saying and doing were of equal importance as there is a potential for considerable gulf between them in practice, and this was highlighted constantly during the training course with the OAIS authors.

The role of multiple Designated Communities also arises in reference to audit processes: it is possible and sometimes preferable to name different Designated Communities for different collections within a single institution. During an audit, this

process would be necessary especially if different Designated Communities meant that their respective collections had different workflows, AIP contents, and other such processes or facets. One interview subject, a US-based data curation specialist, noted that his diverse collections have different assumed uses and that this plays an active role in how he and his colleagues define their processes and services:

> Same questions come up of—so, if your designated community's medievalist versus geologist versus astronomer, is there some common layer in that structure where those—all those designated communities have the same requirement? My answer is yes, in our view, but again, to ask that question in that way is really important and to understand at what layer of the stack there is that kind of commonality, it's probably consistent with the OAIS…. But if you're thinking about the service a little higher up the stack, like an image viewing tool, for example, there probably are differences in some sense between what a medievalist wants to do with an annotation versus a geologist, and again, there can be some commonalities, but it's at that layer where I think you start to see the differentiation and the sifting out of different kinds of requirements.

What is key here is the notion of layering: at some layers there is a commonality of services, and at some subject specificity may dictate a need for different approaches to serve different Designated Communities. There is power in having commonalities across multiple Designated Communities: it demonstrates to non-preservation-oriented administrators and funders a poignant need for certain services in a way that a single Designated Community with a single set of needs cannot.

In relation to these layers, this same researcher also said:

> And if we think about it from that conceptual viewpoint, the question for me, or for [my institution]… has been what are you trying to accomplish, and for whom, and it's the 'for whom' part that's the designated community aspect of it. And it does depend on what layer of a management, data management stack, or for some kind of infrastructure, infrastructure stack you're talking about.

This further suggests that only at certain layers does the Designated Community play a large role in defining processes. If this is true, I would suggest the impact occurs at a very macro-level; within the functions of those tasked with detailing workflows and policies; and for those working with Consumers and Producers. However, it is worth noting that audit processes for a repository indicate something along the lines that an auditor should be able to pick any member of the Archive staff, at any layer, and ask them basic questions about the central policies of the Archive. I would posit that, given the foundational nature of the Designated Community and its role in the basic six requirements of an OAIS-informed repository, OAIS authors would intend for this information to be mentioned within such policy statements. Therefore, the Designated Community is meant to play some role in all layers of an Archive.

The idea of a stack of expertise and that the very foundational knowledge about Designated Communities is only essential for certain functions within the Archive raises a question about the relationship between Designated Communities and the compartmentalization of work. In the previous section, I discussed the ways in which OAIS and OAISes conceive of audiences and the fragmentation of labor roles in archives between analog subject specialists and digital specialists. Here, I suggest that Designated Communities also play a role in the compartmentalization of work by further entrenching a

difference between the grand Archive-wide labor and the subject-specialist, Designated Community-specific labor.

5.3 Designated Communities: Scripts and Discourses

Why are Designated Communities so important? Obviously, within the space of OAIS they are tantamount. Though not the most-mentioned aspect of OAIS in interviews, the very project of an Archive as described by OAIS is predicated on this notion: an OAIS is not functional without a Designated Community. Within discursive practice and my project of tracing power, this term designates a space of actual and potential power if for no other reason that the fact that "…[r]epresentation has consequence-- how people are represented is how they are treated" (Hall 1997; Madison, 2005). The ability to designate a group of people grants power to the archive when looking at historical politics of oppression governing group identity formation. In this instance, the group's identity is formulated and regulated by an outside body, the Archive, with no writing or scholarship as of the writing of this dissertation about the potential to include community members in this designation process. I situate a collective process of negotiating a Designated Community between an Archive and its Consumers within progressive work in archive, library, and museum practices as cited in Chapter 2. The one sidedness in group creation also lends itself to a form of imperial inclusivism, where inclusivity dictated by one group rather than co-created in partnership is likely to result in subordinate groups being ''swept into the interpretive schema of another tradition'' (Eck, 1993 in Caswell, 2013, page 282). OAIS can be used by any kind of organization wishing to do preservation, but the cost realities are such that it is an expensive project and those with functional repositories tend to be official memory organizations, often with links to public- and state-sponsored institutions. These

institutions are products of formal processes that are guided by the ideologies of dominant groups and thus guide how subjectivities are represented in the digital objects within the Archive (McCarthy, 1994). This constitutes a layer of representational discourse. An additional layer is the historically obvious fact that those not represented within Designated Communities, often nondominant or subordinate groups, are not likely to be documented within artifacts in the Archive at all, and vice-versa. Designated Communities are groups and the application of theories about group formation and representation are a fertile ground for future research.

At the same time, *Designated Communities* has the power to be among the more recuperative terms with OAIS regarding traditional discourses. The very notion of different Designated Communities creates a space with the power to resist notions of "universal human being[s], unmarked by difference" (Noble, 2012, page 81), a space in which difference can be co-constructed and preserved as difference by community members and the Archive in more inclusive, pluralistic ways. Stripped of specificity, the assumed Designated Community of many official state- and public-sponsored memory institutions, the universal subject, is in fact a representation of dominant norms of neo-liberal individualism: it is white, it is male, it is Western, etc. But the creation of thoughtful Designated Communities has the power to restore subjectivity, to allow nondominant subjects to maintain "opacity," in the language of postcolonial scholar Edouard Glissant (Britton, 1999). That is, it allows nondominant subjects to appear in official memory without becoming objects of close reading and fetishized imaginaries for dominant publics (Best and Marcus, 2009).

Bishop and Star (1996) describe the potential for agency that arises from technologies that undergird the construction of digital libraries: "One potentially positive implication of the increased mutability of document surrogates and bibliographic instruments, however, is that DLs could allow users to move beyond standardized records to create and manipulate knowledge structures in document surrogates that suit their own needs and practices" (page 321).The power of the of the Archive to create Designated Community and the ways in which this guides what information is collected and how it is structured belies the kind of agency that Bishop and Star describe here. The important thing is that the potential for agency exists in technologies and imaginaries of digital libraries, yet constructions of digital archives as informed by OAIS serve to constrain such possibilities.

There are difficulties in creating these types of Designated Communities in practice, above and beyond the ways in which practitioners struggle to narrow "the world" to a particular group of imagined users. Part of the difficulty in using OAIS as a recuperative technology lies in the epistemologies and ideologies of those in charge of actually designating the communities. Without casting aspersions on the digital preservation community about the commitment to the creation of equitable memory practices or lack thereof, there are other difficulties in designating users and creating services for these ideals. Practical struggles that emerged with philosophies of participatory design highlight some of these difficulties and serve as a reminder that, among other things, inclusion is not merely the opposite of exclusion. Oudshoorn, Rommes, and Stienstra (2004) note that many macro-sociological approaches to design, which are time and resource intensive enough for the creation of a single cog—the Designated Community—within an OAIS,

nonetheless "…cannot explain how technological artifacts come to incorporate barriers to specific groups of users" (page 31). This speaks to the creation of Designated Communities within an OAIS, but also to the creation process of OAIS itself. In addition to designing for themselves—Archives creating Designated Communities but also OAIS authors creating OAIS—designers often design for idealized users like the "econs" of the pop economics book *Nudge* (Thaler and Sustein, 2009). In that book, the authors claim that many economics projects fail because economists imagine perfect universalized subjects who behave robotically, predictably, and most notably, rationally at all times, for example always saving prudently rather than spending for pleasure. In discussing the disconnect between products and end users in relation to video game designers, Kerr (2002) notes that:

> A common theme emerging from this work is that while designers may try to design for certain ideal users this design process is often based upon partial or indeed mis-information about end users. Silverstone and Haddon point out that designers' knowledge of users is often tacit, contradictory and untested and in this uncertain environment organisational cultures and powerful sub-groups can compete to determine design. (Kerr, 2002, page 280)

I suggest that the ways in which some interviewees, and indeed authors of literature about implementations of OAIS cited in Chapter 2, speak about Designated Communities even though their institutions have not formally defined one means that their conceptions of users are based on tacit knowledge. This is not to say that Archive management does not have real information about the types of users that informs what they build into Designated Communities, but interviews and existing literature seem to indicate that processes like the

one described by Kerr in relation to game design are similarly pervasive in digital preservation repository design as well. In light of Designated Communities, how many of these concerns need to be taken into account? Do Archives have to go out into the world and base Designated Communities on an assessments of actual users? On paper, this is not necessary: the Archive can designate whichever community it wants. During the training course, another trainee in the class asked one of the OAIS authors about such situations. The response was that archives can make such assumptions if they want, but that this inhibits their ability to engender trust among Producers and Consumers. In essence, the response of the OAIS author was something analogous to a neo-liberal conception of market self-regulation: it is not necessary to regulate the construction of Designated Communities because Archives will do this themselves in order to sell their services to Producers and Consumers. Given the level of trust state and public memory institutions engender by their designations as official spaces of memory, particularly in the global north where these OAIS authors are based, I argue that this response is insufficient: in fact, history shows such institutions can do very dubious things and still maintain trust, funding, and users.

Designated Communities are often implicit rather than officially and publicly documented. Because so much of the daily processes in an Archive are imbricated with the construction of AIPs which are in turn built to express particular information to these Designated Communities, an implicit Designated Community means that the operations of the Archives are relatively invisible. Dissemination packages are not necessarily windows into these processes. Galloway, Lovink, and Thacker (2008) note how the digital interface, like the public facing content of an OAIS Archive that allows users to query for and receive

Information Packages, is material reality that structures a discourse embedded with historical relations. In fact, OAIS does not encourage transparency. Instead, the authors of OAIS, in their explanation of of Designated Communities, effectively direct archives to inscript users into their practice and archival constructs. By inscripting users via the Designated Community technology, archives are also directed to create power relationships between the included and excluded users that privilege assumed, default groups of users.

Kerr (2003) additionally writes about the line between configured users versus imagined users to note the difference between more versus less intentional conceptions of projected Consumers. To call users "configured" is somewhat too deterministic, where *imagined* is a better term: it is possible to imagine a user that will behave in a way that an Archive determines, but not actually configure such a user. However, I argue that this underestimates the power of the imaginary given the role constructions like Designated Communities play within the network represented by OAIS. Akrich and Latour (1992) diagram the interplay of human and nonhuman actors to demonstrate how they co-construct reality through impact on society and social relations:

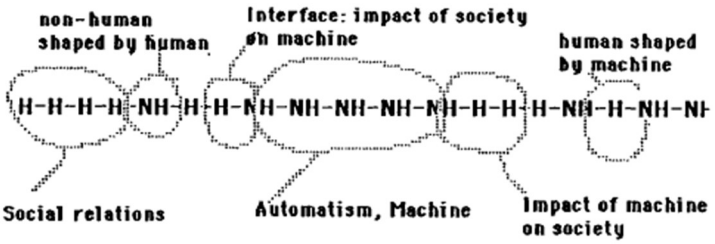

Figure 9.1
The usual categories that sharply divide humans and nonhumans correspond to an artificial cutting point along association chains. When those are drawn, it is still possible to recognize the former categories as so many restricted chains. If we replace H and NH by the name of specific actants, we obtain a syntagm. If we subsitute a specific name for another, we obtain the shifting paradigms.

Figure 7: Human and Non-Human Actors

The presence of the nonhuman actors is key here: OAIS is likely one such nonhuman actor in its role as a technology; I argue that the imaginaries of users like Designated Communities have a role in this space as well. This is echoed by Akrich (1992) who says, "The inscription of representations of users and use in artifacts results in technologies that contain a script: they attribute and delegate specific competencies, actions, and responsibilities to users and technological artifacts (pages 208)." Presumptions about users in turn affect access to and the relationship between users and technologies in what becomes a reflexive, constitutive and self-propagating process that mirrors the creation and role of discourse: the definition of users via a Designated Community impacts the ways in which real users can interact with an Archive and some of the tacit assumptions about them are self-fulfilled as a result of predefined modes of access. While this process is not entirely deterministic as users effect change on technologies all the time, the power inherent to this process and the power of imaginaries should not be overlooked. Van House and Churchill (2008) argue that "…what is remembered individually and collectively depends in part on technologies of memory and the associated sociotechnical practices…"

(page 296). In other words, the creation of Designated Communities that then inform Archive practice is a discursive process wherein Archives are creating the Knowledge Base for future user communities. What makes this particular concept of Designated Community so tricky is precisely that it asks Archives to predict the future, as a digital preservation and curation scholar at a US public university put it, not just technologically but socially and politically; further, the relationship is cyclical.

Science has dominated not just in the creation of OAIS but also, in some ways, the field of digital preservation more generally, and Designated Communities can be used a heuristic to examine the impacts of this positivist domination on evolving practice. This can be seen in comments like this one from a US-based data curation specialist who explained the scientific leanings of terms on OAIS thusly: "So, let me start by, you know, this idea that [designating a community is] easier for the space sciences community or the biology community or so on. I think what that reflects more than anything is that the best examples of data archives that we have so far are in fact disciplinary based." Without making a deterministic claim about this, I would argue that this participant is entirely right that there is a non-coincidental link between the things he describes. This interview subject says "easier for," meaning of course that the logical counterpart of this is that it is "harder for" people not in space sciences or biology, etc. Most participants, when I asked them questions about whether there were different subject-specific or disciplinary responses to OAIS, said that there were noticeable differences of needs and attitudes but that these differences stemmed from other sources (read: not disciplinary differences).

Despite this frequent claim, answers to subsequent questions in those same interviews belied the suggestions that humanist–scientist differences do not play much of a

role. Here is an example from a Canadian digital preservation scholar: "…[Designated Community] is also extremely weakly defined in the OAIS. There is almost nothing about it. It's there. It's there and it's supposed to work. And surely it is a little easier in a science background." Moments later, he continued: "I see very little reason why [the term Designated Community] should be any less workable in cultural heritage than in science data. There is repurposing and norms all over, and this is about prediction of the future, so that's not easy." During the course of the PVWII grant project, this disciplinary division became a focal point when speaking about Designated Communities within a broader conversation about significant properties, a concept that will be discussed in detail in the next chapter. In this context, Jerome McDonough, a digital preservation and STS scholar at a public US university said: "The NASA community is the birthplace for OAIS... The problem outside of NASA is that there isn't always a really well defined designated user community. As opposed to, say, who's the designated user community for the Library of Congress?" Trevor Owens, a US-based digital preservation scholar and practitioner responded by saying: "Scientists seem to assert that they know exactly what they care about, but there's always going to be another humanist example where you're going to want something else to be captured." Chris Melissinos, a practitioner at a major public museum in the US rounded out the exchange by saying: "This user community is also really passionate about exactly what they want and what they think is important. If we're too academic, we can overlook the humanity of the [a digital object] if we're only trying to preserve the bits. There's got to be a balance between academics and [users] in trying to decide what parts to preserve." Here, there is a continuation of trends mentioned above: for example, the hybrid "misconception" of "designated user communities." The language of OAIS is also elided over

the course of the conversation: by the final exchange, a museum practitioner shifts entirely to invoking "users," indicating the point at which OAIS ceased to be the common language of the discussion. The transcript of the entire conversation, only parts of which are reproduced here, began with an explicit reference to OAIS, the lack of information it provides about how to define Designated Communities, and what this means in terms of understanding the sub-groups that might be represented within a Designated Community. By the end, the conversation started by digital preservation scholars from major public US universities had shifted to the larger group, which included practitioners in digital preservation but also in museum and art spaces more generally. The OAIS boundary object disappears as an explicit node of the conversation yet remains in the ways in which participants talked about what it was that an Archive needed to provide users with when dealing with the preservation of video games. This move generally reflects the way that, in the US in particular, OAIS has become an invisible, yet ever-present player in digital preservation discussions and projects.

When speaking with OAIS authors, I posed a question about Designated Communities: is it possible for a community to refuse to be designated? This question is not trivial given historical precedents of oppressive forced grouping in both physical and other ways, such as oppressive catalog designations in Library of Congress subject headings. It is telling, however, that I was not able to convey the meaning behind such a question to a space scientist author of OAIS: the epistemological space between us was too large. But there are other things that are imbricated in this kind of question. When conducting an audit, OAIS authors suggest auditors check on the quality and compliance of Information Packages by finding an actual member of the Archive's Designated Community, plunking

them down in front of the Information Package, and asking them to demonstrate that they understand the content. This is an interesting suggestion because it forces a connection between Designated Communities and real people, something that comes from the audit process more than OAIS itself, wherein it is possible for the Designated Community to be imaginary as a heuristic for planning. But it does raise questions about communities who might refuse relationships with archives. This might be seen in politically unstable countries where there are issues of trust and safety with ruling regimes (Caswell, 2013b). But this can also be seen in spaces where there is a fundamental mismatch between the labor and recording keeping being done by the archive and the wants or needs of the Designated Community.

Consider the example of "missing" US census data. There was, of course, the loss of Hollerith cards in a 1921 fire which resulted in a loss of records from 1890. But the more telling incident here stems from the 1960 census. Researchers who use census data tell what NARA describes as "apocryphal stories" about missing US census data (Adams and Brown, 2000). In this case, the designation "missing" signifies a dearth of data that researchers want. The National Archives notes that these data are not technically missing because they were not meant to be kept: there was no records schedule to preserve the data that researchers want. This is a mismatch: the 1960 census information is not lost, because what is missing was not meant to be kept, yet the researching community definitely misses it. If the Designated Community for this census data is this group of researchers, then the National Archives is not OAIS compliant in the sense that the Designated Community disagrees with the assessment of sufficient authenticity stemming from the preservation activities on the part of the Archive. Or rather, given that the

Archives insist that the data are not lost or missing, does this necessarily preclude these researchers from actually being the Designated Community for this Archive? The question of who gets to be in the Designated Community and what actual communities get to say about this is not explored yet in the OAIS literature and this type of back and forth relationship does not work well with the ways the authors of OAIS conceived of Designated Communities. Given the central function of the term to the Archive and its processes, this has a profound effect on the ability to use OAIS as a recuperative technology for inclusion.

5.4 Conclusions

In this chapter, I considered a single term from OAIS, *Designated Communities*. This term was brought up both in response to specific questions I asked about it but also in relationship to a number of other themes that arose in interviews: audit processes, challenges of digital technologies, the challenge of heterogeneous user groups, the relationship between the sciences and the humanities, etc. Its omnipresence in the data matches its omnipresence throughout the OAIS documentation: it is required for the central functions of an OAIS such that its percolates implicitly throughout the entirety of the information and functional models. It is also a term embedded with many scripts that participates in many discourses. The power the Archive has to make this designation is not trivial, and this is a continuation of the power archives have had for millennia to dictate official memory and who or what gets included in it. Just as archives have been engaged in the practice of forgetting as much as remembering, the term *Designated Community* forces digital Archives to be in the business of necessary exclusion.

Yet this is also a space of potential for OAIS: it can be used in a recuperative fashion if Archives look to recent literature about inclusive and pluralist archives, libraries, and

museums for methods by which they can create Designated Communities with input from actual users and communities. I suggest that this is less likely to happen given the absence of formal models to help describe the process of actually making a Designated Community: when there is no guide, people will fall back on assumptions about users which will only serve to perpetuate the power disparity between archives and communities, particularly communities of sub-ordinate populations. It is precisely for reasons like this that one of the central arguments of this thesis is that additional metrics are needed to guide the processes that arise as part of deploying the OAIS reference model in memory institutions.

CHAPTER 6

SIGNIFICANT PROPERTIES

Introduction

The previous chapters examined findings from interviews conducted with digital preservation practitioners that highlighted, among other things, how situational understandings of OAIS are and the power relationships between the sociotechnical network that is represented by OAIS and its audiences. This chapter expands on the theme of situational discourses and understandings. I broaden the interview pool by bringing in interview data from Producers and transcripted conversations with practitioners and academics who work on the preservation of video game content. Here, I engage with the concept of *significant properties* and how these do and do not fit within the OAIS reference model. Note that the term *significant properties* is not capitalized: that is because it is not an OAIS specific term. Rather, this term is borrowed from traditional archival practice. Coping with this term represents another point of tension between the various disciplines brought together to construct the sub-discipline and profession of digital preservation.

Webb, Pearson, and Koerbin (2013) of Australia's National Library sum up the state of current development of significant properties as actionable within the general realm of digital preservation:

> We have come to a tentative conclusion that recognising and taking action to maintain significant properties will be critical, but that the concept can be more of a stumbling block than a starting block, at least in the context of our own institution.

This simultaneous acknowledgement of the critical yet poorly understood nature of significant properties demonstrates both the importance of the term, but also the barriers

to its productive impact given a lack of definitional clarity: significant properties have become an elephant in the room for digital preservation. Thus, this chapter argues that one method of synthesizing these various definitions is to engage with how this term is used in practice.

This chapter marries data from interviews conducted for the dissertation with data from the Preserving Virtual Worlds II (PVWII) grant alongside the framework set out by Giaretta et al (2009). I use these data to examine how well user-described significance fits within the OAIS entities for Representation Information, Provenance, and how this data fits within the OAIS conception of Authenticity. This work explores a concern laid out by Sierman (2012) in examining the extent to which Transformational Information Properties serve as an adequate substitute for significant properties. I argue through these data that some significant properties fit within the entities of the OAIS reference model, particularly those related to the digital object itself and the software/hardware environments required to make an object function. I also argue that OAIS, as it currently exists, cannot encapsulate all the types of significant properties derived from the interview data. The places where these mismatches occur are places wherein other preservation practitioners and scholars have identified weaknesses in the model related to the changing landscape of digital content, in addition to fundamental disparities in epistemology and values.

6.1 Significant Properties

There is a general consensus that significant properties are important and yet there are no widely adopted methods by which one can determine what is significant for preservation purposes. The lack of a simple and widely accepted definition is one part of the difficulty in actually evolving the term *significant properties* into concrete preservation

and curation strategies. General discourse on the topic refers to properties that are most essential to the understandability of digital objects over time. That is to say, significant properties recognize both the situatedness of digital artefacts and the fact that it may not be possible or practical to save every aspect of every object over time.

The term *significant properties* has been used in preservation and curation literature for over a decade. The most commonly referenced definition, and also an early one compared to others I reference here, is the one by Hedstrom and Lee (2002), who define the term as "those properties of digital objects that affect their quality, usability, rendering, and behaviour". The term is described variously in many places, and Giaretta et al (2009) discuss the difficulty in settling on a single definition for the term in part by exploring some of the myriad definitions that currently exist in disciplinary literature. Such discussions occur across institutions, information types, and research disciplines. Of science data, Sacchi et al (2011) say:

> Although this notion has clearly demonstrated its usefulness in cultural heritage domains, its application to the preservation of scientific datasets is not as well developed.

What precisely is meant by demonstrated usefulness is not entirely clear, as many practitioners in cultural heritage acknowledge the use-value of this notion without being able to advance either a concrete definition of what it means or how to account for it formulaically or machine-readably. One of my OAIS interview participants, a manager of digital preservation at a European national library, suggested that libraries are well equipped to deal with significant properties, "because…as a library we have a lot of experience in describing things so we are very good at metadata". This quote suggests that

she perceives a relationship between descriptive metadata and significant properties. Echoing the findings of Giaretta et al (2009) about the occasionally contradictory nature of various definitions of significant properties, another OAIS interview participant, a research and development officer at a national archive, said "well, it's just technical metadata, isn't it?"

The other difficulty with this term is that it represents a larger schism within the field of digital preservation between practitioners and designers from computer science and those who come from archival or library science. Kevin Bradley (2007) presciently said:

> "All God's children got significant properties," we can sing in unison, but this takes us no further if we cannot define its meaning in such a way that we understand what properties are under consideration, and describe them in a way that is machine-readable and automatically actionable.

This encapsulates the tension between the social, the human and the technical. Because all of these elements are at play in preservation, particularly when it comes to cultural heritage, significant properties serve as a potential flash point within larger preservation discourses, exemplifying the tensions that I have highlighted throughout this dissertation that arise around OAIS and the growth of the field of digital preservation.

6.1.1 Significant Properties and OAIS

The OAIS reference model has long and wide adoption within the digital preservation community. Further, the terms contained therein have come to function as boundary objects across different types of preservation and curation. As such, mapping significant properties to established entities from OAIS is a promising project for moving

concepts of significance into the realm of practice given the large constraint that OAIS provides for the development of any major preservation methodology or standard. Giaretta et al (2009) took on this task as part and parcel of what were to be upcoming revisions to OAIS in 2012, in response to calls by many in the field for clarity on this term and an expansion of its role within OAIS. In the cited article, the authors firstly examined various uses of the term significant properties and secondly proposed a number of existing, and thus more precisely or homogenously defined, terms from within the OAIS reference model that might be used instead of having to either consolidate or propose yet another canonical definition for this term. This was an important project: significance in this way does not appear in earlier versions of the model. In the resulting changes in 2012, the most recent as of the writing of this dissertation, significant properties are mentioned without being sufficiently addressed for some audiences (CCSDS, 2012). Barbara Sierman (2012) compares the most recent version of OAIS with its predecessors and notes:

> The Information Property is related to the commonly known but not always clearly defined term "significant property", but I think more discussion is needed to define better where the differences and similarities between the two concepts lie and how to translate this into the daily practice.

The Transformational Information Property in the 2012 revisions of OAIS is meant to stand instead of significant properties, rather than in place. During my interviews with OAIS authors, some interviewees noted that they decided to side-step this discussion entirely by creating a separate entity that would serve a distinct set of functions because of the sheer number of incommensurate existing definitions for significant properties. The key is that Transformational Information Properties are meant to work in conjunction with other

existing features in OAIS, in lieu of actually defining the term significant property, thus avoiding the need for authors and OAIS as a sociotechnical network to engage within this space. In practice, the outcomes are not so neat: by choosing not to wade into the significant property debate, the OAIS authors are taking an effective stand that indicates that the concept of significant properties does not need to be incorporated within the major standard of the field: it is not useful or important enough. This dictates in part how well it can be taken up by others given the pervasiveness of OAIS and the ways in which practitioners struggle to envision alternative frameworks for their preservation work. Additionally, the solution conceived of by the authors to avoid the term has not stopped practitioners within the profession from continuing to call for OAIS to deal with significant properties more explicitly.

Giaretta et al (2009) propose a framework utilizing existing OAIS entities as containers for the types of data that might be considered significant. A particularly important part of this move is the emphasis on the Designated Community, as authenticity does not exist in a vacuum but is instead a product of the relationship between a potential end-user and the data they might receive from an Archive.

Here, I use the term *a potential end user* purposefully to refer to the broadest set of possible users, because the term Designated Community refers to a specific audience rather than general users. This distinction is particularly pertinent for institutions such as libraries, whose users are a vast and heterogeneous group. In such a case, the Designated Community is necessarily artificially constructed in order to scale preservation practices so they are manageable within the resources of the institution. OAIS does not say how broad or narrow a Designated Community must be.

The previous chapter details the tensions engendered by the term Designated Community in OAIS: it is 'weakly defined' in the sense that the model does not concretely detail how to form and document such a community. While such specificity is not necessarily within the purview of a reference model, the missing piece here is that a subsequent guideline or standard about Designated Communities has not yet been developed, and that many institutions have not, at a site-specific level, formally defined their Designated Communities. Lee (2005) describes the genesis of standards within a particular field: reference models happen first, are high-level and aim at providing foundations for future work. Rather than providing specific guidance, they are meant instead to spur the creation of further standards that will detail implementable workflows. Lee provides this diagram by Cargill (1997, page 92) to explain this genesis:

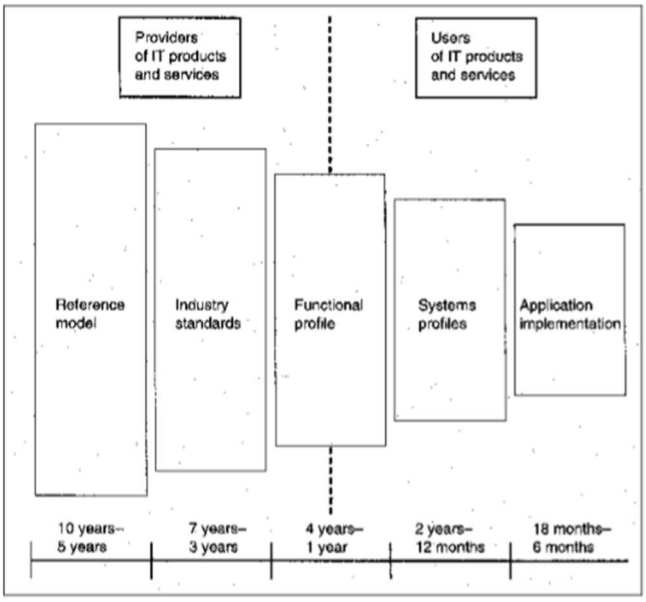

Figure 8: Genesis of Standards

Work that does not address the Designated Community cannot address the significant properties elephant. In dealing with a concept like significance, it becomes necessary to ask

significance for whom, something that is often implied but not always specifically addressed in discussions of significant properties. Yeo (2012) sums this up eloquently:

> However, the determination of 'significant properties' is no less problematical than the debate about notions of value …not least because different user communities will bring different perceptions of what constitutes significance.

The situated nature of the Designated Community and the idea of *preservation for someone* arises from the same discourses of place and time that inform conversations about significant properties. Struggles I identify here are due in part to changes in technological landscape the importance of which authors of OAIS were not able to predict. This is not new: for example, earlier versions of OAIS assumed migration[45] to be the default method for continual use of obsolete formats in future, yet recent years have seen a shift away from migration and normalization towards a more mainstream acceptance of emulation. The 2012 OAIS revisions encapsulated this change. Recent developments in areas like linked data and other forms of distributed content pose a challenge to the current iteration of the OAIS reference model, and practitioners like David Rosenthal (2015) have made calls for attention to this as OAIS heads into a new round of revisions in 2017.

6.2 Preserving Virtual Worlds II

This chapter utilizes data gathered as part of the Preserving Virtual Worlds II grant. PVWII was funded by IMLS and concluded in 2013. It included investigators from the University of Illinois, the University of Maryland, Rochester Institute of Technology, and Stanford University. Investigators examined the concept of significant properties as it

[45] I use migration here in the sense of moving content to a new, perhaps related, format that results in changes to underlying code and structures even if surface aspects remain the same or similar. OAIS uses the term Migration to mean something specific.

applies to video games with the aim of informing preservation practices for complex media, building on previous projects that examined the significant properties of software and a previous game preservation project, Preserving Virtual Worlds I (PVWI) (Matthews et al, 2008; Knight, 2008; McDonough et al, 2010). Broken into two investigative phases, Phase 1 entailed a two-fold method for examining significance. Investigators performed technical and content analyses of a number of video game series. The game series spanned a time frame from 1971 to the present and ranged across different game genres[46]. Simultaneously, investigators conducted interviews with people involved in the design and dissemination of games from the case set; with designers working in other game design studios; and with fans and programmers who have worked on more well-known modifications (mods) of some of the games from the case set. These interviews were qualitatively coded and analyzed by members of the research team across the various institutions involved in the project.

Phase 2 of PVWII focused on the development of tools and metrics to assist in the preservation of the significant properties identified from the research in Phase 1. These included an examination of how such properties could inform decisions about the emulation, migration, and re-implementation of games as well as defining benchmarks for authenticity in playback. The Phase 1 research painted a very complex picture of significance within the realm of games, and a key finding was, unsurprisingly, that significance is highly situated, which echoes findings about the deployment of OAIS as

[46] I use the term genre very loosely and will not attempt to classify the game set by genre; the investigative team for PVWII struggled with this, particularly the way that existing genres popularly used for games are neither commensurate nor mutually exclusive. As PVWII wrapped, the team decided not to include a category for *genre* in the wiki tool designed to collect user-defined significant properties.

discussed in previous chapters. The research data indicated that what is significant about games may not be something inherent to the game's code (bits) or even computing environment (platform, operating systems, controls), but could include elements as varied as underlying data models or general surface affective experiences. As such, PVWII suggested a layered model for looking at games, delineating different aspects of each system wherein different users might locate significance. I will discuss this model in greater detail later in this chapter.

A second deliverable was the creation of a survey tool which borrowed from earlier projects such as the Variable Media Questionnaire (Ippolito, 2003) that aims to capture significant properties as defined by various stakeholders, including designers, players, archivists/preservationists, and curators. This tool is designed as a wiki for the purpose of collecting and automating the analysis of large quantities of data that will serve as a record of the Knowledge Base of different user communities. In combination with game-specific contributions to format registries, another PVWII phase 2 goal, the hope is that preservation of games can be enhanced by crowd participation in the process of gathering and centralizing previously dispersed but necessary information about games.

6.3 Mapping Significance

Given the ubiquity of OAIS, I wanted to understand how complicated multi-part works like video games and virtual worlds fit within the model by allocating significant properties to existing OAIS entities and identifying those which are a mismatch with the model. Video games are exemplars for preservation work because of their complexities: as digital objects, they have numerous dependencies, both analog and digital. They also, by way of their popularity, have enormous user bases. They are often highly proprietary and

yet belong to companies that go in and out of existence, with intellectual property traded, sold, and occasionally abandoned. They also function as heuristics for understanding the essential yet nebulous OAIS concept of Knowledge Base: "…The games of a people reveal a great deal about them" (McLuhan, 1964). The heterogeneous users of video games, which include leisure players, professional players, in-game laborers and others, have the possibility to help illuminate the boundary between necessary and non-essential information that make an object understandable: basic assumptions that players will understand the nature of a simple d-pad for 2-D representations, for example, are necessary for some kinds of play. Understanding that Yoshi is a dinosaur instead of a horse, for example, is not necessary: the essential piece of information there is that the creature can be ridden and this is how game players should approach it.

Video games and other complex types of *work* pose a challenge to digital preservationists for two reasons: first, as mentioned above, the large and general category of significant properties is one that OAIS intentionally avoids. Second, while the notion of Archive in OAIS is very specific, it shares some foundational tenets with the study of traditional paper archiving practice. Several interview participants told me that it is not the job of the archive to collect and preserve external significant properties. The tension here arises from the fact that cultural heritage objects are inherently socially constructed and products of a particular place and time. As Gitelman (2005) points out, humanistic histories look at objects whose role over time changes in different ways than the artefacts of the history of science: art is still art hundreds of years later while scientific objects of similar age are rarely still used as active instruments in the production of new scientific knowledge. Both the science of archives and the positivist scientific discipline that helped

form OAIS are suspicious about the notion of selection particularly as it relates to value. Within traditional archival practice, the term selection is used very narrowly: the scope and collection statement determine the type of content to be collected, and all such content from the organization is archived rather than an archivist selecting certain materials for processing and preservation in a more colloquial sense. Creating Information Packages for complex media requires some measure of more colloquial selection, where selection means precisely this latter concept: that the Archive must choose a set of things to include in the package that encompasses the most significant properties. The purposeful and transparent creation of artificial boundaries is at odds with foundations of archival practice which inform the authors of OAIS and how OAIS gets deployed. This is the second mismatch.

I will focus on interview data related to two games franchises from the PVWII case set: *Carmen Sandiego* and *Civilization*. For both games, multiple creators were interviewed, painting a broad and varied picture of significance as determined by creators. In order to determine how well this significance data can be captured by the entities in OAIS, I parse the data to look specifically for information that could be modeled as Representation Information; the documented as Provenance; benchmarks for authenticity; and that can be modeled as Transformational Information Properties.

In keeping with the requirements of OAIS, I define the Designated Community based on self-descriptive information derived from the interviews with the broader notion that, for this case study, the Designated Community is comprised solely of the interview participants. The process of defining a Designated Community for this portion of the dissertation acknowledges a limitation of this work: the preservation here is for someone, so the model here aims to preserve what this Designated Community identifies as

significant. The process of designating a community acknowledges that these properties may not be significant for everyone and what I propose preserving may not constitute a sufficient or authentic representation of the content for everyone.

Finally, I identify significant properties that do not easily fit within the Representation Information of particular digital objects and discuss why it is that these do not work within the current iteration of OAIS. Some of these properties are related to the tricky OAIS term Knowledge Base; others are distributed in a manner that challenges OAIS's requirement for adequate control of the content.

6.4 Findings

My study examines a difficulty in the day-to-day deployment of earlier frameworks, whose basic constructions often insufficiently account for significant properties that are not inherent to the digital object itself, such as those that are not intrinsic to the code yet still essential to some kind of long-term understanding. The current interest in emulation as a preservation method does in some measure move the preservation community towards an acceptance that things beyond the object themselves are significant and require preservation – in the case of emulation, significance is found in the behaviors of the original computing environment. The work here looks to extend this by forcing the consideration of even broader data about significance that may encompass aspects of the social and cultural and aspects of the Designated Community's Knowledge Base. These data about what really make a game, and relatedly what really constitutes a 'digital object', are not and should not be incompatible to the precisely defined categories existent in OAIS because the model requires updating AIPs as the Knowledge Base of the Designated Community changes over time. This requirement acknowledges that there is more involved in understanding objects

than simply recreating the objects themselves: artefacts are a product of a particular place and time, and are understandable as such. They have an 'aura', a term from Walter Benjamin that Abby Smith (2003) roughly equates with an experiential and affective authenticity. If what is called 'significant' by participants in the PVWII survey and interview data does not map well into the Information Package particular to individual video game Digital Objects, this does not absolve the Archive from the need to treat and preserve this kind of information.

6.4.1 Significant Properties in PVWII Data

Giaretta et al (2009) note that, "The notion of Significant Properties has emerged as a key concept in preservation within the library community but has not been a concept that is much used in the context of the preservation of research data that is not normally viewed as a document." This speaks to disciplinary divisions like those discussed in previous chapters, in particular the assumption from systems designers, engineers, and computer science practitioners that the other professions involved in digital preservation bring with them a level of imprecision and unhelpful adherence to tools and terms borne of a lack of understanding about purposeful systems design.

The CEDARS project defined significant properties as:

> those characteristics [both technical, intellectual, and aesthetic] agreed by the archive or by the collection manager to be the most important features to preserve over time.

It is very much in this vein that the PVWII project team approached significant properties, beginning with a note from the PVWI final report when the project team for the first grant noted: "Without a clear understanding of which aspects of a game are likely to be

considered significant by scholars in the future, it is extremely difficult to choose an appropriate preservation strategy and preserving games without any change in their appearance and play may simply not be achievable in many instances" (McDonough et al, 2010, pages 6-7).

Sergeant (2002) on the other hand proposed that "Significant Properties are those attributes of an object that constitute the complete (for the intended Consumer) intellectual content of that object". While the notion of completeness here is an interesting one in light of the PVWII findings—for example, could a preserved version of Carmen Sandiego be considered complete without its accompanying analog paper encyclopedia—I find the primacy given to an individual object problematic. Singling the object out as divorced from is spatial and temporal context will not guarantee the understandability of the object over time, even if its rendering environment and bits are preserved. This is a problem that both Giaretta and colleagues and the OAIS model acknowledge: that the situation (that determines the situatedness) of an object needs to be preserved. I argue that this is what is encapsulated by the terms *Knowledge Base* of the *Designated Community* within the sociotechnical complex of OAIS, even if the explicit definitions in the OAIS documentation do not indicate this. The often described example is the shift from a dominant language of English to Spanish, at which point the records in traditionally English speaking archives need to be marked as explicitly English, a denotation that was previously unnecessary as English enjoyed a decentered and invisible privileged position as the dominant norm. Another example that stems from the PVWII research is the change in geography and maps over time: *Carmen Sandiego* games frequently involved chasing 'bad guys' across various geographic locations (read: followed that VILE agent from country to

country). The contemporary Knowledge Base of the 1985 game player (assuming this person is part of the Archive's Designated Community) involved maps that, for example, did not contain South Sudan and North Sudan as separate countries and that showed the USSR as a single block rather than the de-federated former Soviet nations contemporary to the writing of this dissertation. Maps are often represented as documents, and documents make for easy pieces of Representation Information to store along with the digital object, all under the OAIS entity entitled "Content Information". But when a digital object like the 1985 original floppy disk version of *Carmen* is seen as imbricated in a complex and ever-changing sociotechnical network (or as a network unto itself), then there are subtler changes that occur that are more difficult to document than an English dictionary or a map. In several Carmen games, South East Asian countries are typified by images of people in conical hats working in rice fields. Even as of the writing of this dissertation, this image still allows game players to identify a certain part of the world, but this knowledge will change rapidly. Damming of the Mekong River and rising water levels associated with global warming trends mean that large swathes of the Mekong River delta, known as one of the top rice producing and exporting areas in the world, are at risk of being flooded with salt water from the sea. These climate changes threaten to end the farming of rice in these areas: if these trends continue unabated, within a few decades this region will no longer be the center of the rice growing industry. With it will go the cultural association of people in conical hats bent over rice fields as production shifts to Africa, where popular imaginaries suggest different visual markers to note time, place, and occupation. At this point, parts of the game that rely on tacit knowledge that recognizes images of conical hats and non-descript green fields (this non-description being due largely to technological limitations at

the time these games were produced) means that the game can no longer be played as it was meant to be: the very behaviors of the digital object break down without enough understanding about the contemporary Knowledge Base of original intended users: without this knowledge, players do not recognize that this image connotes a particular location. Without recognizing a location, a player cannot continue their pursuit of Carmen across the world.

And so this situatedness, I would argue, is a significant property in the sense it was described by PVWII: without this kind of information, the game is not playable over time even if the bits and computing and rendering environment are preserved. A current presumption of game preservation is that a game, by its nature, is meant to be played, so if it cannot be played, we cannot be said to have preserved a working copy (McDonough, 2013).

Significant properties, as identified in interviews from PVWII, could be located at any point in the layered model developed during PVWII for representing video game components.

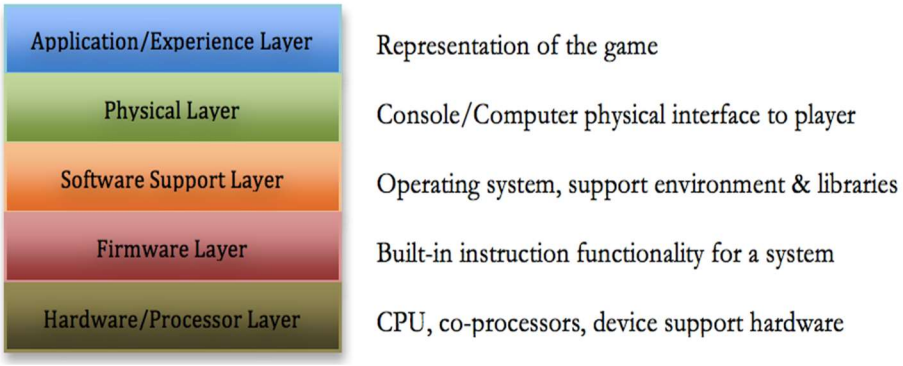

Figure 9. Layers of a video game (Decker et al, 2012)

For example, some video games were designed around specific *software support layers,* layer three on the stack, such as the first *Civilization* game designed to work with early Windows operating systems. The function of the then-novel eponymous windows was incorporated heavily into the game, and constituted a significant property to the developer we spoke with, who mentioned the role this operating system played in the game's development. The Nintendo game *Duck Hunt* notoriously used a special peripheral *hardware* piece, layer five on the stack. A light gun (as opposed to the normally used d-pad and 4 button controller) allowed players to shoot at ducks, as the name of the game implies, and the game is not functional without this piece of equipment. In 2010, a fellow student at New York University did a term project on the preservation of *Duck Hunt* and noted the comparative ease of finding working cartridges and working Nintendo hardware and the relative difficulty in finding the special light gun[47]. As of the writing of this dissertation, six years later, another piece of the hardware layer of *Duck Hunt* has become increasingly difficult to locate: the cathode ray tube (CRT) televisions that the light gun needs in order to function. A CRT monitor might be considered *hardware* or might be considered part of the *physical layer*, layer two in the stack, as part of the physical interface to the player. These twin external hardware dependencies, both of which are essential to a functioning version of the game and so might be considered significant by some Designated Communities, together pose a double difficulty in preserving the game.

Yet most significant properties identified by the interview participants in PVWII fell unambiguously under the top, *application* layer of the stack, which is the representation of

[47] A student of mine at the University of Illinois similarly noted the difficulty in finding the mouse peripheral that came bundled with the Super Nintendo game *Mario Paint.*

the game. As a result, I divide the significant properties in this data into three categories, according to where they can be located in relation to the layered model: two of these lie within the top layer of the stack and the third lies, purposefully, outside the stack altogether.

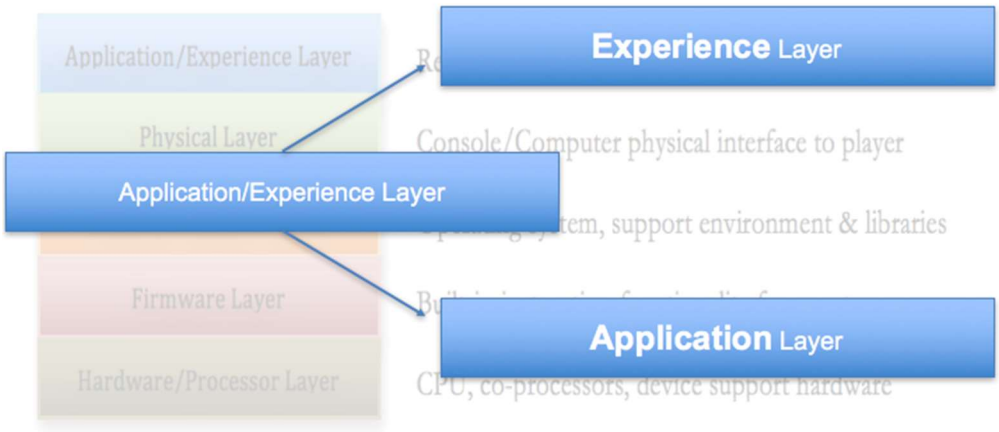

Figure 10. Application and experience as separate layers

I firstly break the Application/Experience Layer into two parts: Application and Experience. These encompass many of the significant properties identified by PVWII participants. The application layer includes things like the game code itself, as well as items like jump tables for early *Mario Brothers* games or historical statistical mortality data that determined how likely a player was to die when playing *Oregon Trail*. The experiential layer encompasses the surface and affective experiences of playing the game, the 'aura': the fact that Carmen Sandiego is first and foremost a scavenger hunt and that it is only kind of a one-player game despite its single avatar, for example. I separate these two because I argue that they are not necessarily related. To be sure, the original code in conjunction with a computing environment were necessary to manifest the original playing experience. But to recreate experiences like this, the mnemonic experience rather than evidential information

(Smith, 2003), to give an authentic representation of the experience of play, the original code is no longer necessary. PVWII investigators posed the question to game designers: how important is the original code if you can generate the same surface appearances and behaviors with a different backend? Most responded that they were not wedded to the original code, but more so to the experience of play. Some noted that the original code itself was 'poor', often due to time constraints. These two things can exist separately: because it is possible to save 1s and 0s and even consoles and media without saving the experience and it is possible to recreate the experience without the 1s and 0s, I separate this layer into two discrete layers.

Finally, I also argue that some kinds of significance, as described by PVWII participants, lie outside the stack altogether—that is to say, they cannot be found in code or computer environment. These include significant properties like those I term *relationally significant*. PVWII investigators asked participants to name their favorite game franchise and to explain what made it so important. One point of significance that was mentioned was the way that a particular title within a game franchise was 'leaps and bounds' ahead of its predecessors and other titles available at the time. I term this *relationally significant* because understanding this statement about what makes a game important requires placing it in context almost like archival bond with other games of its time. To understand this property of a game does not require a playable copy, although it might entail placing a playable copy up against playable copies of its contemporaries to demonstrates its advancements. But there are other ways to represent and benchmark this: for example, placing visuals from saved game files, videos of play, or machinima in relation to visuals of its contemporaries. The constant feature here is that is impossible to understand the

"advancedness" of a game by looking at the game itself: it has to be seen in relation to other things.

PVWII interviews also raised other affective aspects of game play as significant. For *Carmen Sandiego* and *Oregon Trail,* two franchises that are often termed *edutainment* games to the chagrin of their producers, interviewees expressed that understanding them in the educational context of the mid-1980s is important to understanding the experience of play. Like the tangible difference between playing a quarter-guzzler like the original *Donkey Kong* and playing a game on a home console (one designed to be short to eat money, the other designed to be long to engender customer loyalty to a product), there is a tangible difference between playing *Carmen Sandiego* solo at home, with Google at your fingertips to answer questions, versus playing it in its original environment: picture several kids clustered around a tiny and expensive computer in a school, because no one had a computer at home. One person might have controlled the keyboard; another would have held the accompanying encyclopedia. One interview participant who worked on programming for *Carmen Sandiego* said that seeing these games in context was how he envisioned ideal preservation for his games, while acknowledging the difficulty in manifesting something like the mnemonic impressions of a particular time and place.

6.4.2 Significant Properties as per OAIS Interviews

While coding the interviews with OAIS practitioners, authors, and scholars, I identified two nodes regarding significant properties. Firstly, there were explicit mentions of significant properties. There were only three instances in which significant properties were brought up by my interviewees, and this makes sense given that the dominant theme of these interviews was OAIS and the fact that *significant properties* is not an OAIS term. I

mentioned one of these instances earlier in this chapter: a technology officer at a European national archive suggested that significant properties for the data she had to preserve, mostly video and audio files, were technical metadata. The three explicit significant property instances echo the dominance of the OAIS authors in shaping how people within the realm of digital preservation continue to respond to and understand their work in relation to OAIS. In one instance, I asked an interview participant about significant properties specifically because I knew this participant had made public statements about them in relation to OAIS in the past. In this particular discussion, the interviewee mentioned significant properties in relation to enrolling analog professionals within libraries in digital work. The interview subject said:

> …we have more analog material… and lots of people were trained to deal with analog material and fewer people are trained to deal with digital material. And as you can't just give them the sack [laughs], you need to deal with them, you train them or whatever, so that takes a long time and I think that's one of the problems all main libraries are dealing now with that they have staff that's not quite prepared for digital material. So that the thinking about OAIS starts within a… small group of people… and we tell them 'I think you should interpret it like this or like that' and what you don't see is that they try to translate it to their analog environment and sometimes that does not work because it's digital. So it's difficult to translate I think, although the model itself is very clear, I think it's rather straight forward, but when you go the significant properties, well, endless discussions.

This interview participant, someone who is both a library practitioner and actively involved in OAIS revisions and related standards, describes OAIS as "relatively simple". In

this case, the designation of simplicity is meant, as much as anything, to indicate how not simple the concept of *significant properties* is. The situation in which she is working is already a fraught one to some extent: the library has a large analog collection and many analog employees, and moving into the digital space requires people to learn new skills. The interview participant says that it is not possible to simply fire the analog professionals who cannot make the shift to digital, and this is meant literally as well as to suggest that to do so would be wrong. This interview took place in Western Europe, where labor laws would prohibit such actions. Indeed, this was not the only person out of the European interviewees to make just such a comment about analog employees. The fact that firing analog-skilled employees is stated consistently across different institutions, however, is suggestive: someone at some point thought firing reticent, or "conservative" as three other participants described them, employees was the best solution moving forward in the digital era. And it is under this umbrella discussion, speaking about employees who work with analog materials, who cannot make analogies between their previous work and their digital futures, who struggle with a simple model, often because they "only read the first 80 pages [of OAIS]" according to the same interviewee, that the subject of significant properties arises. As suggested elsewhere, this notion comes from library and archive traditions, and therefore clashes with data and systems design origins that dominate the construction of OAIS. This is the unresolved tension a reviewer noted in response to an article I submitted on the subject to a major preservation-oriented conference. And perhaps it is the perception by OAIS authors that significant properties come from libraries and archives that predicates its continued exclusion from OAIS.

A second mention of OAIS came from a US-based data scientist who said:

> I mean, if OAIS didn't exist, you know, people would still need to preserve things and they would come up with some other framework, and obviously it would be not exactly the same as OAIS. It would probably have a lot of the same ideas in it. There were, you know, obviously… concepts that I used before I ever saw OAIS, but when I saw it, I thought, "Oh, yeah, this maps to this in OAIS." And OAIS has concepts in it from earlier versions of OAIS that aren't the same anymore like format migration isn't called format migration anymore, it's called transformation. And significant properties are now like transformational information property, you know, and things like that.

This suggests a familiarity with the process of OAIS creation and revisions, such that this person is aware of the fact that Transformational Information Properties are the official term meant to deal with significant properties. This interview subject speaks from a place of privilege: as a data science scholar, this person was already familiar with the type of terminology that is contained within OAIS, and is happily fluent in its lingua franca. In fact, of all my participants, this one had the fewest complaints about OAIS, expressing most answers in form similar to the quote above.

The comment by the US-based data scientist about the relationship between significant properties and Transformational Information Properties is a common misconception, if it can be called that. It may simply be a casual simplification. While Transformational Information Properties are meant to encompass some aspects of significant properties, they are not a replacement. Defined in the 2012 revisions as an:

> [i]nformation [p]roperty the preservation of the value of which is regarded as being necessary but not sufficient to verify that any Non-Reversible Transformation has

> adequately preserved information content. This could be important as contributing to evidence about Authenticity. Such an Information Property is dependent upon specific Representation Information, including Semantic Information, to denote how it is encoded and what it means. (The term 'significant property', which has various definitions in the literature, is sometimes used in a way that is consistent with its being a Transformational Information Property). (CCSDS, 2012 page 1-16)

It is stated quite clearly that this definition is meant to cover only some versions of significant properties. Depending on the definition of significant properties one employs from among the myriad ones in existence, some of these properties are contained within entities that predate the 2012 revisions, including under the Digital Object itself as well as in places like the Preservation Description Information entity, without necessary reference to a Non-Reversible Transformation.

These are the two distinctly interesting explicit mentions of significant properties in my interview data. The more populous node, however, was for what I identified as implicit significant properties. I applied this label to any discussions wherein an interviewee mentioned some aspect of a digital object without which that object would not be understandable, functional, authentic, or worth preserving; in other words, properties labeled by the participants with any of the descriptors from the myriad definitions of significant properties at the outset of the chapter. These include a number of references that echo the PVWII data. One practitioner mentioned a concern about the dependency on outside objects for understandability, in particular external technologies. This US-based museum practitioner also said:

> Yeah, like Windows '95, we need a place to track that information and because there is a many-to-many relationship there, it makes sense to record that in a structured way where we have some kind of master record of all these technologies.

This comment was in reference to the difficulty of creating mutable AIPs using the software programs the institution uses to track all the information needed to keep digital art functional. The substance of the comment mirrors discussions with video game creators who referenced the significance the role of the operating system, coincidentally also Windows '95, played in the creation of a title within an iconic video game franchise.

Likewise, the experiential aspects of digital objects also arose in the OAIS interviews. One participant, a digital preservation manager at a private US university, said:

> …Maybe we need to be more clear about it's not just about providing [access] to the files, it's about providing an experience… I mean, I like to think about it as being able to present the same content to the user…we could have documented that content, regardless of the experience through which they receive that content, even if the content is an experience… I don't know. It's complicated…And I also feel like… just in general… there's so much interaction, and the experience of being able to work and build, something like that.

This is not to say that engaging with analog materials is not experiential: indeed, reading a paper book is an experience, and reading a Dickens novel as a set of serialized chapters over the course of months is not quite the same experience as reading the entire work at once when it has been collected into a single volume. But in this case, the interviewee is expressing something fundamental about the interactivity of many types of digital content. I take this ethos to be the same one that motivates the response on the part of video game

programmers that it is the look and feel and even social experience of playing a game that is more important to preserve than the code. This is precisely the difficulty that preservationists face with dynamic and interactive content like Twitter: preserving the text and even the still images embedded within Tweets along with time stamps and GIS data is not, in and of itself, a particularly difficult task: these are fairly simple information types for which the profession has begun to develop fairly robust means of preservation. Rather, the difficulty lies in the preservation of a surface experience, which is as much a significant property as the content of the text itself. Cases like video games and Twitter offer heuristics that demonstrate one of the chief difficulties in the realm of preservation: it it very hard to predict the future. If, in future, the notion of limiting forms of speech to 140 characters has disappeared, then the intellectual labor of using a combination of Internet slang (LOL, SMH, U, etc) and emojis to express long thoughts ceases to be visible. In fact leaving the contemporary moment is not necessary to see how situated tacit knowledge is necessary to understand the experience of digital objects: this concept is particular not only to a time but also a place. China's social media tools that also limit character numbers offer far greater latitude in terms of some kind of depth or length given the way that ideographic languages work versus something like a Roman or Cyrillic alphabet or even Japanese kanji.

The greater difficulty is how to demonstrate, through the DIP, the temporal nature of Twitter: just showing the volume of Tweets around a particular subject or hashtag does not connote the process of Tweets evolving in a particular moment. There is a qualitative difference between the long, slow accrual of anecdotes to go with the hashtag #dissertationproblems versus the rapid-fire and time-sensitive torrent of Tweets about

#hasjustinelandedyet, in reference to a woman who posted a seemingly offensive Tweet moments before taking off on a long flight, during which time her story went viral.

These are particularly challenging significant properties for preservation, and I argue that this is partly because the work of conveying this information back to users is a function that happens in a couple entities within the OAIS information model. First, an Archive must store sufficient information within its AIPs to know enough about an object to be able to convey this information: this includes something like the OAIS/FRBR mapping constructed as a result of PVWI wherein the model suggests linking to an outside source for Context and Provenance information (McDonough et al, 2010). Perhaps in a case like #hasjustinelandedyet, the AIP would contain not only the Tweets with text, images, and time stamps, but also references to articles both popular and scholarly explaining what happened. The digital preservation manager at a private US university quoted above describing the interactive nature of technology mentions the practice of documenting the experiences of users, and for very complicated media that is one of the few (perhaps the only) option at this point in time. Another interview participant, a researcher at a European national archive, said:

> I looked at technological hardware preservation. I looked at simulation—yeah, migration and emulation then documentation. Documentation is kind of like a separate thing but I felt because so many of these other things there are so many reasons why we can't really do that yet. I feel like documentation is basically what we're left with.

The second entity that is implicated in this process is the DIP. Conveying these types of experiential significant properties to Consumers requires creative work through the DIP—

this information must be disseminated in a way that connotes the affective experience of the digital object or collection. The DIP is one of the more poorly defined entities within OAIS in large part, as one interview subject said, because it requires a prediction of the future. While querying Twitter or the Library of Congress' Twitter archive now, Consumers are contemporary enough to many situations that a mere return of text and time stamps is sufficient. But at some point in the future, when the Knowledge Base has changed and people no longer understand what Twitter is or how it works, this DIP has to change. There is recent work that provides formal modeling of DIPs displayed as a set of services and exchanges with Consumers (E-ARK, 2015) and this work acknowledges the need for DIPs to change according to queries by Consumers; it suggests tracking these and potentially adding them back into the AIP using the PREMIS standard for documentation[48]. But even though this is a more specific and prescriptive standard that follows OAIS, it does not and perhaps cannot help to address what will need to be somewhat imaginative solutions to conveying the experiences of interactive and dynamic digital content. For something like Twitter in future, this may be an interactive display that shows the speed, rapidity and global spread of Tweets as dots on a map or time-lapsed video of a Twitter interface. This entire concern is imbricated in the complexity of Designated Communities and Knowledge Bases. Archives are supposed to track Knowledge Bases and update content when Knowledge Bases change. This is a difficult task, not only because there are no current guidelines that deal specifically with this[49], but also because change is both a hard thing to

[48] Note that PREMIS is another standard that was built to conform with OAIS and one that claims to be moving 'beyond OAIS' (Dappert, 2015) in it's latest revision which was made public in the summer of 2015.

[49] Although one interview subject suggested the outcomes of the SCAPE project: "a lot of the idea in the SCAPE approach of preservation monitoring and planning is predicated on

notice in the moment and a more difficult thing to document after the moment has passed. That there is no one solution is part of what makes this kind of thing hard to standardize; that there should be guidelines anyway is probably obvious given the complexity of the task.

There are also ways in which it may be possible to overstate the difficulty of the digital preservation task: it may be that at this particular juncture, the preservation of surface and affective experience is not possible, particularly not at scale. One interview participant, a senior digital preservation consultant at a boutique US firm, noted that these preoccupations can serve to paralyze the field in such a way that getting to grips with what should be relatively simple tasks like bit-level preservation still have not been definitively addressed:

> Yeah, I think it actually… and this isn't OAIS's fault, it's just I think this field has suffered from -- in my opinion, it has suffered from too much fixation on those kinds of issues and not just doing the absolute minimum to get you to a point to have a future opportunity to visit those questions when the need really arises. We don't even have good bit preservation nailed down, and that should be very easy. It's really simple, it's dumb, just do it, and stop talking about it, please. I'm so tired of it.

This participant also noted that concerns about significant properties are more challenging for some kinds of content that others. For audiovisual materials, she argued: "Watch it and listen to it, and look at it." Another participant, senior special collections archivist at a private US university, said, "So, for us to be able to push [a digital object] into something

---

evolution of and instruction of the designated community in technology, in semantics, in usage, in requirements."

where we have, you know, huge, huge disk space, and to be able to say well, at least you know, it's safe, the original is safe. I would think that would be like a big plus to people, just to be able to provide that as a service for their materials." Keeping the 1s and 0s safe is a most basic requirement, and this might be seen as sufficiently significant in many cases, particularly if this is explicitly stated in users and donor agreements. Yet at the same time, multiple people have pointed out, including the authors of OAIS that I spoke to, that 1s and 0s alone are rarely sufficient, particularly when longer time scales are involved. A digital library director at a private US university summed it up nicely: "I mean, files are not that useful without something…"

6.4.3 Mapping Significant Properties in OAIS

The previous two sections have detailed some of the significant properties that arose in conversations with game programmers and OAIS practitioners and scholars. Here, I will demonstrate what maps well to the existing OAIS entities, informed by the work of Giaretta et al (2009) and the 2012 OAIS revisions. I will also show what works less well.

The figure below is an image from OAIS that details the contents of the AIP. I have highlighted in purple the entities wherein some significant properties could be located and I speak about some of these in the examples that follow.

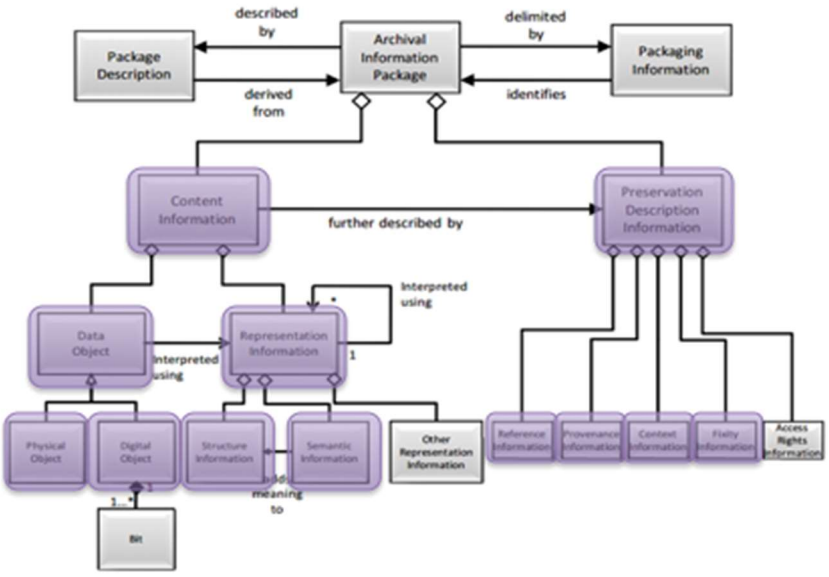

Figure 11: Significant Properties in OAIS

Some significant properties fit well within the Content Information entity in the AIP model. Content Information includes the Data Object itself, which can be comprised of both Digital Object(s) (bits) and Physical Objects. Source code of games fits here as do some essential physical ephemera, like *Carmen Sandiego*'s analog copy protection *World Encyclopedia*. Ephemera can be documented as a separate object and related to the digital data via the Context Information entity.

Access software, and by extension, access hardware may be documented as part of the Data Object itself or as Structural Representation Information. Changes in the Designated Community's Knowledge Base may be documented as Semantic Information, although there are limits. Including software as part of the digital object itself is something that OAIS does not do very well yet, according to some practitioners. One of the interview subjects has argued vociferously and publicly for its inclusion as part of the object itself in the 2017 revisions. Semantic Information can document significant properties like a

language shift from English to Chinese, for example. Preservation professionals dispute whether or not it is the role of the repository to document changes in common knowledge, such as geographical names and borders or popular imaginaries in the case of *Carmen Sandiego,* as discussed in the previous chapter.

Sometimes what is significant about a game is its relationship to other games. One game programmer said, "*Doom*, for example, it made some of these huge graphics and texturing leaps and bounds, [these were] obviously… a product of its time." "Leaps and bounds" progress in one game necessarily relates it to a history wherein a game was markedly different that its contemporaries, as noted previously. Another significant property noted by interviewees is the relationship of a particular title to a larger franchise, for example a particular release of *Civilization* in relation to all versions. This was stated explicitly but is also tacitly implied when participants spoke about franchise games by punctualizing an entire series into a single sociotechnical entity, saying things like, "*Civilization* is one of my really favorite games of all time," as opposed to naming a particular version or release of *Civilization*.

In OAIS, this relationality can be mapped as Context Information within the Preservation Description Information entity. What is meant by Context Information is unclear to some interview participants; its description in the OAIS literature is similar to archival bond. Therefore, a repository can only express this Significant Property as Context Information if it holds enough games to demonstrate how a particular game relates to others.

Many interviewees acknowledged that preserving the affective and social aspects of games is a most challenging task. Playing games in arcades is a fundamentally different

experience than playing at home; these locations impact game design, for example the simplicity of original *Donkey Kong* versus the deeper interaction of *Super Mario Brothers*. The creators and players describe the school-setting of the earliest *Oregon Trail* and *Carmen Sandiego* titles as a significant property. The need to understand the time and place in which a game was made and/or played might be easiest to understand with a game like *September 12th,* a news game predicated on the events of September 11, 2001. The twin difficulties are encapsulated by two quotes from different game developers. The first, a contemporary developer, said, "…it's hard to differentiate between what is like your nostalgia and what is sort of useful, right?" A second quote, from a developer of a game series that is no longer in production, said, "So you really have to sort of capture the essence of the time. Now I don't obviously have a good answer for that, but somebody should think about it."

These Significant Properties do not fit well within OAIS. This may be because documenting this type of information in relation to a particular object has not always been seen as the province of the archive itself. In some cases, the preservation of some non-code significant properties of a game is more desirable than preserving working code itself: a video of game play, a textual narrative of a walk-through may better capture the experience than working copies of obsolete technology. In fact, these expanded descriptions of what might be significant about a game challenge the very assumption that a baseline for a game's authentic preservation is its functionality or playability.

6.5 Conclusions

Some significant properties, as suggested by interviewees from my OAIS research and PVWII respectively, fit well within the existing OAIS entities. For others, one could

argue for their inclusion within existing entities although it may mean stretching the capacity and meaning of these entities beyond what was envisioned by the designers of OAIS. This latter is not to indicate that such actions would be wrong: indeed, it is the role of the reference model to inform things in the future which likely entails moving into spaces the original authors could not envision. Further, as I have argued throughout this dissertation, OAIS is more than a single standard defined by a single document; it is a complex of authors' and users' scripts and other standards like PAIMAS and PREMIS. Within the larger sociotechnical system, entities and terms take on meanings that, while not exactly the same as those in the OAIS glossary, are still valid.

Data from PVWII suggest that social and affective attributes of games are considered significant by designers and players. These significant properties are largely expressed as relational properties: they obtain in relation to objects, events, spaces, and times outside the object and often outside the archive or repository. These relationships are also nuanced in nature: certain properties are more important than others, or are only important in certain cases (for example, to particular Designated Communities). In fact, the situatedness of significant properties suggests that, for popular content like video games, the notion of Designated Communities is too vague and it is more important to think about archived objects in the context of Ranganathan's (1931) third law: every [digital object] its [user].

Some significant properties fit within OAIS; others will require either new metrics or changes to the existing standard. These findings are echoed by similar comments from OAIS interview subjects, and this is all the more pertinent given both the variety of participants in this latter study and the fact that the conversations I had with them were

very different in nature and subjects from the PVWII interviews. The similarities between the two data sets speak to the salience of these themes.

What was surprising about this project was just how much data I struggled to map to OAIS: my original hypothesis when I began this mapping project was that all Significant Properties should fit within OAIS, given its commitment to changing Knowledge Bases over time. For example, the process of documenting context is nothing more than moving additionally pre-inscribed affordances of a digital object into the circumscribed setting of the Archive. In the language of scripts, adding information from the Knowledge Base of the Designated Community to the AIP as additional documentation is taking what is normally afforded to the actants forming the Designated Community and pulling it into the AIP. This is merely an extension of an on-going balancing act, of finding the line for sufficiency in deciding how much to document: this is precisely why I call for the creation of metrics to help drawing these artificial boundaries so that this work can be made machine-actionable for work at scale. My conclusion is, therefore, that all significant properties do not fit within existing OAIS entities and I echo the calls of other preservation scholars that changes are needed in the ways in which we think about the responsibilities of repositories, especially given the potential for distributed digital preservation in linked data environments. Additionally, I posit that these difficulties will be exacerbated in areas where OAIS already does not work as well. A couple of interview participants noted that the scripts within OAIS presume a level of infrastructure. While Seles (2016) demonstrates how this plays out in situations where Archives are located in geographical regions where the legal, electrical, and network infrastructure are missing, some of my interview participants pointed out

that, even in wealthy first world contexts, institutions wherein preservation is not a primary function will lack many of the structures presupposed by OAIS.

I argue that the ability of the OAIS authors to choose what not to deal with explicitly within OAIS connotes a position of relative power. This is what prevents significant properties from becoming a dominant discourse in the field and serves to relegate them to comments like those by Webb, Pearson, and Koerbin (2013). There is power in being able to ignore something or to deem it inconsequential. If the gradations implied by significant properties are not acknowledged to exist, then the Information Packages saved by the Archive become the sole representative of a particular historical moment without needing to acknowledge other views for other audiences. I argue that a concept like significant properties opens the space for broader participation in official memory making processes. This again ties to audience, the power of the Designated Community, and the power the Archive has in being not only allowed to but required to create it. Further, when Designated Communities are not explicitly stated or created, as interview participants admit is often the case, what is the default? This raises the privileged invisibility of the presumed audience, the dominant population presumably informed by assumptions of the OAIS makers and Archive makers (Akrich, 1995), and of the assumption that what is selected by the archive encompasses what is most important about an object and its place in time and space and history.

I situate claims that significant properties are situated and sometimes outside the digital object and it computing environment within a growing body of archival science literature referenced in Chapter 2 that speaks to the situatedness of archival content and what is needed to contextualize it (MacNeil and Mak, 2007). The juridical and legal

undercurrents of archival conceptions of authenticity are balanced by work in practice, where archivists understand that evidence, for example, aids in interpretations about the world (Caswell and Gilliland, 2015) and that archives may have the role of preserving mnemonic devices in addition to evidence (Smith, 2003). What is necessary is for digital preservationists to decide whether what is wanted is particular bits of information or impressions of the past. This is perhaps the difference between the archive and the museum, for example, and the fact that a single tool like OAIS needs to provide guidance to all is demonstrative of its inherent inadequacy or at the very least a sign that memory institutions should[50] begin to reconceive their role in a digital future.

[50] I say 'should' here because this is a prescription on my part, not a techno-deterministic statement that digital technologies will necessarily lead to a convergence of memory practices. Rather, this is a space to counter some of the hegemonic and imperial discourses that manifest in different ways in different types of institutions: this loosening of role signifies a loosening of boundaries in which it is possible to change some problematic practices.

CHAPTER 7

CONCLUSION

Introduction

In 2017, OAIS will be revised again according to the timetable of required periodic updates by the International Standards Organization. This will be its third go around of revisions as an official ISO standard, with previous revisions having occurred in 2007 and 2012; obviously it has been undergoing informal revisions as long as the idea of what it would become was well formed enough that early drafts could reasonably be seen as the same sociotechnical object and not aborted predecessors. The European OAIS author with origins in space science gives 1992 as the year for this.

Within my interviews, participants disagreed as to whether or not these revisions resulted in substantive change. The manager of digital preservation at a European national library said, "there are some important concepts that are introduced in the new one." Other participants allow for small changes to particular concepts or terms, saying things like, "So, there were some small changes to the OAIS when it got renewed, but nothing huge, right? A couple of things to figure and some stuff; a couple of new terms" or "like for me the changes in the Magenta Book were not a big deal or of any huge interest to me. I started using the Magenta Book as my reference, because it was the new reference, but the changes in there did not have a significant impact on, you know, how we would build systems." Others, however, argue that new changes do not really arise as part of the official revision process: a digital preservation consultant at a boutique US firm said, "It doesn't -- because it was for so long the accepted practice it hasn't been updated to reflect new developments I would say."

These quotes indicate two things. One is that OAIS as a sociotechnical complex is so large that it engenders considerable inertia. This means that it does not and cannot change substantively in a short space of time, with a handful of revision periods constituting a short space of time. The second issue at play here is whether or not it would be desirable for OAIS to change substantively. To be sure, practitioners want to see changes in OAIS. David Rosenthal (2015) has been vocal about this, as have his partners in the online space designed to elicit public commentary about OAIS ahead of its revisions. The desire for wide community participation in revisions was palpable: one interview participant expressed a desire for a public forum to discuss potential changes, and since I interviewed her just over a year ago, she has been instrumental in creating such a space. There was a feeling among interview participants that 'the community' was not well consulted for revisions, with a digital preservation and curation scholar at a US public university saying that according to his informal observations, "There were very little… community input, as far as I can tell." Additionally, within my data there is a well-populated coding node about OAIS needing to change.

Yet for all this desire, there is a recognition that as a foundation of the field of digital preservation, in addition to being a keystone in the sociotechnical network that includes implementation documents, professionals, and related standards like XFDU, 16363, PAIMAS, etc., it is not perhaps wise or possible to affect radical changes to OAIS itself. It would be rather like pulling a key bottom block from a Jenga tower. A digital preservation and curation scholar at a US public university described the situation thusly, "But, for the

most part, the presumption is these are the terms as people are using them, so we probably shouldn't make too many changes to them..."

So what does this mean for the 2017 revisions and beyond? One Canadian-based digital preservation scholar said:

> So, I mean, I know that in the last review cycle, essentially nothing was changed, I mean almost nothing… Well, yeah. The change that would be really beneficial would be almost a different model in that perspective.

In other words, where we go from here is some place different. While the same participant argued that OAIS does not talk much about foundational concepts, I have argued throughout this dissertation that OAIS constitutes a real part of the foundation of the profession and sub-discipline of digital preservation. As multiple participants have said, the field has reached closure on the subject of OAIS. A digital preservation and curation scholar at a US public university compared it to gas pedals on the right and brakes on the left. He finished by reminding me, though, that closure does not last forever.

From here, big changes to OAIS and to the field of digital preservation more generally require building new and different things on top of OAIS, and letting OAIS sink further into the background; in other words, focusing on the top of the Jenga tower. As the same scholar above said, who knows if the field will be using OAIS terms in 20 years' time; perhaps, certain terms may be the only vestige of it left at that point, particularly if standards like ISO 16363 do not succeed in engendering new interest in OAIS. The other solution to the grumblings about the inadequacy of OAIS (and the "Twitter bashing") is to knock the Jenga tower over and start from scratch. More than one participant expressed a desire to see someone suggest something completely new, but no one could offer a viable alternative

that is currently in existence. I suggest that this latter, while potentially beneficial, is enormously difficult to conceive of given the pervasive influence OAIS has had on the development of the field. As the throw away remark about the relationship between Archivematica and OAIS earlier in this dissertation shows, even when people do not see OAIS, it is often there. Developing new foundations for digital preservation is like trying to build a new boat from the one you are currently floating on, or imagining a functional non-capitalist alternative while living in the current neoliberal moment: the scripts are embedded so deeply it is hard to be aware of where they start and stop, and it is hard to get away from them.

This conclusion will explore alternatives and futures of OAIS and digital preservation. I begin with brief summaries of the major findings of the previous chapters. I will then talk about answers to the questions I posed to frame this research: both the formal research questions that shaped this entire project and the questions that arose as part of the theoretical framework that has guided how I have understood the results of the project. Finally, I will discuss the futures: futures of OAIS and the field of digital preservation, and research futures that this project suggests.

7.1 Dissertation Findings

In this section, I highlight some of the major findings described throughout this dissertation. I suggest, based on the literature, that OAIS is not a static object but is better seen instead as a sociotechnical complex that encompasses numerous other standards, technologies, and actors. OAIS stems from the space sciences although it borrows heavily from archives to purchase the authority and connotations of longevity that are part of popular archival discourse.

In Chapter 4, I argue that OAIS conceives of a variety of professional roles within the Archive Management entity, and when OAIS gets deployed in institutions, employees of the organization whose labor does not fit within the entities of the functional model are disempowered. This happens not only because of the privileging of digital material in some environments, but also because a lack of fluency with OAIS language among some workers serves to exclude them further within organizations in a way that mirrors how they are literally excluded from the model.

Aside from the internal staff of an Archive, OAIS speaks about both Producers and Consumers, and in many cases, interviewees suggested that external content producers and users are not included within the linguist fold of OAIS, while OAIS's inherently alien terms are referred to in some relationships between the Archive and Producers/Consumers precisely so the "foreign" language can be employed to connote authority. The fact that OAIS calls itself an Archive also lends it authority in exchanges with donors and users, and it borrows from other archival discourses that mean it is subject to some similar limitations to those faced by traditional archives.

While OAIS uses a number of scripts from archives, none of the participants I asked about the causes of the geographical differences in attitudes to OAIS suggested the difference between modern archival practice in Europe and the US as a cause. One interview participant, a digital repository manager in a US museum, explained the use of an archival science-based tool, Atom, at his institution, saying:

> Atom originally was designed for archival collections, so it was more – and specifically European archival terms, so it was like fond and sub-fond and I don't

> remember the other terms but not even – I don't even know if you can convert it to American archival language…

His language indicates an awareness of the differences in practice wrought in part by the Jenkinsonian traditions in US archives, yet this was never suggested as a potential cause for the disparity in attitudes towards OAIS between Europe and the US. Instead, the two overarching causes that were suggested by interviewees related to the centralization of funding in Europe and the related centralization that governs preservation projects. This results in much larger scale investigations than in the US. I suggest that at the root of this centralization are politics of European unity through shared cultural initiatives, part of European Union politics of "unity in diversity."

Chapter 5 explored the term *Designated Community* as it is defined literally in the model but also as it is described by the authors; perceived by users; and implied by the ISO 16363 audit standard. The fact that there are so many gradations in how this term is employed supports my contention that OAIS is not just a document but instead a sociotechnical system. One of the most significant findings of this dissertation is the relationship between Designated Communities and actual users. In the chapter, I discuss why Designated Communities are not users and when: in OAIS they can be entirely constructed and not a description of real users, while for ISO 16363, real users are all but required to check that an Archive has made its content "understandable" to a member of the Designated Community. I also contend that the concept of Designated Communities is fundamentally at odds with certain memory professions, particularly those in institutions that serve diverse populations and that handle popular content. While the OAIS authors offer an explanation about the genesis of the Designated Community requirement, their

explanation in fact means that they are telling libraries and archives that they should be actively inscripting their archive in precisely the way Akrich uses this term. OAIS is essentially requiring these memory institutions to state their intended and preferred community and implicitly build practice around that. Implicit practice becomes explicit at the point of the audit. This process gives power to the Archive and favors the Designated Community users. These are the ones who get helped and get usable access. The non-Designated Community members do not. Many organizations have not formally stated a Designated Community, and this has power implications because the default tends to be an empowered one: an assumed dominant population often reflective of the Archive makers themselves. The default is rarely if ever a subordinate default. Being able to change the Designated Community is another manifestation of the power of the Archive. Whether or not an Archive wants to or actually does change Designated Communities, the salient matter is that OAIS allows them to: they can and every time they do, this has some constitutive power and creates a new self-perpetuating discourse.

In Chapter 6, the difficulties of the long-term become more apparent. It is hard to predict the future, and significant properties have different potential meanings for digital objects in the future depending on the user and social context. At least, the ability to represent what was significant at one period in time might be more challenging in future.

Some significant properties fit well with OAIS, while others do not. The ability to exclude significant properties from OAIS denotes power on the part of the authors. This term is symbolic of the tensions between memory practices and the sciences. This is interesting, particularly in light of Van House and Churchill's (2008) contention that, in spaces where the two must work together, the advantage of resources and influence lies

with the sciences. There is also tremendous power in the ability to erase differences, in getting to choose a single canonical form rather than allowing for differences. By relegating significant properties within OAIS to elements that are things already inherent to the digital object itself, the authors of OAIS preclude bringing into the archive a mechanism for the interpretation of differences. Such a mechanism would stem precisely from how the scripts in technologies are unpacked in difference places of use rather than what is contained within the technologies themselves, still bundled up: as Akrich says, it is not even possible to see the scripts if the technology is never deployed. Only through using something do its seams become visible. OAIS, in excluding significant properties, excludes the seamfulness that is required for more pluralistic archives.

In conclusion, power relationships are manifested throughout the OAIS documentation and in the ways in which it gets deployed in local contexts. Requirements like the construction of a Designated Community, the need to speak in particular OAIS language, and the ability to exclude significant properties place considerable power in the hands of the Archive. The directive to script preservation and dissemination services for a preferred set of users, the Designated Community, encourages the reification of existing imperial tendencies that existed in pre-digital, analog archives.

7.1.1 Research Questions

This research project began with two research questions, and here I will describe some answers as suggested by the findings. First, I asked about the values contained within OAIS and what methods or practices were prescribed by it. A script built into OAIS is that it is above all things flexible and all-encompassing—multiple interviewees declared that anyone should be able to use it for anything. Yet this is not what people find, even some of

those same interviewees—for whatever reason, when OAIS gets implemented, there are things that people do not like about it and, more importantly, things that they identify as not working. This is a source of tension: people say in the same interview that it should work for everyone while spending considerable time discussing the cultural disconnects between OAIS and their daily realities. I argue that part of this tension results from the role that the OAIS sociotechnical complex has played on the genesis of the sub-discipline of digital preservation more generally: OAIS and the evolution of digital preservation as a profession with boundaries, standard practices, curricula, etc. happened concurrently. OAIS is bound up in the very basic ways that people think about preservation, even when they are not aware of it, and this makes looking for real alternatives very difficult. This is problematic because of some of the practices that OAIS encourages or mandates among practitioners, mostly notably to script daily practice to favor some users over others.

Secondly, I asked: in what ways has the adoption of this model as an organizational system for the preservation of digital content in library, archive and museum spaces served to challenge or reproduce the hierarchies and discourses of traditional archives and memory institutions?

While OAIS is a reference model, and therefore not a technology in the sense this word is most commonly deployed by its own creators, it is still a technology of control in the way in which this term is employed in post-structuralist critiques like Foucault's theory of governmentality. Cast in this light, it is necessary to question the ways in which OAIS is a reflection of a need to govern or control that 'facilitates the exercise of…power' beyond simply being widely adopted for its great efficiency or productivity; in fact, the number of cases wherein inferior technologies win primacy wars abounds (Marglin, 1978, page 88). In

considering OAIS in such a light, its machinations can be seen as techniques of capital as well, in the ways in which digital preservation practices tends to create new silos and hierarchies of work as I described in Chapter 4. Gorz described precisely the fragmentation of labor I saw in some locations in Marxist terms:

> Capitalist techniques were not meant to maximize the production and productivity *in general* of all workers *whatsoever*. Instead they were to maximize the productivity for capital of workers who had no reason to give of themselves, since an enemy had dictated the aims of their production. To make them bow to this will it was not enough that they should lose the ownership of the means of production… They had to lose what their professional and practical knowledge and skills had given the so far: the power to run the machines without the assistance of a hierarchical corps of engineers, technicians, maintenance experts, foremen, and so on. (Gorz, 1978, page 56)

This is borne out by interview data with both OAIS practitioners themselves as well as those working in such institutions whose daily work does not necessarily focus on the digital. Smith (2003 and 2007) argues in much the same vein as Schaefer (Owens, 2014) that the way to address this kind of divide is for all participants in an organization to learn to understand the technologies in use. This is especially true for subject and analog specialists who are anxious about the perceived deskilling happening with their replacements: new technology-savvy employees that lack content-specialist knowledge or, in the case of some employees, automated cataloging and descriptive metadata tools that obviate their jobs entirely.

In fact, the Marxist critique plays out in interview data in other ways as well. One of the stark features of OAIS adoption is precisely the functionalist essentialist ideology that accompanies its adoptions: it was the only tool that supported the inevitable needs of the institutions in the late 90s and early 2000s, and in most cases, no alternatives were considered. Common essentialist narratives about the need for everyone to think, and by extension eventually become, like a digital preservationist are expressed by multiple interview subjects with differing job titles and descriptions.

It is important not to overstate the disruption caused by digital technologies in traditional memory institutions and practices. To do so smacks of a kind of techno-determinism. Yet a disruption has occurred, and it is important to question the eventual outcomes of this disruption precisely because it is still taking place and outcomes are not decided yet: this can be seen in the interview participants who describe how nascent their digital preservation efforts are. Archives may find a way to continue as they have for thousands of years with digital technologies.

> Things operate in circuits of value, which themselves can be spatially located and temporally varying. Dear treasures in one part of town may be garbage, and laughable, in another. Indeed, Pierre Bordieu noted that there is perhaps no more telling indication of cultural capital than the ability to bestow value where there had previously been presumed to be none. (Acland, 2007, page xv)

The way in which the Designated Community concept in OAIS has been employed is a perfect example of this: it takes a process that was implicit in previous archival practice and explicitly prescribes it: in fact, archives are "rewarded" with official certification for excluding users in creating a well-defined Designated Community.

While this dissertation is not an explicit gender analysis, I would be remiss if I did not mention some of the gendered aspects of the deployment of this technology and how these too lend themselves to the power dynamics already identified in this dissertation. There are feminized and masculinized discourses within digital preservation practice, particularly revolving around notions like authenticity. Authenticity is a central concern of digital preservation practice and much of the alternative constructions for significant properties within OAIS are aimed at establishing just this. Smith cites Walter Benjamin's notions of authenticity in archives and describes Benjamin's view that authenticity is a quality that is intrinsic to the object itself. Indeed, she makes the comparison that "Authenticity is like virginity: it can be lost, but it can never be regained" according to Benjamin (Smith, 2003, page 176). Yet the dominant discourses about the preservation of digital content and its fragile authenticity are dominated by men: these discourses and technologies originate in techno-masculine spaces like space data. In eschewing a term like significant properties that comes from the more feminized memory institutional sector, the OAIS sociotechnical complex constructs a discourse in which Authenticity is more or less inherent to an object itself. It is at least inherent to the AIP and the Archive of which it is a part: it is fragile, breakable, and must be guarded by the archive.

There are distinct power relations even in the poorly defined space of DIPs. Galloway notes that the quest to make a perfect visualization manifests a desire for a representation that is imperial itself (Galloway, Lovink, and Thacker 2008). I suggest a relationship here because in order to make good DIPs in the future, as the examples of Twitter and Carmen Sandiego demonstrate, Archives will have to get creative. And yet, it will be considerable work on the part of the Archive to ensure that the solutions do not

compound the existent imperial tendencies of archives. Such DIPs will need to connote mnemonic devices that express the affective experience of using a digital technology at a particular place and time as well as show the Archive's own values and scripts. This is an avenue of necessary work in the future, because the default alternative is merely a continuation of existing imperial discourses within archives.

The evidence in the interviews is that people are using OAIS more than they realize. It is easy to overlook the influences of OAIS like those mentioned above precisely because OAIS is pervasive and foundational to field of digital preservation. There are ways in which OAIS could be employed for purposes that resist the construction of the power relations described above. There is also a need to find a way to develop some salient alternatives to OAIS: I see this as necessary not only from a utopic desire to see a more plural and democratic form of archive, but also as a practitioner myself. I recognize the difficulty of the labor that goes into record-keeping and memory-making, and alternatives are necessary in the form of concrete metrics and machine-actionable standards that assist the practitioners in their daily work. In these last sections, I discuss the futures that might assist in this area: futures for OAIS and research beyond.

7.2. Futures of OAIS: Alternatives and Closure

> Invention involved the development of new technological artifacts by inventors and provided the basis on which economic transformation might proceed. Innovation on the other hand involved introducing these new technological artifacts into a routine economic cycle of production and consumption and thus reorganizing the world around these new inventions…. Both invention and innovation involved dealing over

> time with "messes" (i.e. systems of ill-defined but interrelated problems) and resolving the messes in some fashion… (Kaghan and Bowker, 263)

Multiple people noted in interviews that the uptake of OAIS was largely dependent on timing: it happened in an era where many institutions had begun to realize they needed assistance in thinking through the process of preserving the growing onslaught of digital materials. I do not suggest a causal or inevitable relationship, either between "the times" and OAIS or vice-versa, but suggest that conditions including time did play a large role in its becoming the invisible foundation it has become today. Because of this, alternatives are hard to find.

The very notion of standards is that they exist as a process of closure and stabilization of boundaries (Ole et al, 2006): they purposefully close off infinite alternatives and focus the attention of a field in a particular direction. Lee (2005) notes in his work that closure has been reached on the subject of OAIS, and indeed had been as early as the early 2000s. This explains one reviewer's comment to me about OAIS being "uncontroversial." Closure is also neither static nor permanent. The OAIS sociotechnical complex is continually redrawing its boundaries to maintain its position. It does so through the creation of new standards and new tools. This is purposeful: this is precisely the role of the reference model as it seeks to inform future practice and derivative, industry specific standards, protocols, and solutions more generally, in the way the Lee (2005) and Cargill (1997) describe. OAIS has effectively closed a space that leaves very little room to imagine alternatives—people could not think of alternatives where they were thought necessary. Others thought that nearly anything could be fit within OAIS so that even 'alternatives' would not be too alternative, and I would argue this is partly a result of the OAIS discourse

of flexibility. The boundaries around OAIS are flexibly drawn and actively maintained in a way that allows them to continually encompass changes—in a super concrete and simplistic way, this happened in 2012 with revisions that addressed concern for the preservation of computing environments and the rise in popularity of emulation. It happened again with the release of ISO 16363 and ISO 16919 and the subsequent push by European memory organizations and political bodies to require certification through these standards. Like capital, this becomes a very hard space in which to theorize alternatives. This kind of pervasiveness requires a paradigm shift in order to create the intellectual space in which to conceive of alternatives (Kuhn, 1996)[51].

When the majority of the power lies with one set of actors (I argue that the dominant actor here is still CCSDS and systems science), the solutions imagined by dominant actors for others are inflected by assumptions of otherness based on self-identity. This approach to design is all the more pertinent when the dominating central actor (Law, 1999) is one that already yields more power in heterogeneous negotiations (Van House and Churchill, 2008). In fact, the pervasive inability of interviewees to name or imagine alternatives to OAIS suggest that it has become a black box. If a black box is defined as, "…a frozen network element, often with properties of irreversibility" (Walsham, 1997, page 468), the language of freezing constitutes an absolute closure of boundaries, the absolute zero of standardization processes.

And yet, there is an inherent tension between what is expected of a reference model and complex cases like digital preservation, wherein the stakeholders are so diverse and

[51] I used the language of Thomas Kuhn in describing practices in the sciences without making any claims about whether preservation work is a 'science'.

the relationships so frictional that closure in the classical sense cannot be reached. Ole et al (2006) describe standards in such cases:

> These cases are described as worlds too complex to be closed and ordered according to a single mode or logic. There is only partial order, interacting in different ways, or interconnected and overlapping subworlds, ordered according to different logic. The interconnectedness of the subworlds means that while one is trying to make order in one subworld by imposing a specific logic, the same logic is causing disorder in another: each order also has its disorder (Berg and Timmermans 2000; Law 1999). (Ole et al, 2006, page 566).

This is all the more pertinent given the heterogeneity of stakeholders engaged in digital preservation, and that digital preservation is a meta-sub-discipline within a meta-discipline. What is functional in one part of this vast pantheon is less so in others. The shift, as described by Ole, is a move towards adhoc collections of local solutions, and OAIS was developed within such an environment: it was developed openly by a large consortium rather than by a single standards body while simultaneously aiming to standardize approaches as a replacement for the adhoc approaches that existed prior to OAIS in the late 1990s and early 2000s. This is where the space of alternatives lies. While people could not articulate alternatives, one interview subject, a Canadian-based digital preservation scholar, suggested that some solutions lie within careful systems design. The implication here is perhaps that such sweeping standards are not serviceable now that the field and digital technologies have developed: for particular projects and places, the design needs to be more or less focused on local needs. OAIS is very functional in certain disciplinary settings. While it is possible to make changes to it to make it functional elsewhere, perhaps

such changes would render OAIS less functional in the spaces where it works right now. In this case, the alternatives are really that: other standards that exist alongside OAIS to serve the spaces it cannot. This could include cultural heritage spaces, but it could also include distributed networks and distributed content, as well as projects like preserving the intellectual labor of digital collections without duplicating the work of preserving the objects themselves. All of these areas of researcher, however, are in nascent stages if work has begun on them at all. For real alternatives, it is necessary to look to future research. I suggest some possible avenues below.

7.3 Future Research Avenues

In keeping with the hybrid nature of the discourses that inform OAIS, these research avenues are similarly hybrid. I suggest the need to challenge the assumption within archival spaces that artificial boundary making is undesirable; rather, embracing the artificiality of the task of rendering something like a virtual world into a static object of preservation requires good metrics that celebrate the interpretive work required of such a task. I also suggest that while documentation is a good stopgap for preserving those affective experiences that we cannot currently render technologically, we should continue to look for better solutions. One space that needs particular attention is the *Knowledge Base* term within OAIS.

7.3.1 The need to document beyond the digital object and beyond

This research notes the fundamental ways in which OAIS is not sufficient to document the kinds of relationships and affective experiences associated of digital objects that make them significant cultural items. I argue for the need for more context. Figuring out what constitutes sufficient context is difficult precisely because in digital preservation

there is no starting point and no stopping point. As preservationists, we must construct both of these for any digital preservation project and in so doing we create a punctualization, an archival package that represents an artificially static object in place of an ever-moving network of actors, objects, and technologies. This type of language is uncomfortable for many memory practitioners who see their roles as guardians rather than as creators. The word "artificial" applies a negative connotation to a creative act. I argue against this application and suggest that a first step towards realizing this kind of practice is to embrace the fact that the work preservationists do with digital objects is inherently artificial: it is artificial to name a fixed object in place of something like a video game or virtual world, and representations of this in the future will be artificial as well. But if this type of work can be seen less under the traditional archival umbrella of evidence and more along the lines of imparting an affective experience, the field of archival science can open itself to the possibilities of better, more inclusive work through networked technologies.

The need for metrics to define the creation of artificial boundaries should be obvious. To begin with, the very nebulousness of this task makes this type of work difficult in computational terms. This dissertation is another call for the end of precisely the type of simplified yet arbitrary guidelines that dictate that web crawlers, when documenting the Internet, stop at a depth of 3 links, for example. In suggesting more nuanced approaches, it is necessary to develop tools that assist practitioners in their necessary daily work of bounding and documenting. Acknowledging the artificiality of this type of work is important, because it promotes a transparency of process and denounces claims to neutrality or naturalness in the work we do: in fact, archivists are always engaged in constructive and creative processes of reinterpretation, and the goal here is to make that

explicit in the digital realm (MacNeil and Mak, 2007). In so doing, it is equally important to avoid essentializing the objects of preservation and to avoid creating narratives that serve to reinforce normative practices and relationships without question. We need to avoid creating imperial visualizations.

While one of the very benefits of such an approach would be to bring fragments and ephemera away from the margins, such work must be undertaken in a way that appreciates the ability of the fragment to purposefully resist incorporation into larger, often dominant, narrative structures (De Villiers, 2012). This means we must also allow certain things to be hidden, and indeed an archives and special collections librarian at a small liberal arts college in the US expressed discomfort in the use of tools like FTK and the potential to uncover things which donors did not wish to share. There is an ongoing tension between the desire for transparency and the simultaneous desire not to compel exposure, particularly as this has been forced onto certain populations more violently than onto others. In developing metrics of expanded context we must avoid the use of discourses that "make confession and disclosure … less than liberatory." (Cvetkovich, 2003; Sedgwick 1997).

7.3.2 The Fetish of Documents

And while documentation is good, we need to move beyond it. There are limits to what can be done with documentation: it is not a perfect solution.

> This shift from an approach grounded in what something *is* to how something *works* changes the analysis of material evidence from iconographic reading to indexical reading, leading us into the lifecycle of production, use, control, resource consumption, labor, cost, environmental impact and so on — so that an

> artifact's materiality is read as a snapshot moment within continuous interdependent systems. (Drucker, 2013)

This compelling statement from Drucker echoes what I have argued numerous times throughout this dissertation: important behaviors related to the object may or may not be inherent to the object itself. Concrete examples might be something like the sound and time it took to flip and load a floppy disk while playing a game in 1985. The materiality of this experience is lost when using an emulated or migrated copy of the executable code. In some ways, it probably should be: in digital preservation terms, either of those preservation strategies, emulation or migration, is more tenable long term than keeping working machines for every kind of media. So how do we create such "snapshots"? I argue that this requires work in the area of DIPs, beyond the formal description of information exchanges created by the E-ARK (2015) project.

Current thinking about such topics in digital preservation amount to what I call a *fetishization of documentation*, which is closely related to the "panacea of metadata", a phrase Giaretta himself uses in the introduction to his advanced digital preservation text (2011). We are in an age of 'supermodernity' that is characterized by documentation that arises when linked data and digital archives are not the panacea they were assumed to be in the realm of preservation and memory (Acland, 2007). Documentation is an important digital preservation strategy, like emulation or migration. The assumption is that when something is too complicated for its bits to properly maintained, or when it is admitted that even saving the bits is not sufficient, the current answer is to document the object: fill the open text notes field in metadata standards with descriptive information, take screenshots and video. This impulse results in an interesting tension that arises between the graphical

moves away from the document in the digital era and the simultaneous rise of the document/metadata as the solution to all our problems about replicating behaviors, acting benchmarks, or connoting significance.

Bann says, "History certainly abhors continuity" (in Ernst, 2013, page 45). There is simultaneously a tension between the push in archives for continuity and the need to recognize the space between ourselves and objects of the past. Documents as discrete, non-continual entities represent this inherent discontinuity. Tying this to OAIS, I see the need for more work in the area of the Knowledge Base. Concretely, how do we document this and track its changes without this activity becoming an all-consuming sort of project? A real exploration of what that looks like must take into account the inherent discontinuity of such a project as well as the general inability of human beings to see changes as they happen.

7.4 Conclusions

The research projects proposed above all function within a discursive space created by OAIS. This is partly because of a pragmatic recognition that OAIS and the field of digital preservation are so tightly bound that, given the resources available, working within these structures makes some sense. It was this kind of pragmatism that drove me to study OAIS as an exemplar of digital preservation to begin with. Further, I believe with the right subsequent standards, it is possible to use OAIS to do some new and different things. Many of my arguments take as an underlying assumption the fact that significance is situated, and I believe this is an unintentional script within OAIS as illustrated by concepts like the Designated Community and Knowledge Base: OAIS recognizes how important situations are, and this is a space of potential. OAIS stacks power very much in favor of the Archive,

and the discourses common within digital preservation stack power very much in favor of OAIS. This being the case, it is also possible to use the power-laden parts of OAIS to create more inclusive projects. This might include metrics for the creation of Designated Communities that purposefully seek out sub-ordinate populations for active inclusion. None of this is possible, however, without being very aware of the politics and power relationships that undergird this technology which is used so pervasively. Without understanding the complicated term *Designated Community*, for example, it is difficult to render OAIS a resistive technology. As such, the primary take away from this dissertation is an un-boxing of OAIS: I have attempted to lay bare elements of this profession that are normally decentered and invisible. This type of investigation is necessary and must be on-going: without active intervention in this space, it is entirely possible that the digital futures will carry the same scripts as the analog past.

BIBLIOGRAPHY

Acland, C. (ed) (2007). Residual Media. Minneapolis, MN: University of Minnesota Press.

Adams, M. & Brown, T. (Winter 2000). Myths and Realities of the 1960 Census. Prologue Magazine, National Archives. Retrieved from: http://www.archives.gov/publications/prologue/2000/winter/1960-census.html

Akrich, M. (1992). "The De-Scription of Technical Objects". In *Shaping Technology/Building Society: Studies in Sociotechnical Change*, eds. Bijker, W. and Law, J. Cambridge, MA: The MIT Press.

Akrich, M. (1995). User representations: Practices, methods and sociology. In Managing technology in society: The approach of constructive technology assessment, eds. Rip, A., Misa, T.J. and Schot, J. London: Pinter Publishers.

Akrich, M. & Latour, B. (1992). "A Summary of a Convenient Vocabulary for the Semiotics of Human and non-Human Assemblies." In *Shaping Technology/Building Society: Studies in Sociotechnical Change*, eds. Bijker, W. and Law, J. Cambridge,MA: The MIT Press.

Ayoung, Y., & Tibbo, H. (2011). Examination of Data Deposit Practices in Repositories with the OAIS Model. IASSIST Quarterly, 35(4), 6-13.

APARSEN (2015). Report on a Common Vision of Digital Preservation. Retrieved from: http://www.alliancepermanentaccess.org/wp-content/uploads/sites/7/downloads/2015/02/APARSEN-REP-D11_5-01-1_1.pdf

Ball, A. (2006). "Briefing Paper: The OAIS Reference Model" UKOLN: University of Bath.

Balsamo, A. (2011). *Designing culture: the technological imagination at work.* Durham, NC: Duke University Press

Barrett, M. (1991). *The Politics of Truth: From Marx to Foucault*. Cambridge, UK: Polity Press.

Bastian, J. (2003). *Owning Memory: How a Caribbean Community Lost Its Archives and Found Its History.* Westport, CT: Libraries Unlimited.

Bates, M. (2015). The Information Professions: Knowledge, Memory, Heritage. Information Research: An International Electronic Journal, 20(1).

Baudrillard, J., & Glaser, S. (1994). *Simulacra and Simulation*. Ann Arbor, MI: University of Michigan Press.

Becker, C., Kulovits, H., Guttenbrunner, M., Strodl, S., Rauber, A., & Hofman, H. (2009). Systematic planning for digital preservation: evaluating potential strategies and building preservation plans. International Journal of Digital Libraries 10, 122-157.

Bennett, T. (1995) *The Birth of the Museum: History, Theory, Politics.* London, UK: Routledge.

Best, S., & Marcus, S. (2009). Surface reading: An introduction. Representations, 108(1), 1-21.

Bishop, A. & Star, S. L. (1996). Social Informatics of Digital Library Use and Infrastructure. ARIST 31, 301-401.

Boast, R. (2010). "Museum as Contact Zone Revisited," Museum Anthropology, 33:2: 56-70.

Bowker, G. & Star, S.L. (1994). Knowledge and Infrastructure in international information management: Problems of classification and coding. In *Information acumen: The understanding and use of knowledge in modern business* ed. Bud-Frierman, L. London, UK: Routledge.

Bowker, G. C., & Star, S. L. (1998). Building Information Infrastructures for Social Worlds-The Role of Classifications and Standards. Community computing and support systems, (1519), 231-248.

Boyle, J. (2003). "The Second Enclosure Movement and the Construction of the Public Domain." Law and Contemporary Problems 66:33, 33-74.

Bradley, K. (Summer 2007). Defining Digital Sustainability. Library Trends 56(1): 148-163.

Britton, C. (1999). *Edouard Glissant and postcolonial theory: Strategies of language and resistance*. Charlottesville, VA: University of Virginia Press.

Brock, A. (2009). Life on the Wire. Information, Communication and Society, 12(3), 344-363.

Caplan, P. (2004). 'DAITSS overview.' URL: http://www.fcla.edu/digitalArchive/pdfs/DAITSS.pdf

Caplan, P. (2007). "The Florida Digital Archive and DAITSS: A Working Preservation Repository Based On Format Migration." International Journal of Digital Libraries 6(4).

Cargill, C. (1997). *Open Systems Standardization: A Business Approach*. Upper Saddle River, NJ: Prentice Hall.

Caswell, M. (2013a). On archival pluralism: what religious pluralism (and its critics) can teach us about archives. Archival Science, 13(4), 273-292.

Caswell, M. (2013b). Rethinking Inalienability: Trusting Nongovernmental Archives in Transitional Societies. American Archivist, 76(1), 113-134.

Caswell, M. (2014). Seeing Yourself in History: Community Archives and the Fight Against Symbolic Annihilation. Public Historian, 36(4), 26-37.

Caswell, M., & Gilliland, A. (2015). False promise and new hope: dead perpetrators, imagined documents and emergent archival evidence. International Journal Of Human Rights, 19(5), 615-627.

Charmaz, K. (1983). "The Grounded Theory Method: An Explication and Interpretation." In *Contemporary Field Research*, ed. Emerson, R. Prospect Heights, IL: Waveland Press.

Chun, W. (2006). *Control and Freedom: Power and Paranoia in the Age of Fiber Optics*. Cambridge, MA: MIT Press.

Clark, K. (January 31, 2001). "The Politics of Schemas: Parts 1 and 2." In O'Reilly XML.com. Sebastopol, CA: O'Reilly Media, Inc.

Cockburn, C. (1983). *Brothers: male dominance and technological change.* London, UK: Pluto Press.

Consultative Committee for Space Data Systems (2002). *Reference Model for an Open Archival Information System (OAIS)*. CCSDS 650.0-B-1, Blue Book. Issue 1.

Consultative Committee for Space Data Systems (2004). *Producer-Archive Interface Methodology Abstract Standard (PAIMAS)*. CCSDS 651.0-M-1, Magenta Book.

Consultative Committee for Space Data Systems (August 2009). *Reference Model for an Open Archival Information System (OAIS)*. CCSDS 650.0-P-1.1. Pink Book. Issue 1.1.

Consultative Committee for Space Data Systems (June 2012). *Reference Model for an Open Archival Information System (OAIS).* CCSDS 650.0.M-2, Magenta Book. Issue 2.

Coombe, R. & Weiss, L. (2015). Neoliberalism, Heritage Regimes, and Cultural Rights. In L. Meskell (ed.) Global Heritage: A Reader. Hoboken, NJ: Wiley-Blackwell, 43-69.

Cvetkovich, A. (2003). *An archive of feelings :trauma, sexuality, and lesbian public cultures.* Durham, NC: Duke University Press.

Daniels, M. (2013). Where the OAIS ends: Archival principles and the digital repository (Conference Paper) Archiving 2013 - Final Program and Proceedings 2013, 10th IS&T Archiving Conference, Archiving 2013; Washington, DC; United States, 12-14.

Dappert, A. (2015). Digital Preservation Metadata and Improvements to PREMIS in Version 3.0. Webinar.

Data Seal of Approval: http://datasealofapproval.org/en/

de Jong, A., Delaney, B., & Steinmeier, D. (2013). OAIS Compliant Preservation Workflows in an AV Archive. White Paper, Netherlands Institute for Sound and Vision.

De Villiers, N. (2012). *Opacity and the closet: queer tactics in Foucault, Barthes, and Warhol.* Minneapolis, MN: University of Minnesota Press

de Vries, H. J. (2003). Standardization: A Business Approach to the Role of National Standardization Organizations. Boston, MA: Kluwer Academic Publishers.

Dean, M. (1999). *Governmentality: power and rule in modern society*. London, UK; Sage Publications.

Decker, A., Egert, C., Phelps, A., & McDonough, J. P. (September 2012). Technical properties of play a technical analysis of significant properties for video game preservation. In Games Innovation Conference (IGIC), IEEE International, 1-4.

Douglas, J. (2010). Origins: evolving ideas about the principle of provenance. In *Currents of archival thinking,* eds. Eastwood, T., MacNeil, H. Santa Barbara, CA.: Libraries Unlimited.

DRAMBORA: http://www.repositoryaudit.eu/

Drucker, J. (2013). Performative Materiality and Theoretical Approaches to Interface. Digital Humanities Quarterly, 7(1).

Duranti, L. (2010). "The appraisal of digital records: the return of diplomatics as a forensic discipline." International Conference on Surviving the Digital Wild Frontier. Singapore, Singapore.

Duranti, L. & Thibodeau, K. (March 2006). The Concept of Record in Interactive, Experiential and Dynamic Environments: the View of InterPARES. Archival Science, 6 (1), 13-68.

E-ARK. (2015). E-ARK DIP Draft Specification. D5.2, Revision 4.

Ernst, W. (2013). *Digital Memory and the Archive.* Minneapolis, MN: University of Minnesota Press.

Fairclough, N. (2006) *Language and globalization.* London, UK; Routledge.

Flanders, J. (2013). The Literary, the Humanistic, the Digital: Toward a Research Agenda for Digital Literary Studies in https://dlsanthology.commons.mla.org/the-literary-the-humanistic-the-digital/

Foucault, M., trans. Swyer, R. (1972). *The Archaeology of Knowledge*. London, UK: Tavistock Publishing.

Fuchs, C. (2008). *Internet and Society: Social Theory in the Information Age*. New York, NY: Routledge.

Galloway, A. R., Lovink, G., & Thacker, E. (2008). Dialogues Carried Out in Silence: An E-mail Exchange. Grey Room, (33), 96-112.

Giaretta, D. (2011). *Advanced Digital Preservation*. Berlin, Heidelberg: Springer-Verlag.

Giaretta, D., Matthews, B., Bicarregui, J., Lambert, S., Guercio, M., Michetti, G., & Sawyer, D. (2009). Significant Properties, Authenticity, Provenance, Representation Information and OAIS Information. iPRES 2009: the Sixth International Conference on Preservation of Digital Objects. California Digital Library, UC Office of the President

Gitelman, L. (2006). *Always already new: media, history and the data of culture.* Cambridge, MA: MIT Press

Gorz, A. (ed) (1978). *The Division of Labour*. Brighton, UK: Harvester Press.

Green, N. (1990). *The Spectacle of Nature: Landscape and Bourgeois Nature in Nineteenth-Century France.* Manchester, UK: Manchester University Press.

Hackett, Y., Underwood, W., & Eppard, P. (2008). Part One—Case and General Studies in the Artistic, Scientific and Governmental Sectors: Focus Task Force Report. International Research on Permanent Authentic Records in Electronic Systems (InterPARES), 2.

Hall, S. (1997). *Representation: Cultural Representation and Signifying Practices*. London, UK: Sage.

Halavais, A. (2009). *Search Engine Society*. Cambridge, UK: Polity Press.

Hamilton, K. (2015). Black Boxes. The Hedgehog Review.

Hedstrom, M. and Lee, C.A. (2002). "Significant properties of digital objects: definitions, applications, implications", Proceedings of the DLM-Forum 2002.

Herbst, P. G. (1974). *Socio-technical design: Strategies in multidisciplinary research*. London, UK: Tavistock.

IFLA Study Group on the Functional Requirements for Bibliographic Records. (1998). Functional Requirements for Bibliographic Records, Final Report. UBCIM Publications, New Series; v. 19. Munich: K.G. Saur Verlag.

International Standards Organisation, 'How Does ISO Develop Standards?' Accessible at: http://www.iso.org/iso/home/standards_development.htm.

Ippolito, J. (2003). Accommodating the Unpredictable: The Variable Media Questionnaire. In *The Variable Media Approach: Permanence Through Change*, eds Depocas, A., Ippolito, J., and Jones, C., 47-55.

Kaghan, W. & Bowker, G. (September 2001). "Out of Machine Age? Complexity, Sociotechnical Systems and Actor Network Theory." Journal of Engineering and Technology Management 18(3-4), 253-269.

Kendall, L. (2002). *Hanging Out in the Virtual Pub: Masculinities and Relationships Online.* Berkeley, CA: University of California Press.

Kerr, A. (2002). Representing Users in the Design of Digital Games. Proceedings of Computer Games and Digital Cultures Conference, ed. Frans Mäyrä. Tampere: Tampere University Press.

Kim, J., Warga, E. & Moen, W. (October 2012). Digital Curation in the Academic Library Job Market. In: ASIS&T, ASIS&T 75th Annual Meeting. Baltimore, MD, USA, 26-30.

Knight, G. (2008). Framework for the definition of significant properties. Retrieved from http://www.significantproperties.org.uk/documents/wp33- propertiesreport-v1.pdf

Knight, G. & Hedges, M. (2007). Modelling OAIS compliance for disaggregated preservation services. International Journal of Digital Curation, 1(2), 62-72.

Kuhn, T. S. (1996). *The structure of scientific revolutions* (3rd ed.). Chicago, IL: University of Chicago Press.

Lage, K., Losoff, B., & Maness, J. (2011). Receptivity to library involvement in scientific data curation: a case study at the University of Colorado Boulder. portal: Libraries and the Academy, 11(4), 915-937.

Latour, B. (1987). *Science in action: how to follow scientists and engineers through society.* Milton Keynes, UK; Open University Press.

Lavoie, B. F. (January/February 2000). Meeting the challenges of digital preservation: The OAIS reference model. OCLC Newsletter, No. 243, 26-30.

Lavoie, B. F. (2004) The Open Archival Information System Reference Model: Introductory Guide. Microform & Imaging Review, 33(2), 68-81.

Law, J. (1992). “Notes on the Theory of the Actor-Network: Ordering, Strategy, and Heterogeneity.” Systems Practice 5(4): 379-393.

Lee, C. (2005). “Defining Digital Preservation Work: A Case Study of the Development of the Reference Model for an Open Archival Information System.” Doctoral Dissertation. The School of Information, University of Michigan, Ann Arbor.

Lee, C. (2009). "Open Archival Information System (OAIS) Reference Model." In *Encyclopedia of Library and Information Sciences*, Third Edition, eds Bates, M. and Maack, M. Boca Raton, FL: CRC Press, 4020-4030.

Lindlar, M. (2013). Time to Change Effects and Implications of Digital Preservation in an Organizational Context. IASSIST Quarterly, Fall Winter 2013.

MacNeil, H. & Mak, B. (2007). "Constructions Of Authenticity." Library Trends 56(1), 26-52.

Madison, D. S. (2005). *Critical ethnography: method, ethics, and performance.* Thousand Oaks, CA: Sage Publications.

Marglin, S. (1978). “What do bosses do?” In *The Division of Labour*, ed. Gorz, A., Brighton, UK: Harvester Press.

Marino, M. “Reading the Transborder Immigrant Tool,” HASTAC. January 18, 2011. http://www.hastac.org/blogs/markcmarino/reading-transborder-immigrant-tool-mla-11-cross-post.

Matthews, B., McIlwrath, B., Giaretta, D., & Conway, E. (2008). The significant properties of software: A study. JISC report, March.

McCarthy, C. (1994). Multicultural Discourses and Curriculum Reform: A Critical Perspective. Educational Theory, 44(1), 81-98.

McDonough, J. (2008). "Structural Metadata and the Social Limitation of Interoperability: A Sociotechnical View of XML and Digital Library Standards Development." Balisage: The Markup Conference 2008. Montreal, Canada, Aug. 12-15, 2008. In Proceedings of Balisage: The Markup Conference 2008. Balisage Series on Markup Technologies, vol. 1.

McDonough, J. (2011). Packaging videogames for long-term preservation: Integrating FRBR and the OAIS reference model. Journal of The American Society For Information Science & Technology, 62(1), 171-184.

McDonough, J. (2011). 'Knee-deep in the data': Practical problems in applying the OAIS Reference Model to the preservation of computer games (Conference Paper) Proceedings 2012 45th Hawaii International Conference on System Sciences, 1625-1634.

McDonough, J. (2013). A Tangled Web: Metadata and Problems in Game Preservation In The Preservation of Complex Objects: Volume 3. Gaming Environments and Virtual Worlds. JISC.

McDonough, J., Kirschenbaum, M., Reside, D., Fraistat, N. & Jerz, D. (Fall 2010). "Twisty Little Passages Almost All Alike: Applying the FRBR Model to a Classic Computer Game." Digital Humanities Quarterly 4(2).

McDonough, J., Olendorf, R., Kirschenbaum, M., Kraus, K., Reside, D., Donahue, R., Phelps, A., Egert, C., Lowood, H., & Rojo, S. (2010). Preserving Virtual Worlds Final Report. Retrieved from IDEALS online 2011-11-30.

McLuhan, M. (1964). *Understanding Media: The Extensions of Man.* New York, NY: McGraw Hill.

McMeekin, S. (2011). With a Little Help from OAIS: Starting down the Digital Curation Path. Journal of the Society of Archivists, 32:2, 241-253.

Meghini, C. (August 2013) Data preservation. Data Science Journal Volume 12(10), GRDI51-GRDI57.

Miaskiewicz, T., Sumner, T., & Kozar, K. A. (April 2008). A latent semantic analysis methodology for the identification and creation of personas. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* 1501-1510.

Nakamura, L. (2002). *Cybertypes: Race, Ethnicity, and Identity on the Internet*. New York, NY: Routledge.

Nakamura, L. (2009). Don't Hate the Player, Hate the Game: The Racialization of Labor in World of Warcraft. Critical Studies In Media Communication, 26(2), 128-144.

National Science Foundation Cyberinfrastructure Council (2007). Cyberinfrastructure Vision for 21st Century Discovery. Available from http://www.nsf.gov/pubs/2007/nsf0728/nsf0728. Pdf

Nesmith, T. (Spring 1999). "Still Fuzzy But More Accurate: Some Thoughts on the 'Ghosts' of Archival Theory," Archivaria 47.

nestor: http://www.langzeitarchivierung.de/Subsites/nestor/DE/Home/home_node.html

Nicholson, D., & Dobreva, M. (2009). Beyond OAIS: Towards a reliable and consistent digital preservation implementation framework. DSP 2009: 16th International Conference on Digital Signal Processing, Proceedings, art. no. 5201126.

Noble, S. (2012). Searching for Black Girls: Old Traditions in New Media. Doctoral Thesis. Graduate School of Library and Information Science, University of Illinois, Urbana-Champaign.

OCLC/CLR. (February 2007). Trustworthy Repositories Audit & Certification: Criteria and Checklist, Version 1.0.

Oudshoorn, N., Rommes, E., Stienstra, M. (January 2004). Configuring the User as Everybody: Gender and Design Cultures in Information and Communication Technologies. Science Technology Human Values, 29(1), 30-63.

Owens, T. (2014). We're All Digital Archivists Now: An Interview with Sibyl Schaefer. The Signal Digital Preservation Blog. Library of Congress.

Owens, T. (2015). Digital Sources & Digital Archives: The Evidentiary Basis of Digital History (Draft). Retrieved from: http://www.trevorowens.org/2015/12/digital-sources-digital-archives-the-evidentiary-basis-of-digital-history-draft/

PACG (2011). Educating for the archival multiverse. American Archivist, 74, 69–101.

Poster, M. (2004). "The Mode of Information and Postmodernity." Chapter 27 in *The Information Society Reader*. London, UK: Routledge, 399-410.

Presner, T. (2014). "The Ethics of the Algorithm: Close and Distant Listening to the Shoah Foundation Visual History Archive," forthcoming in: History Unlimited: Probing the Ethics of Holocaust Culture (Cambridge: Harvard University Press, 2015), paired with Stephen Smith, "On the Ethics of Technology and Testimony: A response to Todd Presner."

Ranganathan, S. R. (1931). The five laws of library science.

Rawson, K.J. (2014). Transgender Worldmaking in Cyberspace: Historical Activism on the Internet. QED: A Journal in GLBTQ Worldmaking 1(2), 38-60.

Rose, G. (2011). *Visual Methodologies: An Introduction to the Interpretation of Visual Material.* London, UK: Sage Publications.

Rosenthal, D. (2015). The case for a revision of OAIS. Digital Preservation Coalition OAIS Wiki. Available at: http://wiki.dpconline.org/index.php?title=The_case_for_a_revision_of_OAIS

Rubin, N. (2010). Preserving Digital Public Television Final Report. Library of Congress.

Sacchi, S., Wickett, K., Renear, A., & Dubin, D. (2011). A Framework for Applying the Concept of Significant Properties to Datasets. In Proceedings of ASIS&T 2011.
Schwartz, J. & Cook, T. (2002). "Archives, Records, and Power: The Making of Modern Memory," *Archival Science*, 2:1/2, 1-19.

SCAPE- Scalable Preservation Environments: http://scape-project.eu/

Schiller, D. (2006). *How to Think About Information*. Urbana, IL: University of Illinois Press.

Schumann, N., & Recker, A. (2012). De-mystifying OAIS compliance: Benefits and challenges of mapping the OAIS reference model to the GESIS Data Archive. IASSIST Quarterly, 36(2), 6-11.

Sedgwick, E. K. (1997). Paranoid reading and reparative reading, or, You're so paranoid, you probably think this introduction is about you.

Sekula, A. (1986). 'Reading an archive: photography between labour and capital', In *Photography/Politics: 2,* eds. Holland, P., Spence, J., and Watney, S. London, UK: Comedia, 153-161.

Seles, A. (2016). The Transferability of Trusted Digital Repository Standards to an East African Context. Doctoral Thesis. Department of Information Studies, University College London. London, UK.

Sergeant, D. (2002). Interpretation of the OAIS Model. Retrieved from http://www.erpanet.org/events/2002/copenhagen/presentati ons/dmserpanet.ppt

Sherratt, T. (2015). "Seams and edges: Dreams of aggregation, access & discovery in a broken world" ALIA Online 2015: Sydney, Australia.

Sierman, B. (2012). "OAIS 2012 Update" in Digital Preservation Seeds Blog accessible at http://digitalpreservation.nl/seeds/oais-2012-update/

Sierman, B. (2014). The SCAPE Policy Framework, maturity levels and the need for realistic preservation policies. iPRES 2014: Proceedings of the 11th International Conference on Digital Preservation, 259-266.

Smith, A. (Summer 2003). Authenticity and Affect: When is a Watch Not a Watch? Library Trends 52(1), 172-82.

Smith, A. (2007). Valuing Preservation. Library Trends 56(1), 4-25.

Spence, J. (2006). Preserving the cultural heritage: An investigation into the feasibility of the OAIS model for application in small organisations. Aslib Proceedings: New Information Perspectives, 58(6), 513-524.

Star, S. & Griesemer, J. (1989). Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Professionals in Social Studies of Science, 19(0), 387-420.

Subotic, I., Schuldt, H., & Rosenthaler, L. (2011). The DISTARNET Approach to Reliable Autonomic Long-term Digital Preservation. In DASFAA 2011, Part II, LNCS 6588, eds. Yu, J.X., Kim, M.H. and Unland, R., 93–103.

Thaler, R. & Sunstein, C. (2009). *Nudge: improving decisions about health, wealth and happiness.* New York, NY: Penguin.

Tillett, B. (September/October 2003). What is FRBR: A conceptual model for the bibliographic universe. Technicalities, 25(5).

Trist, E. (June 1981). "The evolution of sociotechnical systems: a conceptual framework and an action research program." Occasional paper No. 2.

Tsing, A. (2005). *Friction: an ethnography of global connection.* Princeton, NJ: Princeton University Press.

Tuchman, G. (1978). ''Introduction: The Symbolic Annihilation of Women by the Mass Media,'' in Hearth and Home: Images of Women in the Mass Media, eds. Tuchman, G., Kaplan Daniels, A., and Benet, J. New York, NY: Oxford University Press.

Van House, N. & Churchill, E. (Sept. 2008). Technologies of Memory: Key Issues and Critical Perspectives. Memory Studies 1(3): 295-310.

Van Malssen, K. (2010) Planning Beyond Digitization: Digital Preservation of Audiovisual Collections. in Zorgen Voor Onzichtbare Assets: Over Het Behoud van Digitale AV-Collecties, ed. Annemieke de Jong. Beeld en Geluid and AVA_Net, Netherlands.

Vardigan, M. and Whiteman, C., (2007). ICPSR meets OAIS: applying the OAIS reference model to the social science archive context. Archival Science, 7(1), 73–87.

Walsham, G. (1997). Actor-Network Theory and IS Research: Current Status and Future Prospects. In Information Systems and Qualitative Research: Proceedings of the IFIP TC8 WG 8.2 International Conference on Information Systems and Qualitative Research, 31 st May-3rd June 1997, Philadelphia, Pennsylvania, 466-480.

Webb, C., Pearson, D., and Koerbin, P. (2013). 'Oh, you wanted us to preserve that?!' Statements of Preservation Intent for the National Library of Australia's Digital Collections. D-Lib Magazine 19:1/2.

Wilson, A. (2006). An Innovative Shared Preservation Service: The SHERPA-DP Project: Securing a Hybrid Environment for Research Preservation and Access - Digital Preservation. Archiving Conference, Archiving 2006 Final Program and Proceedings, 1-6.

Yeo, G. (2010). "'Nothing is the same as something else': significant properties and notions of identity and originality." Archival Science, 10(2), 85-116.

APPENDIX A: IRB INFORMATION

Consent to Participate

Encoding Power: Digital Spaces using the Open Archival Information System (OAIS) Reference Model

You are invited to participate in a research study on the use of the OAIS reference model in preserving cultural heritage and audio/visual content. This study is being conducted by doctoral student Rhiannon Bettivia in the Graduate School of Library & Information Science from the University of Illinois Urbana Champaign.

This study will take approximately one to two hours of your time. You will be asked to participate in an interview regarding your experience with the OAIS reference model and your judgments regarding the importance of various aspects of model in your daily work. This interview will be recorded and transcribed, but your name and identity will not be included in the transcription and any personal names occurring in the course of the interview will be replaced with pseudonyms in the transcription. The original audio recordings will be deleted once transcribed. Audio recordings are preferred for this study; if you do not wish to be recorded, you may decline.

Your decision to participate or decline participation in this study is completely voluntary and you have the right to terminate your participation at any time without penalty. You may skip any questions you do not wish to answer.

Your participation in this research will be completely confidential. I will not collect or store personal identifying information regarding you or include any personal identifying information in research results that we disseminate. Interview transcriptions will be maintained in secured storage in the researcher's office. Possible outlets of dissemination for our research results may include publication through journals such and presentations at conferences in areas of digital preservation. Although your participation in this research may not benefit you personally, it will help me understand which aspects of the OAIS reference model are significant to the preservation community and assist in the preservation of culturally significant materials.

There are no risks to individuals participating in these interviews beyond those that exist in daily life.

Your decision to participate, decline, or withdraw from participation will have no effect on your status at or future relations with the University of Illinois.

If you have questions about this project, you may contact

RHIANNON BETTIVIA
Doctoral Student
Graduate School of Library & Information Science
University of Illinois at Urbana-Champaign
501 E. Daniel Street, MC-493
Champaign, IL 61820-6211
(917) 334-3519
rbettivi@illinois.edu

JEROME P. MCDONOUGH
Associate Professor
Graduate School of Library & Information Science
University of Illinois at Urbana-Champaign
501 E. Daniel Street, MC-493
Champaign, IL 61820-6211
(217) 244-5916
jmcdonou@illinois.edu

If you have any questions about your rights as a research participant in the study, please contact the University of Illinois Institutional Review Board at 217-333-2670 (collect calls accepted if you identify yourself as a research participant) or via email at irb@illinois.edu.

You will be given a copy of this consent form for your records.

I have read and understand the above consent form, I certify that I am 18 years old or older and, by signing this form, I agree to participate in this study.

I consent to have my interview audio recorded (please check one) : yes ____ no ____

I consent to be contacted for brief follow up questions, not to take more than 15-20 minutes of time via email or phone (please check one): yes____ no ____

If yes, please note preferred method of contact for follow up:_____

Signature Date

UNIVERSITY OF ILLINOIS
AT URBANA-CHAMPAIGN

Office of Vice Chancellor for Research

Institutional Review Board
528 East Green Street
Suite 203
Champaign, IL 61820

August 13, 2014

Jerome McDonough
Library & Information Science
202 LIS
501 E Daniel St
M/C 493

RE: *Encoding Power: Traditional Archives in Digital Spaces using the Open Archival Information System (OAIS) Reference Model*
IRB Protocol Number: 15092

EXPIRATION DATE: 08/12/2017

Dear Dr. McDonough:

Thank you for submitting the completed IRB application form for your project entitled *Encoding Power: Traditional Archives in Digital Spaces using the Open Archival Information System (OAIS) Reference Model*. Your project was assigned Institutional Review Board (IRB) Protocol Number 15092 and reviewed. It has been determined that the research activities described in this application meet the criteria for exemption at 45CFR46.101(b)(2).

This determination of exemption only applies to the research study as submitted. Please note that additional modifications to your project need to be submitted to the IRB for review and exemption determination or approval before the modifications are initiated.

We appreciate your conscientious adherence to the requirements of human subjects research. If you have any questions about the IRB process, or if you need assistance at any time, please feel free to contact me or the IRB Office, or visit our website at http://www.irb.illinois.edu.

Sincerely,

Rebecca Van Tine

Rebecca Van Tine, MS
Assistant Human Subjects Research Specialist, Institutional Review Board

c: Rhiannon Bettivia

telephone (217) 333-2670 • fax (217) 333-0405 • email IRB@illinois.edu