

Mary Courtney Bailey. Bridging the Gap: Handling Born-Digital Records in Manuscript Repositories. A Master's Paper for the M.S. in L.S degree. April, 2013. 80 pages.  
Advisor: Jacqueline Dean

Due to the rapidly burgeoning volume of born-digital records, it behooves archives to determine how they can best bridge the gap between handling analog and handling born-digital records. In addition to analyzing existing case studies of repositories that already accession and process electronic records, the study presented in this paper used both a survey instrument and semi-structured interviews with archivists to investigate whether and how manuscript repositories are handling born-digital materials. The intent of this study has been to pinpoint some of the problems that plague manuscript repositories in particular and to identify some practical steps that have already been taken at similar repositories and should be replicated and suggest courses for further study and action. The problems of handling born-digital materials, especially when it comes to providing access, are not unique, and the profession would be well served by finding a space for collaboration to solve these thorny issues.

#### Headings:

Archival materials

Digital preservation

Electronic information resources -- Access control

Electronic records

Libraries -- Special collections -- Manuscripts

Manuscripts -- Collections

BRIDGING THE GAP:  
HANDLING BORN-DIGITAL RECORDS  
IN MANUSCRIPT REPOSITORIES

by  
Mary Courtney Bailey

A Master's paper submitted to the faculty  
of the School of Information and Library Science  
of the University of North Carolina at Chapel Hill  
in partial fulfillment of the requirements  
for the degree of Master of Science in  
Library Science.

Chapel Hill, North Carolina

April 2013

Approved by

---

Jacqueline Dean

## Table of Contents

Introduction. ....	2
Literature Review.....	10
Prepare.....	11
Identify.....	12
Select. ....	13
Get. ....	17
Store.....	19
Protect. ....	24
Manage.....	26
Provide.....	29
Case Studies.....	37
Research Design.....	41
Findings. ....	44
Survey of Repositories Not Currently Processing Born-Digital Materials. ....	44
Interviews with Repositories Currently Processing Born-Digital Materials. ....	46
Feedback from Jump In Initiative Participants. ....	51
Discussion. ....	54
Conclusion.....	64
Bibliography.....	69
Appendix A. ....	74
Appendix B. ....	75
Appendix C. ....	77
Appendix D. ....	78

## Introduction

---

*“If we do not solve the problems surrounding e-records, who will?”*<sup>1</sup>

---

In his famous 1995 article for *Scientific American*, Jeff Rothenberg warns, “digital information lasts forever – or five years, whichever comes first.”<sup>2</sup> Elizabeth Dow echoes this point, recognizing that while analog materials could be accessioned and then handled at whatever point in the future when it best fit into the curator’s schedule, electronic records have their own schedule that must be followed unless a repository is willing to risk losing these electronic records.<sup>3</sup> When Rothenberg expanded on this article four years later, he identifies four primary modes of loss of electronic records: “physical decay of media, loss of information about the format, encoding, or compression of files, obsolescence of hardware, and unavailability of software.”<sup>4</sup> While Rothenberg and others certainly called attention to the issue of electronic records in the 1990s, in fact, archivists have been dealing with machine-readable records at least since the 1970s. For example,

---

<sup>1</sup> Quote from a public university archivist, in response to an open-ended question on the 2007 survey by Susan Davis. Susan E. Davis, “Electronic Records Planning in ‘Collecting’ Repositories,” *American Archivist* 71 (Spring/Summer 2008): 183.

<sup>2</sup> Jeff Rothenberg, “Ensuring the Longevity of Digital Documents,” *Scientific American* 272 (January 1995): 42.

<sup>3</sup> Elizabeth H. Dow, *Electronic Records in the Manuscript Repository* (Lanham, MD: Scarecrow Press, 2009), xiv.

<sup>4</sup> Jeff Rothenberg, *Ensuring the Longevity of Digital Information* (Washington, DC: Council on Library and Information Resources, 1999), 3, <http://www.clir.org/pubs/archives/ensuring.pdf>.

the National Archives and Records Administration (NARA) received its first electronic records in 1969.<sup>5</sup> A few years before Rothenberg's piece, Margaret Hedstrom writes of a joint meeting between the Society of American Archivists' Committee on Automated Records and Techniques and the National Association of Government Archives and Records Administrators' Committee on Information Technology that determined that "archivists need to be open to 'radical thinking' about the role of archives because successfully dealing with electronic records may demand a transformation of the basic purpose of archives and the methods archivists use."<sup>6</sup> Hedstrom goes on to identify barriers to success, concluding that "fear of change, aversion to risk, a custodial mentality, and a failure to recognize electronic records as critical to the future success of archives" are all significant impediments that reinforce the notion that the purview of archivists is limited to the paper realm.<sup>7</sup>

Due to expanding varieties of digital objects and the increasing debate over whether electronic records fit into the traditional paradigm of archival practice, there was a second wave of literature about born-digital materials that was generated during the last decade of the twentieth century and the first decade of the twenty-first century. Charles Dollar has been very vocal in his call to archivists to look to the future rather than the past. In a 1992 address, he challenges, "We must get our archival heads out of the sands

---

<sup>5</sup> Kenneth Thibodeau, "Preserving Digital Memory at the National Archives and Records Administration of the U.S." (paper presented at the Workshop on Conservation of Digital Memories, Second National Conference on Archives, Bologna, November 20, 2009), 1, <http://www.mybestdocs.com/thibodeau-k-preser-dig-memNARA091120.pdf>.

<sup>6</sup> Margaret Hedstrom, "Electronic Records Program Strategies: An Assessment," in *Electronic Records Management Program Strategies*, ed. Margaret Hedstrom (Pittsburgh: Archives & Museum Informatics, 1993), 1.

<sup>7</sup> Ibid., 4.

of practices devised for medieval charters and papal decrees. We must realize that clinging to old practices in light of the volume of new records is *not* a noble defense of principle or archival tradition, but an act of willful neglect.”<sup>8</sup> Setting aside disagreements over archival principles, the problem with these writings from the perspective of manuscript repositories is that they focused on the evidential value of business and government records without much acknowledgment of the role that personal papers play in the documentation of a society. Writing in *Archives and Manuscripts* in 1994, Adrian Cunningham proclaims his article as “a first attempt at redressing this imbalance in the literature” regarding personal records versus government and organisational records.<sup>9</sup>

Due to the rapidly burgeoning volume of born-digital records, it behooves archives to determine how they can best bridge the gap between handling analog and handling born-digital records. In his foreword to the report the Council on Library and Information Resources published in December 2010 entitled *Digital Forensics and Born-Digital Content in Cultural Heritage Collections*, Charles Henry estimates that 90 percent of the records currently being created are born digital.<sup>10</sup> Given the huge number of born-digital materials currently being generated and the long history of their consideration in the archival literature, one has to wonder why manuscript repositories have yet to reach a consensus about the best methods for handling born-digital collections. In her book

---

<sup>8</sup> Philip C. Bantin, “Strategies for Managing Electronic Records: A New Archival Paradigm? An Affirmation of Our Archival Traditions?” *Archival Issues* 23, no. 1 (1998): 17.

<sup>9</sup> Adrian Cunningham, “The Archival Management of Personal Records in Electronic Form: Some Suggestions,” *Archives and Manuscripts* 22 (May 1994): 95.

<sup>10</sup> Matthew G. Kirschenbaum, Richard Ovenden, and Gabriela Redwine, *Digital Forensics and Born-Digital Content in Cultural Heritage Collections* (Washington, DC: Council on Library and Information Resources, 2010), vii, <http://clir.org/pubs/reports/pub149/pub149.pdf>.

*Electronic Records in the Manuscript Repository*, Elizabeth Dow offers one explanation. She suggests that electronic records “aren’t nearly as seductive as the old records” and refers to them as the “ugly babies of our professional future”<sup>11</sup> – not beautiful, but still needing just as much care and attention as the pretty babies, such as a letter by a famous author, written on monogrammed stationery. It appears that archivists may have fallen prey to some of the predilections that are documented in a survey of Canadian historians, which reports that original sources “engage the senses, not just the mind.”<sup>12</sup> Of course, born-digital materials will never fulfill this sensory longing.

Perhaps another of the stumbling blocks to having manuscript repositories embrace born-digital materials is semantic. Tom Hyry and Rachel Onuf point out that the words “personal papers” are both imprecise and “anachronistic. If they are going to continue to be used to identify the non-work-oriented materials generated by an individual, they will need to be explicitly redefined and expanded.”<sup>13</sup> For their part, the British Library calls personal digital objects that are the equivalent of personal papers eMANUSCRIPTS (eMSS).<sup>14</sup> In this paper, born-digital materials, electronic records, and digital objects will all be used to refer to objects that have been generated in a digital format.

---

<sup>11</sup> Dow, xii.

<sup>12</sup> Wendy Duff, Barbara Craig and Joan Cherry, “Historians’ Use of Archival Sources: Promises and Pitfalls of the Digital Age,” *Public Historian* 26, no. 2 (Spring 2004): 19.

<sup>13</sup> Tom Hyry and Rachel Onuf, “The Personality of Electronic Records: The Impact of New Information Technology on Personal Papers,” *Archival Issues* 22, no. 1 (1997): 41.

<sup>14</sup> Jeremy Leighton John, “Adapting Existing Technologies for Digitally Archiving Personal Lives: Digital Forensics, Ancestral Computing, and Evolutionary Perspectives and Tools,” iPRES 2008: The Fifth International Conference on Preservation of Digital Objects (29-30 September 2008), [http://www.bl.uk/ipres2008/presentations\\_day1/09\\_John.pdf](http://www.bl.uk/ipres2008/presentations_day1/09_John.pdf).

Laura Millar helps to identify five key issues regarding electronic records in her 2010 book on archival principles and practices. She lists: “technological dependence and obsolescence; mutability; the potential loss of context; the effects of decentralized information management, security and privacy; and cost.”<sup>15</sup> Perhaps the most challenging aspect of this list from the perspective of manuscript repositories is that these factors reside outside of their control. Software and hardware dependencies and/or obsolescence are set in motion by the record creators, therefore, many repositories do not have involvement with or control over decisions such as whether to use open-source or interoperable software. The issues of context and decentralized control also revolve around the record creators, with the repository being resigned to deal with whatever amount of metadata is provided by the donor and with whatever complications or losses of data occurred before materials crossed the archival threshold. While the repository is directly responsible for providing security that ensures records cannot be damaged or changed and that personally identifiable information is not distributed to inappropriate parties – just as archives have done for centuries with analog records – the mechanisms for causing these problems are very different for electronic records, such as viruses or bit rot that can affect the integrity of computer files. Finally, cost also remains a factor outside the control of repositories, with budgets often dependent on institutional funding or soft money from grants, the prices for technological components set by market forces, and the staffing requirements very dependent on individual abilities and interests. All of these uncertainties help to explain why manuscript repositories have not been quick to

---

<sup>15</sup> Laura A. Millar, *Archives: Principles and Practices* (New York: Neal-Schuman Publishers, 2010), 207.



take up the mantle of preserving born-digital materials.

OCLC Research has made several contributions in recent years to the field of handling born-digital materials. In 2010, they surveyed the Association of Research Libraries, the Canadian Academic and Research Libraries, the Independent Research Libraries Association, the Oberlin Group, and the U.S. and Canadian members of the RLG Partnership. Jackie Dooley and Katherine Luce identify three actions that need to be taken by the special collections community to address born-digital archival materials:

- (1) “Define the characteristics of born-digital materials that warrant their management as *‘special collections.’*”
- (2) “Define a reasonable set of *basic steps* for initiating an institutional program for responsibly managing born-digital archival materials.”
- (3) “Develop *use cases and cost models* for selection, management, and preservation of born-digital archival materials.”<sup>16</sup>

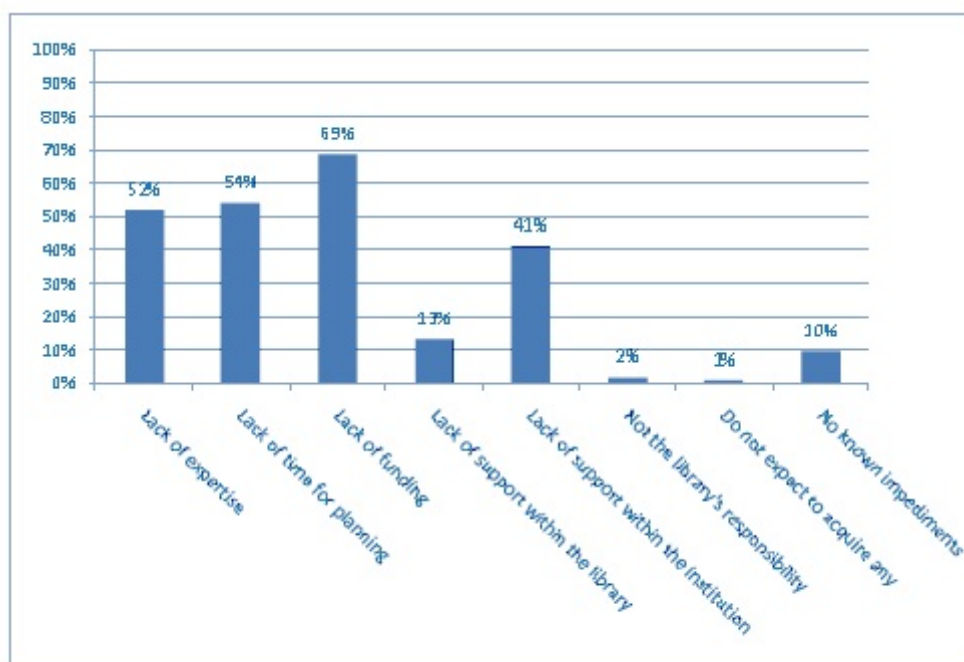


Figure 1.25. Impediments to born-digital management (Q. 63, n = 157)

<sup>16</sup> Jackie M. Dooley and Katherine Luce, “Taking Our Pulse: The OCLC Research Survey of Special Collections and Archives,” *OCLC Research*, October 2010, 13, <http://www.oclc.org/resources/research/publications/library/2010/2010-11.pdf>.

One of their survey questions revealed information about the impediments to the management of born-digital materials, with the results showing that lack of funding, lack of time for planning, lack of expertise, and lack of support within the institution pose common roadblocks (see Figure 1.25).<sup>17</sup> As already noted, lack of funding is often beyond the reach of a repository to solve, and lack of institutional support will require some intentional advocacy by archivists to reverse, but lack of time for planning and lack of expertise can more immediately be addressed on the repository level.

In 2012, Ricky Erway produced two reports for OCLC Research that contribute to the conversation about born-digital materials. In *You've Got to Walk Before You Can Run*, she suggests that a simple three-step survey can begin to address electronic records that are already in the collections of an archive: (1) find the physical media already in the repository, (2) count and describe these media, and (3) prioritize the further treatment of collections.<sup>18</sup> Simplifying the situation and identifying a place where repositories can begin attacking the problem of born-digital materials should enable more institutions to join this field. The second in this series of Demystifying Born Digital reports suggests the creation of SWAT sites – “software and workstations for antiquated technology” sites that have the expertise to share regarding the handling of digital media.<sup>19</sup> The benefit of

---

<sup>17</sup> Ibid., 60.

<sup>18</sup> Ricky Erway, *You've Got to Walk Before You Can Run: First Steps for Managing Born-Digital Content Received on Physical Media* (Dublin, OH: OCLC Research, 2012), 3-4, <http://www.oclc.org/content/dam/research/publications/library/2012/2012-06.pdf>. Erway elaborates on each of these steps, and step two includes processes such as identifying file formats and calculating the overall size of the digital collection. She also includes on page 5 eleven technical steps to follow with physical media.

<sup>19</sup> Ricky Erway, *Swatting the Long Tail of Digital Media: A Call for Collaboration* (Dublin, OH: OCLC Research, 2012), 3, <http://www.oclc.org/content/dam/research/publications/library/2012/2012-08.pdf>.

this plan is that it does not necessitate all repositories becoming technical experts on all varieties of born-digital materials from all eras. If this sort of collaboration can be engendered, this holds great possibility.

In addition to analyzing existing case studies of repositories that already accession and process electronic records, the study presented in this paper used both a survey instrument and semi-structured interviews with archivists to investigate whether and how manuscript repositories are handling born-digital materials. While not focusing too closely on the technical issues of born-digital records and not summarizing all of the debate in the archival literature about issues that relate to the processing of born-digital materials, the intent of this study has been to pinpoint some of the problems that plague manuscript repositories in particular and to identify some practical steps that have already been taken at similar repositories and should be replicated and to suggest courses for further study and action.

### Literature Review

---

*“This is a huge issue, especially for small institutions.  
I see this issue as a black hole in the fabric of history.”<sup>20</sup>*

---

Given the fact that electronic records have consumed the attention of a good portion of the professional archival literature across many continents for over three decades, this literature review is by no means exhaustive. It would be beyond the scope of this paper to include a comprehensive review. Instead, it provides a representative sampling of the discourse, focusing especially on issues that are relevant to the handling of born-digital materials by manuscript repositories.

For its Digital Preservation Outreach and Education program, the Library of Congress defines six core principles for its curriculum.<sup>21</sup> These topics were adapted and expanded by Cal Lee for the Closing the Digital Curation Gap project, a partnership between the University of North Carolina at Chapel Hill and Jisc that was funded by the Institute of Museum and Library Services.<sup>22</sup> These topics serve as the primary framework for the review of scholarship on born-digital materials. The final section of this literature review is an overview of various case studies that have been conducted about handling

---

<sup>20</sup> Quote from a private college archivist, in response to an open-ended question on the 2007 survey by Susan Davis. Davis, 183.

<sup>21</sup> “Digital Preservation & Outreach Curriculum,” Library of Congress, accessed 12 March 2013, <http://www.digitalpreservation.gov/education/curriculum.html>.

<sup>22</sup> “Getting Started Guides,” Digital Curation Exchange, last updated 8 March 2013, <http://digitalcurationexchange.org/gp>.

born-digital materials, along with a summation of the lessons learned from these investigations.

### ***1. Prepare***

In her list of steps and strategies to begin addressing the problem of electronic records in manuscript repositories, Elizabeth Dow counsels that a repository should work to develop or amend policies before committing to the long-term preservation of born-digital records.<sup>23</sup> In her 2006 article about the acquisition of the Michael Joyce papers at the Harry Ransom Center, Catherine Stollar Peters echoes this opinion, seeing policies as evidence of an institutional commitment to the project.<sup>24</sup>

While the theory of developing policies before the acquisition and processing of born-digital materials holds merit, in practice, this does not usually seem to be the case. In her 2007 survey of “collecting” repositories, including both public and private academic institutions and historical societies, Susan Davis found that only 24 percent of the institutions had a policy in place regarding the acquisition of digital records. Of that subset, 57 percent of those policies mirrored the policies for traditional collections.<sup>25</sup> She quotes a respondent from a public university who summarized the situation: ““We are passively accepting born-digital materials. We don’t even have a plan for preservation of the digital surrogates we are creating. We barely have enough staff to cover reference and manage limited processing. All planning, policy, etc. take a back seat to day-to-day

---

<sup>23</sup> Dow, 15.

<sup>24</sup> Catherine Stollar Peters, “When Not All Papers Are Paper: A Case Study in Digital Archivy,” *Provenance* 24 (2006): 34.

<sup>25</sup> Davis, 178.

efforts to keep up with basic activities.”<sup>26</sup> In a similar vein, Ben Goldman, writing about his work at the American Heritage Center, contends that beginning the work first can be a very valuable means of shaping the necessary policies and procedures for a repository.<sup>27</sup> Some questions cannot be answered (or even anticipated) unless the repository is already doing work with born-digital materials.

## **2. Identify**

The step of identifying includes both determining what born-digital materials might already be within collections and deciding what born-digital materials should be accessioned. In many cases, electronic media have been accessioned in hybrid collections without proper documentation having been generated about their existence. In his case study on the Beinecke Rare Book and Manuscript Library, Michael Forstrom defines such “fugitive media” in this way: “there has been no significant precustodial intervention, the digital content has not been appraised prior to acquisition, and the media is part of a collection consisting chiefly of paper-based materials.”<sup>28</sup> In defining the process of surveying digital materials, Elizabeth Dow acknowledges that a physical survey of an accumulated collection is prohibitive; instead, “surveying digital materials depends on determining the context of the materials’ creation and use.”<sup>29</sup>

---

<sup>26</sup> Ibid., 180.

<sup>27</sup> Ben Goldman, “Bridging the Gap: Taking Practical Steps Toward Managing Born-Digital Collections in Manuscript Repositories,” *RBM: Journal of Rare Books, Manuscripts, and Cultural Heritage* 12, no. 1 (2011): 21.

<sup>28</sup> Michael Forstrom, “Managing Electronic Records in Manuscript Collections: A Case Study from the Beinecke Rare Book and Manuscript Library,” *American Archivist* 72 (Fall/Winter 2009): 461.

<sup>29</sup> Dow, 2.

But the broader issue here boils down to one of determining when born-digital materials should be identified for long-term preservation in a manuscript repository. Elizabeth Dow points out that such archives have typically embraced the life cycle model when dealing with paper records, usually acquiring materials only when they have become inactive records. However, due to the “fragility or impermanence of digital documents,” Dow argues the continuum model provides a much more viable representation, which also “implies that archivists should identify digital materials of archival value and assert some authority over them at creation, or before.”<sup>30</sup> Adrian Cunningham concurs and adds that archivists frequently interact with highly sought after donors before any agreements are signed,<sup>31</sup> so conversing about software platforms and file naming conventions should not be seen as an extraordinary measure. He also argues that “much of the impetus for continuum thinking has come from the emergence of electronic records.”<sup>32</sup>

### 3. *Select*

Given the complexity of the long-term preservation of electronic records, there is an ongoing debate within the archival community about which records need to be maintained in a digital format. In his 2007 book *Records Management*, David Stephens

---

<sup>30</sup> Dow, 8.

<sup>31</sup> Cunningham, “Archival Management of Personal Records,” 101. For reference, Cunningham also provides a very succinct summary of the continuum model in a footnote in his chapter “Ghosts in the Machine: Towards a Principles-Based Approach to Making and Keeping Digital Personal Records,” in *I, Digital: Personal Collections in the Digital Era*, ed. Christopher A. Lee (Chicago: Society of American Archivists, 2011), 89.

<sup>32</sup> Adrian Cunningham, “Waiting for the Ghost Train: Strategies for Managing Electronic Personal Records Before It Is Too Late,” *Archival Issues* 24, no. 1 (1999): 58.

argues that in a business environment, data should only be preserved in a digital format if conversion to an analog format “would severely diminish its value or render it unusable in order to satisfy required (rather than ‘nice to have’) business requirements.”<sup>33</sup>

Elizabeth Dow uses colorful imagery to describe the appraisal of electronic records, referring to “the specter of the certain death of digital documents” and suggesting that this threat of the imminent demise of electronic records actually gives a curator more latitude to reject donations of questionable materials due to the effort required to maintain them – though she also cautions that, unlike paper records, collections that are rejected today are not likely to be available for reconsideration in the future.<sup>34</sup>

This leads into another debate raging in the archival community – whether archivists should have precustodial interventions with donors in order to identify records that should be kept and to attempt to ensure that these born-digital materials will persist until the time when they can be deposited in a manuscript repository. Elizabeth Dow acknowledges the side of the debate that worries such interventions might preclude the otherwise unselfconscious documentation by a donor, but she concludes that this risk is better than the alternative of having no viable records to ingest at the end of a person’s career.<sup>35</sup> Adrian Cunningham argues that “self-conscious record keeping” already exists outside of the realm of electronic records, so this should not be a reason to avoid the precustodial interventions that might ensure the continued viability of this evidence. He

---

<sup>33</sup> David O. Stephens, *Records Management: Making the Transition from Paper to Electronic* (Overland Park, KS: ARMA International, 2007), 237.

<sup>34</sup> Dow, 3.

<sup>35</sup> Ibid., 99. She also later admits that donors already shape donations by what they include and what they exclude from collections in any format. Ibid., 117.



also suggests that functional appraisal is the answer to discerning at the stage of creation which digital objects will hold historical significance. And where some discount the viability of precustodial intervention by suggesting it is too time consuming, Cunningham argues that it is merely a reallocation of time that otherwise would have been spent later in the process.<sup>36</sup> Tom Hyry and Rachel Onuf directly counter Cunningham's arguments, suggesting that precustodial interventions would skew appraisal decisions toward individuals who gain fame early in their lives and would force these decisions to occur without the perspective that comes with the passage of time.<sup>37</sup> But a manual recently published by the Society of American Archivists supports proactive involvement with donors, arguing that "much of the metadata used by archivists to add value to the digital records and manuscripts is best captured before it comes to the archives."<sup>38</sup>

The digital housekeeping practices of the creators of born-digital materials strongly influence the ability of a repository to determine which digital objects warrant preservation. As a result, repositories such as the Beinecke Rare Book & Manuscript Library have developed suggested guidelines for authors who intend to deposit their work. The minimum steps identified are: (1) save old physical media that contains unique files, (2) back up files, (3) use consistent file naming conventions, and (4) organize files logically. They go on to suggest guarding against obsolescence, insuring interoperability, adopting standards put forth by national and international organizations,

---

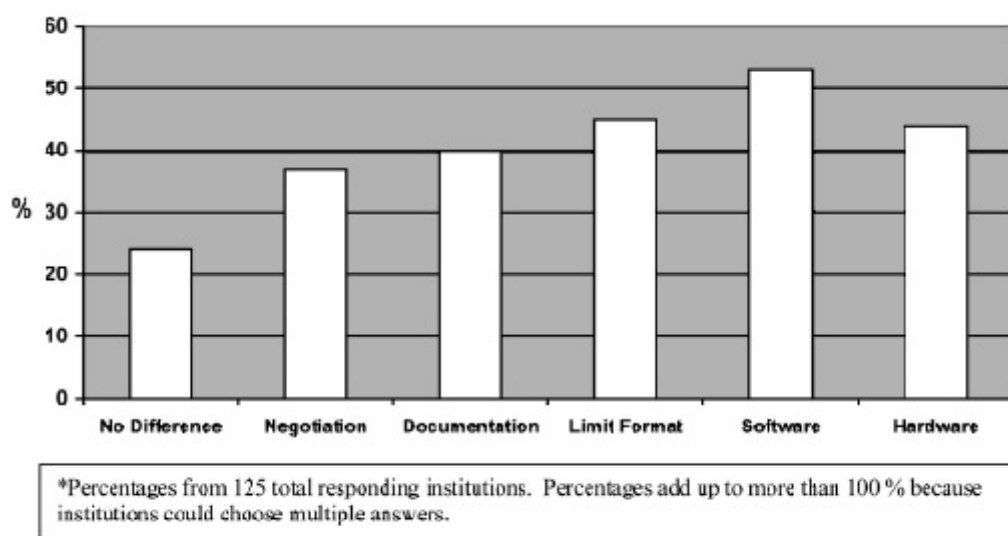
<sup>36</sup> Cunningham, "Waiting for the Ghost Train," 60-62.

<sup>37</sup> Hyry and Onuf, 43.

<sup>38</sup> J. Gordon Daines, "Processing Digital Records and Manuscripts," in *Archival Arrangement and Description*, ed. Christopher J. Prom and Thomas J. Frusciano (Chicago: Society of American Archivists, 2013), 95.

and ensuring backwards compatibility.<sup>39</sup>

However, manuscript repositories most often receive collections from donors who have not had extensive collaboration with the repository during the creation of the records. Susan Davis points to the complications that arise from receiving electronic records in a multitude of formats and without adequate accompanying metadata.<sup>40</sup> The results of her 2007 survey (as depicted in Figure 7 below) indicate that repositories may alter their acquisition procedures to reflect the particular concerns raised by born-digital materials by conducting more extensive negotiations with the donors of digital objects, asking for additional documentation, limiting the acceptable formats of electronic records, and specifying software and/or hardware requirements.<sup>41</sup>



**FIGURE 7. Different procedures for acquisition (Survey Question 9).**

<sup>39</sup> Beinecke Rare Book & Manuscript Library, “Authors’ Guidelines for Digital Preservation,” Yale University, <http://www.library.yale.edu/~nkuhl/AuthorsGuidelines.pdf>. For an draft version of the Born Digital Archival Acquisition Collection & Accession Guidelines from the Beinecke, see Naomi L. Nelson et al., *Managing Born-Digital Special Collections and Archival Materials*. SPEC Kit #329 (Washington, DC: Association of Research Libraries, 2012), 125.

<sup>40</sup> Davis, 169.

<sup>41</sup> Ibid., 182. For a good example of a donor agreement addendum for electronic records from the David M. Rubenstein Rare Book and Manuscript Library at Duke University, see Nelson et al., 122-23.

Some of the loudest voices arguing for earlier interventions by archivists come from the “new paradigm” theorists, who also tend to favor a documentation strategy for appraisal and the continuum rather than the life cycle model of records. At the same time, they support a system wherein the records creators remain the custodians of the digital objects and archivists serve as consultants.

#### ***4. Get***

During the 1990s, two schools of thought emerged about the appropriate locus of custody of electronic records. Luciana Duranti and Terry Eastwood from the University of British Columbia sought to apply traditional archival and diplomatics theory to electronic records and argued that the archive should serve as the holding place for electronic records, just as it has for centuries for paper records.<sup>42</sup> The Pittsburgh Project headed by David Bearman and Richard Cox argued that there needed to be a “new paradigm” in archival thinking to handle electronic records, embracing the continuum model of records and asserting that a noncustodial role was the only realistic one for repositories.<sup>43</sup> As Adrian Cunningham points out, the problem with the noncustodial position is that it does not encompass the need for archives for personal papers: “Governments and organisations may exist for indefinite periods of time or have cooperative successor organisations. Private individuals have an unfortunate habit of dying and leaving relatives who refuse to have any truck with the ongoing custody of the

---

<sup>42</sup> See, for example, Luciana Duranti, “Concepts and Principles for the Management of Electronic Records, or Records Management Theory is Archival Diplomatics,” *Records Management Journal* 20, no. 1 (2010): 78-95. This article is a re-publication of a 1999 article.

<sup>43</sup> See, for example, David Bearman, “An Indefensible Bastion: Archives as a Repository in the Electronic Age,” *Archives and Museum Informatics Technical Report* 13 (1991): 14-24.

deceased's records and who, in any case, probably could not be entrusted with the responsibility."<sup>44</sup>

The 2012 survey by the Association of Research Libraries (ARL) posed the question, "Which of the following strategies does your library employ when ingesting born-digital records stored on legacy media?"<sup>45</sup>

	Current Strategy	Planned Strategy	N
Storing legacy media as is (without transfer to new media or server storage and/or keeping it with analog collection)	47	1	48
Developing a collection of legacy hardware that can be used to retrieve data from legacy media (e.g., 5.25" floppy drives, zip drives, etc.)	27	10	37
Outsourcing the process of retrieving the data from legacy media	15	20	35
Building new systems that replicate the function of the legacy systems (e.g., emulation, virtual systems)	8	9	17
Participating in a collaborative that is developing a collection of legacy hardware	3	8	11
Other strategy	13	9	22
Number of Responses	57	32	60

While the respondents to this survey were limited to the members of the ARL, the results of this question certainly indicate that electronic records already are being accessioned by repositories and that there are some mechanisms in place to extract digital records from legacy media. For one example, a draft version of a digital processing manual for the

---

<sup>44</sup> Cunningham, "Archival Management of Personal Records," 99. Although Cunningham does an about-face several years later to embrace the distributed custody model for government and business records, he still maintains personal records will most often need to be taken into archival custody. See Adrian Cunningham, "Journey to the End of the Night: Custody and the Dawning of a New Era on the Archival Threshold," *Archives and Manuscripts* 24 (November 1996): 312-21.

<sup>45</sup> Nelson et al., 35.

Bentley Historical Library can be viewed online.<sup>46</sup>

One of the leading proponents for the use of digital forensics to acquire born-digital content is Jeremy Leighton John of the British Library. In speaking of personal papers, he identifies three requirements, each of which can be satisfied by using digital forensics tools:

(i) to capture as far as possible the whole contextual space of the personal computer (the entire hard drive or set of hard drives for example) and not just independent individual files, thereby strengthening authentication; (ii) to replicate and retain exact copies of the original files, recognising their historical and informational value (and not just rely on digital facsimiles, even if these match modern standards for interoperability); and (iii) to meet the special requirements for a confidentiality that is sensitive and reassuring to potential depositors as well as being technically convincing.<sup>47</sup>

Although there is not at this point one tool that can satisfy all of the needs of a manuscript repository, John points out that there are products that can handle everything from disk imaging to file browsing.

## **5. Store**

With all of the attendant complications of preserving and providing access to electronic records, some repositories are choosing to create hard copy records and store those as the record copies. Adrian Cunningham contends that this is an acceptable solution for some institutions, provided that there is no loss of necessary functionality

---

<sup>46</sup> Michael Shallcross, "Bentley Historical Library: Guidelines for the Manual Processing of Born-Digital Materials," University of Michigan, November 16, 2011, [http://deepblue.lib.umich.edu/bitstream/handle/2027.42/96439/BHL\\_DigitalProcessingGuidelines\\_20111116-DRAFT.pdf?sequence=1](http://deepblue.lib.umich.edu/bitstream/handle/2027.42/96439/BHL_DigitalProcessingGuidelines_20111116-DRAFT.pdf?sequence=1). The fact that this manual is less than two years old and is no longer being used by the Bentley Historical Library demonstrates something about the rapidity with which procedures regarding born-digital materials change.

<sup>47</sup> John, 1.

through the conversion to analog and that the process of creating and maintaining the paper records is cheaper than the alternative digital choice.<sup>48</sup> Elizabeth Dow also suggests that converting born-digital materials to an analog format is a reasonable preservation solution for small repositories, so long as the “essential contextual information” is preserved in this printout. However, she also acknowledges that this conversion carries an opportunity cost, for the analog surrogates share none of the functionality of their digital counterparts (e.g., being able to manipulate data in a database or follow a hyperlink in a document).<sup>49</sup> Ben Goldman, however, contends this is not a scalable procedure, for “there is not enough paper in the world to print, en masse, all the electronic records we have acquired (and will likely acquire in the future), nor would the solution even be appropriate for more complex types of digital files, such as databases, Web sites, or multimedia.”<sup>50</sup>

Having caused a stir in 2005 with the “More Product, Less Process” (MPLP) approach to archival processing that he and Dennis Meissner defined, Mark Greene followed in 2010 with an article that applies the MPLP principles to more aspects of archival administration. Greene provides a prominent voice in the archival community arguing that electronic records can be handled in much the same way that paper records

---

<sup>48</sup> Cunningham, “Archival Management of Personal Records,” 103.

<sup>49</sup> Dow, 62-63.

<sup>50</sup> Goldman, 14. Given the 1993 decision by the D.C. Circuit Court of Appeals, where the court dismissed the argument by the Reagan and Bush administrations that preserving hard copies of emails satisfactorily maintained the record, it is interesting that this line of thinking has persisted. The opinion states, “if only the hard copy is preserved in such situations, essential transmittal information relevant to a fuller understanding of the context and import of an electronic communication will simply vanish.” United States Court of Appeals for the District of Columbia Court, Case No. 93-5002. Scott Armstrong et al v. Executive Office of the President. Decided 13 August 1993. Accessed 16 March 2013 at [http://www.citizen.org/litigation/article\\_redirect.cfm?ID=620](http://www.citizen.org/litigation/article_redirect.cfm?ID=620).

are arranged and described. While suggesting that MPLP should also apply to description, Greene also embraces crowdsourcing as a mechanism for supplementing the less verbose descriptions that would be created through this process.<sup>51</sup>

Charles Dollar was one of the early voices emphasizing the importance of maintaining provenance and original order for electronic records. In order to accomplish this, he suggests that archivists need to “participate in the design of information resource directories or metadata systems and ensure that they in fact contain all of the contextual information essential to a full understanding of the records in question.”<sup>52</sup> He goes on to suggest that these systems should be responsible for generating description, rather than applying the traditional description process to electronic records.<sup>53</sup>

Although arrangement and description is ordinarily considered an access issue, given the scope and application of the literature to this point, this topic seems to fit more appropriately within the Store section. So an outgrowth of the debate over the life cycle versus the continuum models of records has been concern over whether traditional methods of archival description can adequately describe electronic records. Kathleen Roe is one who argues that there needs to be adaptation because the model of the physical arrangement of paper records does not translate directly to born-digital materials. She suggests that the records creator needs “to identify the records systems within which electronic records function. This focuses attention on how the intellectual relationships

---

<sup>51</sup> Mark. A. Greene, “MPLP: It’s Not Just for Processing Anymore,” *American Archivist* 73 (Spring/Summer 2010): 191-92.

<sup>52</sup> Charles M. Dollar, *Archival Theory and Information Technologies: The Impact of Information Technologies on Archival Principles and Methods* (Macerata, Italy: University of Macerata, 1992), 51.

<sup>53</sup> *Ibid.*, 62.

among databases or electronic files supported an organization's functions and activities."<sup>54</sup>

Philip Bantin provides a useful comparison of the description of electronic records within the continuum model and within the life cycle model. Where proponents of the life cycle model argue that traditional archival description provides the best means of protecting the authenticity of records, advocates of the continuum model suggest four reasons that alternative description methods are warranted. For one, they point out that effective description should take place during the life of the record, not when it becomes inactive. They also point out that prose descriptions do not easily reflect the complex relationships of digital objects. They acknowledge that the physical review of files to determine content and context is not viable for handling the scale of records produced in electronic environments. Lastly, they suggest that record system metadata is an existing alternative for description.<sup>55</sup>

This question of whether metadata provides adequate description has received quite a bit of attention in the archival literature.<sup>56</sup> Writing in *Archivaria* in 1993, Margaret Hedstrom explains that description for all types of records should allow users to identify and locate records, understand the record and interpret its content, and establish the authenticity of the record; apart from interaction with users, the description should

---

<sup>54</sup> Kathleen D. Roe, *Arranging & Describing Archives & Manuscripts* (Chicago: Society of American Archivists, 2005), 64.

<sup>55</sup> Bantin, 27-29.

<sup>56</sup> Elizabeth Dow identifies the three critical features of metadata – it describes content, context, and structure. Dow, 33.



also help manage the record.<sup>57</sup> She contends that archivists would be better served by capturing metadata generated in the records systems rather than generating it themselves.<sup>58</sup> David Wallace contends that in order for this to happen, archivists need to be involved in the creation of “electronic record-keeping systems,” which he differentiates from “data management,” which prioritizes timeliness and reusability of data rather than documenting transactional evidence.<sup>59</sup> But where Wallace sees metadata capable of doubling as archival description, Heather MacNeil disagrees. She compares metadata to a diary and description to “a biography, that, in narrational style, examines a life already lived, from a perspective broader than that in which it was lived.” She goes on to suggest that the volume of data generated by metadata systems is so vast that it “may in fact obscure, rather than illuminate, the broader administrative context and thereby bias the users’ understanding of the records’ meaning.”<sup>60</sup> MacNeil directly challenges Hedstrom’s argument, contending that using metadata as archival description actually perverts the primary purpose of metadata and thereby “contravenes the archivist’s primary duty to protect and preserve the inherent characteristics of archives – their impartiality, authenticity, and interrelatedness – which derive from the circumstances of their creation.”<sup>61</sup> In analyzing this debate that took place on the pages of

---

<sup>57</sup> Margaret Hedstrom, “Descriptive Practices for Electronic Records: Deciding What is Essential and Imagining What is Possible,” *Archivaria* 36 (Fall 1993): 55.

<sup>58</sup> Hedstrom, “Descriptive Practices,” 58.

<sup>59</sup> David A. Wallace, “Managing the Present: Metadata as Archival Description,” *Archivaria* 39 (Spring 1995): 18.

<sup>60</sup> Heather MacNeil, “Metadata Strategies and Archival Description: Comparing Apples to Oranges,” *Archivaria* 39 (Spring 1995): 25.

<sup>61</sup> *Ibid.*, 27.

*Archivaria* in 1995, Wendy Duff cautions, “before archivists abandon archival description, they require research that compares the retrieval performance of the two types of systems: one containing descriptions consisting of metadata and the other with descriptions supplied by archivists.”<sup>62</sup> Unfortunately, her focus on users has not been matched.

## **6. Protect**

Charles Dollar offers a simple definition of the preservation of electronic records: “ensuring their readability and intelligibility in order to facilitate data exchange over time.”<sup>63</sup> Many of the recommendations he made in 1992 for dealing with technology obsolescence are still embraced today, such as advocating for open systems standards and identifying migration paths.<sup>64</sup> In her recent book, Elizabeth Dow identifies five issues that have to be addressed in order to preserve electronic records for the long-term: preserving the hardware, the software, the storage medium, the skills (i.e., being able to use older programs and make sense of stored data), and the information.<sup>65</sup> She goes on to suggest that given the fact that it is impossible to anticipate future uses of electronic records by researchers, the professional goal of archivists should be one of guaranteeing the reusability of these electronic records. Accomplishing this goal will necessitate protecting the records from change, ensuring that migrations render documents that are

---

<sup>62</sup> Wendy Duff, “Will Metadata Replace Archival Description: A Commentary,” *Archivaria* 39 (Spring 1995): 37.

<sup>63</sup> Dollar, 68.

<sup>64</sup> *Ibid.*, 72.

<sup>65</sup> Dow, 22.

“coherent, reconstructible, and functional,” and documenting actions taken both while records reside in a repository as well as at the time of disposition.<sup>66</sup> On a more technical level, David Stephens identifies five types of data preservation practices. They are (1) updating the media on which electronic records are stored, (2) migrating data to new formats, (3) standardizing file formats, (4) recopying media at specified intervals, and (5) emulating the environment in which the digital object was created.<sup>67</sup>

Some of the very characteristics of electronic records make them more difficult to protect. Jacques Grimard describes them as fluid, malleable, and dynamic and argues that they escape fixity.<sup>68</sup> David Wallace argues that “the central change wrought by the computer is the looming mutability of the record itself.”<sup>69</sup> To address these concerns, Ben Goldman provides some very practical suggestions for maintaining the authenticity of digital records, such as running checksums on files upon ingest and periodically after that time and also documenting any actions taken on the digital objects, such as migration.<sup>70</sup>

The Digital Lives research project was led by the British Library, along with University College London and the University of Bristol. In their 2010 report, they make

---

<sup>66</sup> Ibid., 31-32.

<sup>67</sup> Stephens, 238-47. Grimard argues that migration should occur at least at ten year intervals. Jacques Grimard, “Managing the Long-term Preservation of Electronic Archives or Preserving the Medium and the Message,” *Archivaria* 59 (Spring 2005): 166. Standardizing file formats is also commonly referred to as normalization, although Laura Millar refers to it as “migration on ingest.” Millar, 217. Stephens’ use of media recopying as a term is less common than the term refreshing. He mentions emulation as a future preservation solution, one which he describes as a “Digital Rosetta Stone.”

<sup>68</sup> Grimard, 158-59.

<sup>69</sup> Wallace, 14. There is an extensive literature by postcustodial thinkers about the archival paradigm shift caused by electronic records. In addition to writers such as David Bearman, Margaret Hedstrom, and Charles Dollar cited elsewhere in this paper, see also, for example, Sue McKemmish, “Placing Records Continuum Theory and Practice,” *Archival Science* 1, no. 4 (2001): 333-59.

<sup>70</sup> Goldman, 20-21. The American Heritage Center uses the Duke Data Accessioner to accomplish these tasks. It can be downloaded at <http://library.duke.edu/uarchives/about/tools/data-accessioner.html>.

several suggestions about future roles for manuscript repositories, one of which pertains to the issue of authenticity. They suggest that repositories may not always have a custodial relationship to digital objects created by individuals and should look to become “guardians of the authenticity of the originals including digital objects in the wild.”<sup>71</sup> This parallels the noncustodial model of recordkeeping advocated much earlier by the Pittsburgh Project.

## **7. Manage**

While in the beginning many born-digital collections were treated as special projects, that is an increasingly ineffective strategy. As indicated by the 2012 survey by the Association of Research Libraries (ARL), the “trickle” of electronic records has become a “flood,” so archivists “must develop policies and procedures to operationalize the management of born-digital materials, or we risk losing the record of the recent past.”<sup>72</sup> The respondents to this survey indicated four critical developments that will push the management of born-digital materials from the project phase to the program phase:

- “Collaborative solutions for dealing with hardware and software obsolescence.”
- “More, and more appropriate, storage for born-digital materials”
- “Automation of as much of the workflow as possible.”
- “Asset-level access control to enable tiered access to restricted records.”<sup>73</sup>

There is no doubt that money is a dominant factor in how manuscript repositories choose to handle born-digital materials. As Cal Lee concludes in his lessons learned from

---

<sup>71</sup> Jeremy Leighton John, Ian Rowlands, Peter Williams, and Katrina Dean, “Digital Lives: Personal Digital Archives for the 21st Century >> an Initial Synthesis” (2010), vii, <http://britishlibrary.typepad.co.uk/files/digital-lives-synthesis02-1.pdf>.

<sup>72</sup> Nelson et al., 11.

<sup>73</sup> Ibid., 18.

working with electronic records in state government, “resources are limited, meaning is expensive.”<sup>74</sup> The costs include purchasing and keeping updated the technology necessary to preserve and provide access to electronic records along with hiring staff who are competent to work with digital objects. The question becomes whether a repository can find new resources to pay for these costs or whether existing resources have to be reallocated in order to provide additional services. Terry Cook defines “the stark bottom line: unless you can get substantial new financial and human resources, you will need to stop doing something important that you are now doing, and reallocate significant resources to electronic records, period. There is no other way.”<sup>75</sup>

In his 2003 report for OCLC Research, Brian Lavoie identifies three economic decision-makers involved in digital preservation: the rights holder (i.e., the one who holds intellectual property rights), the archive itself, and the beneficiary (i.e., the person(s) who will be better for the long-term preservation of certain digital objects).<sup>76</sup> He goes on to pose this “fundamental economic question” about digital preservation: “*Do sufficient incentives exist for relevant decision-makers to, on the one hand, identify a need to take action to preserve a given set of digital materials, and on the other, provide digital preservation services to parties interested in utilizing them?*”<sup>77</sup> In their 2010 report, the

---

<sup>74</sup> Christopher A. Lee, “Guerilla Electronic Records Management: Records Management Lessons Learned,” *Records and Information Management Report* 18, no. 5 (May 2002): 7.

<sup>75</sup> Terry Cook, “Byte-ing Off What You Can Chew: Electronic Records Strategies for Small Archival Institutions,” *Archifacts* (April 2004): 5, <http://www.aranz.org.nz/includes/download.aspx?ID=111276>.

<sup>76</sup> Brian F. Lavoie, *The Incentives to Preserve Digital Materials: Roles, Scenarios, and Economic Decision-Making* (Dublin, Ohio: OCLC Research, 2003), ii, <http://www.oclc.org/research/projects/digipres/incentives-dp.pdf>.

<sup>77</sup> *Ibid.*, 13.

Blue Ribbon Task Force on Sustainable Digital Preservation and Access suggests that unclear responsibilities, among other concerns, complicate the preservation of digital objects: “Economic analysis of digital preservation of these materials reveals structural challenges that affect all digital preservation strategies: (1) long time horizons, (2) diffused stakeholders, (3) misaligned or weak incentives, and (4) lack of clarity about roles and responsibilities among stakeholders.”<sup>78</sup>

David Bearman and Margaret Hedstrom borrow an image from David Osborne and Ted Gaebler’s book *Reinventing Government* to illustrate how they think electronic records can reinvent the archival profession. They conclude, “electronic records can be a vehicle for archives to move from rowing to steering, towards more enterprising and customer driven approaches to service delivery and towards empowering others to take action in a decentralized records management environment.”<sup>79</sup>

Also believing that born-digital materials bring the possibility of change, though not necessarily with quite as much control, Rick Barry argues that managing born-digital records has a different set of requirements than traditional paper records: “new skill sets, sophisticated, trustworthy, software tools and a great deal of our only inelastic resource – time – to carry out concentrated planning, stakeholder management, and training efforts, all with ever diminishing levels of human and capital resources being allocated to meet

---

<sup>78</sup> Blue Ribbon Task Force on Sustainable Digital Preservation and Access, *Sustainable Economics for a Digital Planet: Ensuring Long-term Access to Digital Information* (La Jolla, CA: Blue Ribbon Task Force on Sustainable Digital Preservation and Access, 2010), 1, [http://brtf.sdsc.edu/biblio/BRTF\\_Final\\_Report.pdf](http://brtf.sdsc.edu/biblio/BRTF_Final_Report.pdf).

<sup>79</sup> David Bearman and Margaret Hedstrom, “Reinventing Archives for Electronic Records: Alternative Service Delivery Options,” in *Electronic Records Management Program Strategies*, ed. Margaret Hedstrom (Pittsburgh: Archives & Museum Informatics, 1993), 98.

these challenges.”<sup>80</sup>

## 8. Provide

Wendy Duff offers an analysis of why providing access has generally been complicated for archives. Speaking at the 2002 DLM-Forum, she suggests that archivists usually focus on the act of record creation rather than on the secondary uses of these records.<sup>81</sup> And, of course, without a consideration of secondary uses, there are not really users that fall into the typical realm of manuscript repositories, for the record creators do not typically make frequent use of inactive records that they have deposited in a manuscript repository. Writing in 1994, Adrian Cunningham asserts that the preservation of electronic records is pointless without adequate provision for user access (along with the requisite training to make good use of this access). He also suggests that providing “networked access” to remote patrons should soon be a viable option.<sup>82</sup> One repository that is currently offering online access to some born-digital collections is the University of Illinois at Urbana-Champaign. However, in the case of the papers of a former chemistry professor, Stanley Smith, the repository warns the online user that some digital documents may no longer work.<sup>83</sup> The results of the 2012 ARL survey indicate that

---

<sup>80</sup> Rick Barry, “Opinion Piece – Electronic Records: Now and Then,” *Records Management Journal* 20, no. 2 (2010): 164.

<sup>81</sup> Wendy Duff, “Understanding the Information-Seeking Behaviour of Archival Researchers in a Digital Age: Paths, Processes and Preferences,” in *Proceedings of the DLM-Forum 2002* (Luxembourg: Office for Official Publications of the European Communities, 2002), 334.

<sup>82</sup> Cunningham, “Archival Management of Personal Records,” 103.

<sup>83</sup> Born-digital and digital surrogates from Stanley Smith Papers, University of Illinois Archives, <http://archives.library.illinois.edu/e-records/index.php?dir=University%20Archives/1505050/>. Based on the work at this repository, Chris Prom wrote a blog entry about their preservation and access plan. Chris Prom, “Simple E-Records Preservation and Access Plan,” *Practical E-Records* (blog), 2011,

“access to collections is not as fully developed as the management of born-digital content.”<sup>84</sup> The results of this survey go on to suggest that the two biggest access challenges are the sensitivity of materials and the lack of IT infrastructure. Along with this is the concern that automated systems are not capable of dealing with complex access restrictions with the same facility as reference desk staff have traditionally done.<sup>85</sup> In her 2010 essay, Ricky Erway raises a related question: “should digital access be subject to the same constraints as analog access?”<sup>86</sup> The documentation from the AIMS project actually divides access to electronic records into four levels:

***discover***, which would allow items to be identified by a search of metadata; ***view***, which would allow metadata to be viewed; ***render***, which would allow browser-renderable representations of content to be displayed (and would also permit searching of content alongside metadata if systems enable this); and ***download***, which would allow associated files to be downloaded.<sup>87</sup>

Despite numerous recognitions that patron access is the end goal, this element of the workflow seems to be the most difficult to solve. Although Dorothea Salo developed this analogy to describe institutional repositories, her concept of the roach motel unfortunately applies just as accurately to most born-digital materials at manuscript repositories: “Data

---

<http://e-records.chrisprom.com/recommendations/supported-formats/simple-e-records-preservation-and-access-plan/>.

<sup>84</sup> Nelson et al., 15.

<sup>85</sup> Ibid., 17.

<sup>86</sup> Ricky Erway, *Defining “Born-Digital”* (OCLC Research, November 2010), 4, <http://www.oclc.org/resources/research/activities/hiddencollections/borndigital.pdf>.

<sup>87</sup> AIMS Work Group, *AIMS Born-Digital Collections: An Inter-Institutional Model for Stewardship* (January 2012), 42, [http://www2.lib.virginia.edu/aims/whitepaper/AIMS\\_final.pdf](http://www2.lib.virginia.edu/aims/whitepaper/AIMS_final.pdf). There is also a useful table found on pages 51-55 that describes various access options as well as the factors that should be considered when deciding which options to provide for each access element.



goes in, but it doesn't come out."<sup>88</sup>

Wendy Duff has done a lot of writing over the years about archives patrons. Although no one seems to have answered her 1995 call to investigate the viability of metadata as a substitute for archival description, she herself paired with Catherine Johnson in 2002 to write about a subset of archive users, historians. In their analysis of participant comments about how they orient themselves to an archives, Duff and Johnson include a revealing quotation about the value of personal contact with a knowledgeable archivist versus merely having digital access to finding aids: “all of the . . . best digitized sources in the world are never going to replace that for me.”<sup>89</sup> Although this article was written before the recognized rise of MPLP, so perhaps the days of archivists well-versed in the intricacies of their collections have already passed, the conclusion of Duff and Johnson is still worth acknowledging: “archivists were easier to use than finding aids and could make connections to relevant material in a way that was impossible to replicate in either the printed or online aids.”<sup>90</sup> The relatively unposed and certainly unanswered question is whether researchers will be comfortable transitioning to a relatively unmediated presentation of born-digital materials, which seems to be the model currently gaining traction. While this could easily be deemed appropriate for known material searching, the mechanisms for perfecting the recall and precision results of more exploratory searches have not been developed. Conveying contextual information also needs to be addressed. The participants interviewed by Duff and Johnson speak both of

---

<sup>88</sup> Dorothea Salo, “Innkeeper at the Roach Motel,” *Library Trends* 57 (Fall 2008): 98.

<sup>89</sup> Wendy M. Duff and Catherine A. Johnson, “Accidentally Found on Purpose: Information-Seeking Behavior of Historians in Archives,” *Library Quarterly* 72, no. 4 (Oct. 2002): 484.

<sup>90</sup> *Ibid.*, 485.

the need to understand a document in the context of the entire collection as well as to gain insight into what a particular collection holds and what it lacks and why.<sup>91</sup> Another article Wendy Duff wrote several years later, in collaboration with Barbara Craig and Joan Cherry, includes one result that is alarming if it transfers to born-digital materials. In their survey, they find that Canadian historians dislike electronic reproductions of records because they find the reproduction process to be error prone.<sup>92</sup> If this distrust of the accuracy and authenticity of digitized records predisposes researchers to also question the validity of born-digital materials, a significant public relations challenge awaits archivists.

There are three primary methods of providing access to digital objects over time: generating an analog version (i.e., printing a hard copy), migrating the digital object to a format compatible with current computer systems, and emulating the original platform in which the digital object was created. For an example of a hybrid collection where the repository chose to print hard copies of electronic records, see the James Welch Papers at the Beinecke Rare Book and Manuscript Library.<sup>93</sup> For a collection where the electronic records are segregated in the finding aid, see the George Whitmore Papers at the Beinecke.<sup>94</sup>

The traditional method of migration calls for digital holdings to be migrated en

---

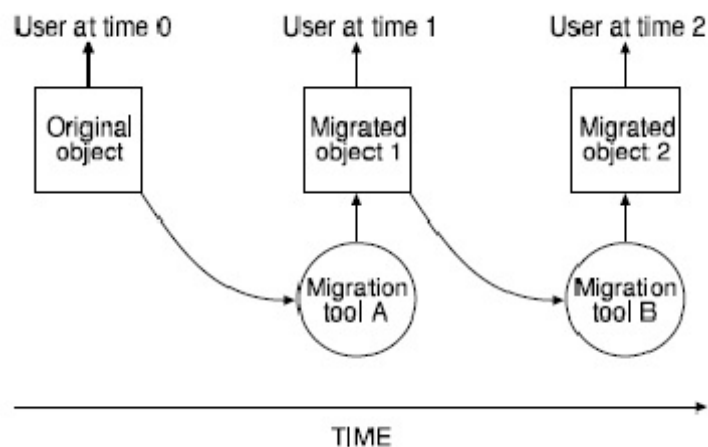
<sup>91</sup> Ibid., 488.

<sup>92</sup> Duff, Craig, and Cherry, 20.

<sup>93</sup> James Welch Papers, Yale Collection of American Literature, Beinecke Rare Book and Manuscript Library, <http://hdl.handle.net/10079/fa/beinecke.welch>.

<sup>94</sup> George Whitmore Papers, Yale Collection of American Literature, Beinecke Rare Book and Manuscript Library, <http://hdl.handle.net/10079/fa/beinecke.whitmore>.

masse when technological changes necessitate the development of a new migration tool.



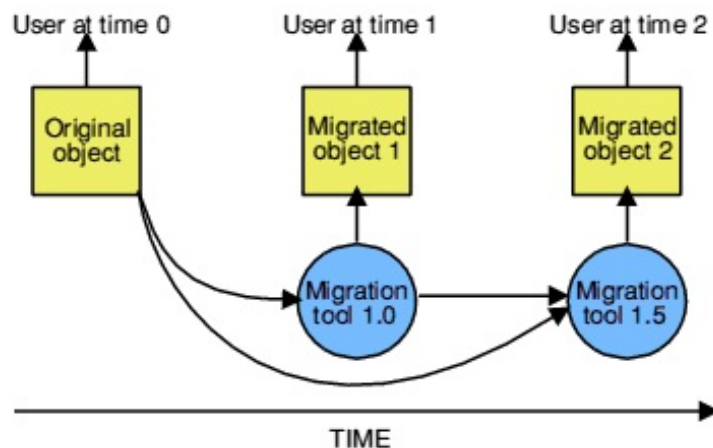
**Fig. 1** A digital object preserved using traditional migration

Three researchers for the CAMiLEON project at the University of Leeds used this graphic to demonstrate the danger of migrating from a copy of a digital object rather than from the original.<sup>95</sup> Just as analog objects degrade through a repeated process of copying, migrating from a previously migrated file can perpetuate any errors that might have occurred in a prior migration. As an alternative, this project, orchestrated by the University of Michigan and the University of Leeds, developed a preservation strategy known as “migration on request.” Through this process, the original bitstream of the digital object is preserved, and it is migrated to a usable format upon the request of a user. The project generated this graphic to illustrate how migration on request can satisfy user needs without perpetuating errors in migration that can build up due to successive migrations:<sup>96</sup>

---

<sup>95</sup> Phil Mellor, Paul Wheatley, and Derek Sergeant, “Migration on Request, a Practical Technique for Preservation,” *The Sixth European Conference on Research and Advanced Technology for Digital Libraries* (2002), 517, <http://www2.si.umich.edu/CAMiLEON/reports/migreq.pdf>.

<sup>96</sup> “Migration on Request,” CAMiLEON, accessed 12 March 2013, <http://www2.si.umich.edu/CAMiLEON/reports/mor/onrequest.html>.



Emulation still seems to be out of the reach of most manuscript repositories today. However, in their 2010 report on their Digital Lives project, Jeremy Leighton John and his colleagues refer to emulation as “an essential approach” and “the preferred access route for many eMSS scholars.”<sup>97</sup> Nonetheless, emulation has not taken hold as a preservation strategy. The 2012 ARL survey posed the question, “Which of the following delivery methods does your library use to provide access to born-digital materials?”<sup>98</sup>

Online access to a digital repository system	42	66%
In-library access on dedicated computer workstation	31	48%
In-library access using portable media accessed through the users' personal computer	22	34%
Third-party access & delivery system	18	28%
Online access to a file space	15	23%
In-library access to records in an emulated environment	1	2%
Online access to records in an emulated environment	1	2%
We do not provide access at this time	13	20%
Other delivery method	10	16%

These results indicate concretely that only one participating repository practices emulation.

---

<sup>97</sup> John et al., xiii.

<sup>98</sup> Nelson et al., 71.

Much of the excitement about born-digital records relates to the searchability that pervades a digital environment. As Elizabeth Dow suggests, the level of search available in electronic records provides “a quality of intellectual access almost impossible to deliver in an analog document.”<sup>99</sup> In response to the 2007 survey by Susan Davis, a public university archivist commented, “I am inclined to accept some digital materials that I might be reluctant to accept in paper format. This is because ephemeral materials take on new value when they are part of a body of material that can be searched using full-text search engines.”<sup>100</sup> NARA seems to be embracing the greater access that can be provided for electronic records. In a 2009 workshop, Kenneth Thibodeau reports that “NARA has decided that the public will need to go to only one place in ERA [Electronic Records Archives] for access to all records which are publicly available, even when there are some restrictions on content. In the public access part of ERA, anyone will be able to find information about any records we preserve, both traditional and digital, federal, presidential, and those Congressional records we are allowed to release to the public.”<sup>101</sup>

Another model for providing access to electronic records that emerged in the early 1990s was that of distributed custody. Margaret Hedstrom and David Bearman were some of the loudest advocates for such a system, arguing,

It is easy to provide copies of electronic records to numerous ‘outlets’ at the same time and through metadata management to support item-level description of records without archivists engaging in item-level description. By employing networks we could greatly expand ability of individual citizens to get information from archives. Distributed points of access could also be supported by a proactive

---

<sup>99</sup> Dow, 25.

<sup>100</sup> Davis, 185.

<sup>101</sup> Thibodeau, 8. The description of NARA’s prototype Online Public Access can be read at <http://www.archives.gov/research/search/about-opa.html>.

reference service staffed by public librarians and other information providers rather than archivists.<sup>102</sup>

However, as already acknowledged in the Get section of this literature review, this noncustodial model is inherently problematic for personal papers.

One arena of the literature on born-digital materials that is lacking is research about the users of these records. This can somewhat be explained by the relative lack of access to these records to this point along with their offering a glimpse into the relatively recent past, which is not necessarily the time period generally most appealing to the patrons of manuscript repositories. Writing in 2004 about historians in the United Kingdom as archival researchers, Ian Anderson suggests, “historians’ publications are one of the most widely distributed means of archives manifesting their cultural and societal value. It is through historians’ research that archival data and information becomes knowledge, developing meaning and understanding about ourselves, our past, and our place in the world.”<sup>103</sup> The white paper produced by the AIMS project acknowledges that the ability to make born-digital materials discoverable and accessible online opens up many possibilities, but doing so also de-personalizes the archival research process by potentially removing the archivist from that process, thereby eliminating one means of ensuring the appropriate access and use of materials and increasing “the risk of misuse or abuse of copyrighted or sensitive information.”<sup>104</sup> Despite the uncharted terrain, it is vital

---

<sup>102</sup> David Bearman and Margaret Hedstrom, “Reinventing Archives for Electronic Records: Alternative Service Delivery Options,” in *Electronic Records Management Program Strategies*, ed. Margaret Hedstrom (Pittsburgh: Archives & Museum Informatics, 1993), 88.

<sup>103</sup> Ian G. Anderson, “Are You Being Served? Historians and the Search for Primary Sources,” *Archivaria* 58 (Fall 2004): 82.

<sup>104</sup> AIMS, 56.

to consider how born-digital materials held in archives might be used. Eric Ketelaar refers to the “affordances of digital technologies” that “stimulate people to create content differently and to use documents differently in different collaborative and distributed networks.”<sup>105</sup> Without consideration of the ways that the work of records creators continues to change as well as the ways that the work of records users continues to change, archives themselves could truly become relics of the past.

## 9. Case Studies

In 2006, the Harry Ransom Center at the University of Texas at Austin acquired the archive of hypertext author Michael Joyce. His papers included both paper-based and born-digital materials, and Catherine Stollar Peters explains that “while the materials would be housed separately, we chose to arrange all of his materials using the same functional series, as opposed to series based on format, to demonstrate the original order in which Michael Joyce created his papers.”<sup>106</sup> The electronic records are stored in a DSpace environment, so the Ransom Center created crosswalks from the DSpace hierarchies to traditional archival levels.<sup>107</sup> Certain metadata fields could not be populated automatically upon ingest, so at first, these were being entered manually; eventually, they abandoned the entering of subject metadata at the item-level because it

---

<sup>105</sup> Eric Ketelaar, “Archives in the Digital Age: New Uses for an Old Science,” *Archives & Social Studies* 1, no. 0 (March 2007): 174.

<sup>106</sup> Peters, 27.

<sup>107</sup> *Ibid.*, 25. Peters also provides a lengthy description of the characteristics of DSpace that made it appropriate for this project. See pp.24-25.

was too time consuming.<sup>108</sup> Another time consuming task was weeding out the duplicate files, which had been generated by the donor as backup files and stored in different locations; however, the Ransom Center decided it took less time to weed them than it would to assign them metadata. At the time Peters wrote this article, the Ransom Center intended to provide access to these files through emulation, but the repository had not yet solved this piece of the puzzle. Her conclusions about digital archivy include a preference for “automated, open-source tools,” a recognition of the need for “specialized knowledge and specialized staff” to handle digital preservation, a realization that archival practices will need to adapt to handle the unique needs of born-digital materials, and the need to have an institutional commitment and “clear policies and procedures” in place before beginning a digital preservation project.<sup>109</sup>

In late 2006, the Manuscript, Archives, and Rare Book Library (MARBL) at Emory University acquired the papers of Salman Rushdie, including over one hundred linear feet of paper materials as well as a large born-digital component consisting of four computers, one hard drive, and several disks.<sup>110</sup> One of the most significant decisions the MARBL faced with this collection was how to provide users access; ultimately, they chose to implement a combination of migration and emulation. They note that an advantage of emulation is that “identifying, categorizing, preserving, and providing access to the materiality of born-digital personal archives can be of equal importance as

---

<sup>108</sup> Ibid., 29.

<sup>109</sup> Ibid., 33-34.

<sup>110</sup> Laura Carroll, Erika Farr, Peter Hornsby, and Ben Ranker, “A Comprehensive Approach to Born-Digital Archives,” *Archivaria* 72 (Fall 2011): 63-64.



attending to the context.”<sup>111</sup> Ultimately, they have provided patrons access points through a searchable database, the emulation, and a traditional finding aid, all of which can be accessed on a computer workstation in the MARBL reading room.<sup>112</sup> They summarize the lessons learned from this project: “the necessity of collaboration; the need to engage with other fields and communities; the role of pre-acquisition consultations with donors and content creators; the importance of triage and appraisal; the value of collection-specific processes and workflows; and the need for co-operative tool development.”<sup>113</sup>

In 2009, Michael Forstrom published a case study about the Beinecke Rare Book and Manuscript Library.<sup>114</sup> He addresses the authenticity requirements of born-digital records and concludes that the InterPARES requirements for authenticity can be applied to electronic records in a manuscript repository and that the rules laid out in *Describing Archives: A Content Standard (DACS)* also apply. He does suggest modifying a descriptive element to incorporate information about “refreshment or ingest into a digital repository.”<sup>115</sup> In a footnote, Forstrom also makes an interesting suggestion for further work needed, saying that it would be useful for electronic records to be linked from the finding aid, although this would necessitate some process for remote authentication of the patrons.<sup>116</sup>

---

<sup>111</sup> Ibid., 79.

<sup>112</sup> Ibid., 80.

<sup>113</sup> Ibid., 88-89.

<sup>114</sup> Michael Forstrom, “Managing Electronic Records in Manuscript Collections: A Case Study from the Beinecke Rare Book and Manuscript Library,” *American Archivist* 72 (Fall/Winter 2009): 460-77.

<sup>115</sup> Ibid., 475-76. The Authenticity Task Force Report of the InterPARES project can be read at [http://www.interpares.org/book/interpares\\_book\\_d\\_part1.pdf](http://www.interpares.org/book/interpares_book_d_part1.pdf).

<sup>116</sup> Forstrom, 477 (note 75).

In 2009, Charles E. Bracker made a donation of 30,000 digital photographs of orchids to the Ball State University Libraries. Because these photographs had not been organized or labeled in any way by the donor, the digital projects librarian and the archivist who worked on this project had to categorize the images and develop a file management strategy. They provide access to selected images through CONTENTdm, but only after the images have been individually examined and edited. The collection has brought added notoriety to the repository, and some valuable lessons were learned about collaboration. They also estimate that the storage space for this born-digital collection is substantially smaller than a comparable collection that has been digitized from analog sources. However, the individual attention that was necessary to create metadata for each image, along with the aforementioned editing, makes this workflow seem unscalable.<sup>117</sup>

In 2011, Ben Goldman wrote of the experiences of the American Heritage Center at the University of Wyoming in beginning to process born-digital materials. He suggests four simple, achievable steps: (1) inventory born-digital materials in the collection to generate an estimate of the quantity of storage space required, (2) determine an appropriate storage mechanism (and plan for one archival master copy and one access copy of each file), (3) transfer digital objects from removable media to the storage system (capturing metadata and documenting actions at the time), and (4) develop policies.<sup>118</sup>

---

<sup>117</sup> Amanda A. Hurford and Carolyn F. Runyon, "New Workflows for Born-Digital Assets: Managing Charles E. Bracker's Orchid Photographs Collection," *Computers in Libraries* 31, no. 1 (2011): 6-10, 40.

<sup>118</sup> Goldman, 16. Goldman elaborates on each of these steps on pp.16-23.

### Research Design

---

*“Electronic information now forms an important part of the documentary memory of our time. We must be able to transmit through time those significant electronic traces of ourselves which form part of a coherent information heritage.”*<sup>119</sup>

---

Upon further investigation of the archival literature, it appears that the literature itself helps to explain why the discussion about born-digital materials has taken so long to take hold within the manuscripts community. In her critique of the new paradigm for electronic records, Linda Henry argues that these writers have created a tight circle wherein they cite each other and rarely look to historical sources in the literature, which has served “to exclude the majority of archivists from the dialogue about electronic records, rather than invite them to participate in it.”<sup>120</sup> She goes on to point out that “their narrow definition of a record and their arguments against archival custody of electronic records pertain, at best, only to organizational archives. These arguments do not hold any promise for noninstitutional archives and manuscript repositories. The new paradigm excludes them.”<sup>121</sup>

Having been exposed to a fair bit of the literature about electronic records during my studies in archives and records management, I also recognize that there are many voices not currently represented. After attending the meeting of the Manuscript

---

<sup>119</sup> Grimard, 167.

<sup>120</sup> Linda J. Henry, “Schellenberg in Cyberspace,” *American Archivist* 61 (Fall 1998): 326.

<sup>121</sup> *Ibid.*, 327.

Repositories Section at the Society of American Archivists (SAA) annual meeting in San Diego, I was further inspired to discover how (or whether) manuscript repositories are currently handling born-digital materials. My advisor, Jackie Dean, sits on the steering committee of the Manuscript Repositories Section, so she put me in contact with Chris Burns, the chair of the section. He shared with me the section's plan for an electronic records initiative, called the Jump In Initiative. The Manuscript Repositories Section challenged SAA members to begin managing born-digital content and specified steps drawn from Ricky Erway's *You've Got to Walk Before You Can Run* report.<sup>122</sup> The three of us discussed what research I could do that would complement this effort by the section. So in addition to conducting a broad literature review, including case studies that were based on manuscript repositories, I chose to survey the membership of the SAA Manuscript Repositories section discussion list.<sup>123</sup> On 3 February 2013, I sent an email to the list requesting that any repositories not currently handling born-digital materials answer a two-question survey online and that any repositories already processing electronic records contact me to set up a time for an interview.<sup>124</sup> The SAA web site lists 884 members on the roster of the Manuscript Repositories section, though I happen to know through automatic replies to my message that some of these people have retired. There are also cases in which more than one individual from a single repository belongs

---

<sup>122</sup> Manuscript Repositories Section, "Jump In Initiative," Society of American Archivists, last modified 30 October 2012, <http://www2.archivists.org/groups/manuscript-repositories-section/jump-in-initiative>.

<sup>123</sup> An application was submitted by the author on 18 December 2012 to the Institutional Review Board at the University of North Carolina at Chapel Hill. After review, a 2 January 2013 message confirmed that this study "does not constitute human subjects research as defined under federal regulations [45 CFR 46.102 (d or f) and 21 CFR 56.102(c)(e)(1)] and does not require IRB approval."

<sup>124</sup> The full text of the message can be viewed in Appendix A.

to the section.

Five members took my survey for repositories that are not currently processing born-digital materials.<sup>125</sup> My intent with this survey was to attempt to pinpoint the roadblocks to handling born-digital materials. Five members who are currently processing electronic records provided me feedback through email or phone interviews. I was also able to arrange interviews with four additional archivists through other contacts.<sup>126</sup> My intent with these interviews was to attempt to ferret out policies or procedures that are working effectively along with challenges that persist. In order to allow for better feedback, I provided the respondents my questions before the interview; I also recorded the phone calls to ensure the accuracy of my notes of our conversation. Arguably, there is a self-selection bias to my pool of respondents; several did mention knowing my advisor or having a connection to the School of Information and Library Science at the University of North Carolina at Chapel Hill. While the numerical response rate to my query was minimal, the information gathered was still revealing.

I also emailed those who have registered for the Jump In Initiative.<sup>127</sup> Out of the thirty-three people registered, seven have provided me additional feedback about their work on this project. In this case, my intent was to gain an early glimpse of their findings and to determine their motivations for participating in the initiative.<sup>128</sup>

---

<sup>125</sup> The instrument that was mounted through Qualtrics can be viewed in Appendix B.

<sup>126</sup> While some questions were tailored to the particular collection, the general questions that were asked of all archivists experienced in processing born-digital materials can be seen in Appendix C.

<sup>127</sup> The text of this message can be seen in Appendix D.

<sup>128</sup> Although this would obviously extend beyond the end of my tenure as a master's student, given the opportunity, ideally I would like to make this research be the first step in a longitudinal study.

## Findings

---

*“Do you know where you’re going to? / Do you like the things that life is showing you  
Where are you going to? / Do you know...?  
Do you get / What you’re hoping for  
When you look behind you / There’s no open door  
What are you hoping for? / Do you know...?”<sup>129</sup>*

---

### ***1. Survey of Repositories Not Currently Processing Born-Digital Materials***

Three of the responses to the survey came from special collections repositories at a university; one came from a government institution; and one came from a public library.

The first question of the survey asked respondents to identify factors that have limited the ability of their manuscript repository to process born-digital records. The options listed were training, costs, concerns about providing access, time, and inadequate administrative support. Each of these factors was rated by the respondents as having some significance in their inability to begin processing electronic records. Each possible response was weighed from 1 to 4, from no significance to highest significance. Given this framework, inadequate administrative support returned the highest result with an average score of 3.4. The next highest result was a 3.2 average score for training to know how to handle born-digital records.

---

<sup>129</sup> “Theme from *Mahogany* (Do You Know Where You’re Going To)” is a song written by Michael Masser and Gerald Goffin and recorded by Diana Ross as the theme to the 1975 Motown/Paramount film *Mahogany*.

#	Question	A small consideration	A significant consideration	A monumental obstacle	Total Responses	Mean
1	Training to know how to handle born-digital records	0	4	1	5	3.20
2	Cost of handling born-digital records	1	4	0	5	2.80
3	Concerns about providing patrons access to born-digital records	2	3	0	5	2.60
4	Time required to process born-digital records alongside backlog of paper records	1	4	0	5	2.80
5	Inadequate administrative support	1	1	3	5	3.40

Statistic	Training to know how to handle born-digital records	Cost of handling born-digital records	Concerns about providing patrons access to born-digital records	Time required to process born-digital records alongside backlog of paper records	Inadequate administrative support
Min Value	3	2	2	2	2
Max Value	4	3	3	3	4
Mean	3.20	2.80	2.60	2.80	3.40
Variance	0.20	0.20	0.30	0.20	0.80
Standard Deviation	0.45	0.45	0.55	0.45	0.89
Total Responses	5	5	5	5	5

The second question of the survey asked the respondents to consider what scenarios would facilitate the processing of born-digital records by their manuscript repository. Once again, the responses were values on a Likert scale from 1 to 4, from no significance to highest significance. By far the option receiving the highest score (3.8) was the development of acquisition, preservation and access policies for born-digital materials. And by far the option receiving the lowest score (2.0) was the provision of patron training in how to access born-digital records.

#	Question	Not an aid	A small aid	A significant aid	An extremely important solution	Total Responses	Mean
1	Being able to bring in a consultant to provide assistance in establishing a workflow for processing born-digital records	0	1	4	0	5	2.80
2	Being able to study established workflows from other similarly sized repositories to learn from their processes	0	1	3	1	5	3.00
3	Being able to acquire grant funding to help cover costs	1	0	1	3	5	3.20
4	Being able to provide training for reference staff in how to facilitate patron access to born-digital records	0	1	4	0	5	2.80
5	Being able to provide training for patrons in how to access born-digital records	0	5	0	0	5	2.00
6	Being able to hire an electronic records archivist who could focus only on the processing and access of born-digital records	1	0	2	2	5	3.00
7	Being able to develop an institutional policy governing the acquisition, preservation, and access of born-digital materials	0	0	1	4	5	3.80

Statistic	Being able to bring in a consultant to provide assistance in establishing a workflow for processing born-digital records	Being able to study established workflows from other similarly sized repositories to learn from their processes	Being able to acquire grant funding to help cover costs	Being able to provide training for reference staff in how to facilitate patron access to born-digital records	Being able to provide training for patrons in how to access born-digital records	Being able to hire an electronic records archivist who could focus only on the processing and access of born-digital records	Being able to develop an institutional policy governing the acquisition, preservation, and access of born-digital materials
Min Value	2	2	1	2	2	1	3
Max Value	3	4	4	3	2	4	4
Mean	2.80	3.00	3.20	2.80	2.00	3.00	3.80
Variance	0.20	0.50	1.70	0.20	0.00	1.50	0.20
Standard Deviation	0.45	0.71	1.30	0.45	0.00	1.22	0.45
Total Responses	5	5	5	5	5	5	5

## 2. Interviews with Repositories Currently Processing Born-Digital Materials

Seven archivists were gracious enough to share their time and expertise.<sup>130</sup>

Several of these repositories, specifically the University of Mississippi Archives and Special Collections and the American Heritage Center at the University of Wyoming, began working with electronic records because they anticipated acquiring collections that have significant amounts of digital content, so they wanted to be proactive and have workflows in place to handle those born-digital materials. At the New York Public Library, their work was sparked by the hiring of a digital archivist as well as the anticipation of receiving more born-digital collections. Much of the work at the

---

<sup>130</sup> Kathryn Michaelis, email message to author, 7 February 2013; Glynn Edwards and Laura Williams, call with author, 8 February 2013; Don Mennerich, call with author, 14 February 2013; Mark Greene, call with author, 18 February 2013; Jeff Thomas, call with author, 28 February 2013; Patrick Cullom, interview by author, 4 March 2013.



American Heritage Center, such as designing a workflow and completing an inventory, also began after the hiring of a digital programs manager, and the expansion of this work is awaiting the hiring of his replacement.

Glynn Edwards made a presentation at the 2012 Rare Books and Manuscripts Conference where she provided an overview of the born-digital workflow in Special Collections at the Stanford University Libraries. It includes using forensic software such as FTK Imager to create checksums for authenticity and to create directory listings along with Archivists' Toolkit for registering the objects and creating finding aids.<sup>131</sup> In the case of the Stephen Gould Papers, Stanford chose to assert intellectual control over the electronic records by mirroring their organization to that of the physical files.<sup>132</sup> Don Mennerich at the New York Public Library is also using forensic tools to extract metadata and assert intellectual control over digital objects.

Some hesitated to compare the processing time of born-digital collections to analog collections because there is no precise metric for doing so, but Jeff Thomas contends that "processing digital documents consumes more time than paper records. Computer files simply take a lot longer to browse through than flipping through paper."<sup>133</sup> While several repositories began their work by generating item-level metadata for digital objects, they have come to the conclusion that this is not a scalable approach; Thomas is

---

<sup>131</sup> Glynn A. Edwards, "Enigma of Email" (presentation at Rare Books and Manuscripts Conference, San Diego, CA, June 2012), 3, <https://docs.google.com/open?id=0B35ZUUIIdskPhWGFYdTFjcDJCYVE>.

<sup>132</sup> The processing plan for the Stephen Jay Gould papers can be seen in AIMS, 99-103. The Processing Workflow can be seen on pages 120-23.

<sup>133</sup> Jeff Thomas, "What Do I Do With The Black Box? Processing the Electronic Records in the Deborah Pryce Papers," *Congressional Papers Roundtable Newsletter*, forthcoming, 13.

especially adamant that the processing and arranging of electronic records must occur no lower than the folder level.

For most of these repositories, providing access is the last piece of the puzzle. To this point, there has been no notable pressure from patrons to provide access to born-digital materials online, likely because these collections tend to be under-described, so users have to luck into locating the resources in finding aids. Two of the repositories are creating PDF access copies of documents. In the case of the Ohio Congressional Archives, this decision was made both because the PDF format is more secure than the Microsoft Word format in which documents were received and because the PDF files can then be grouped into portfolios for easy online access and keyword searching. The digital objects in the Stanford collections are searchable on a media cart in the reading room. This computer is not on a network and there are no ports for external drives, so patrons must flag any items that they wish to print and get the assistance of the staff. The New York Public Library (NYPL) uses Quick View Pro for file viewing and migrates Microsoft Word documents to ensure they do not lose their search functionality and is in the process of setting up a media workstation in their reading room. They may in the future try to virtualize this workstation for remote access.<sup>134</sup> While the NYPL does make an effort to remove personally identifying information such as that in medical records, using Bulk Extractor to redact information out of the disk image, they also recognize that it is impossible to sanitize everything. In the case of digital photographs in the collection of the University of North Carolina at Chapel Hill (UNC-CH), the intention is to upload

---

<sup>134</sup> Seth Shaw, the Electronic Records Archivist at Duke University, is working on a prototype of a system to provide virtual access to records. Given the success of the Duke Data Accessioner that he developed, this is an promising initiative. Interview by author, 7 March 2013.

records into the CONTENTdm system for access but to suppress the image when the donor agreement with a photographer requires permission for a patron to use an image. Patrick Cullom acknowledges that young researchers, especially, presume that they should have a right to access digital stuff merely because the technology makes it easily available, but he goes on to point out that archives have a responsibility, as they always have, to protect the items deposited with them; therefore, an ability to access a digital object does not necessarily equate to a right to access that object.

In completing its work on the AIMS project, Stanford Libraries wrote guidelines for creating agreements that point out the importance of documenting issues relating to ownership, exclusivity, and preservation, to name a few.<sup>135</sup> In all cases with these interviews, the policy work has followed the creation of a basic workflow process. Mark Greene has long advocated early involvement with donors, and he sees no reason not to transfer that practice to donors of born-digital materials. Writing about his time at the American Heritage Center, Ben Goldman speaks of the importance of conducting preacquisition appraisals. Jeff Thomas strongly favors precustodial interventions with Congressional offices to educate them about the importance of creating an organized foldering system and following file naming conventions.<sup>136</sup>

Some helpful collaborations have occurred with organizations outside of the manuscript repository through projects such as AIMS and BitCurator. But most of the work that is being done at this point seems to remain within institutional boundaries, with

---

<sup>135</sup> “Guidelines for Creating Agreements at Stanford University,” in AIMS, 117-19. More general guidelines for a collection development policy can be found on page 6; the key elements of a donor agreement are listed on pages 9-10.

<sup>136</sup> Thomas, 13.

a liaison to the IT department being a more common type of collaboration.

One of the suggestions for improving the handling of born-digital materials is to create an institutional repository that could be responsible for the long-term maintenance of the digital objects that are ingested. Another very practical suggestion is to accumulate now the equipment that will be necessary to access files later; for instance, UNC-CH is building a “Frankenstein” machine that will have the capability to access files from various types of digital media cards and other media formats. Another repository recognizes that more money and more staff are necessary to handle born-digital materials effectively. One respondent points out the need to get electronic records documented in the processing manual for his repository. Mark Greene is emphatic that more people need to be competent and comfortable in working with born-digital materials rather than isolating that expertise, but it seems more common that fewer people are involved in working with electronic records during the initial planning and implementing stages.

Given the fact that for many years NARA provided leadership for the archival profession, I contacted archivists at the two most recent presidential libraries<sup>137</sup> to find out how they handle born-digital materials and to determine if there are any lessons that can be generalized to other types of repositories. While the Clinton Presidential Library does have a database of the emails (plus their attachments) that were generated by the Clinton White House from 1993-2001, Adam Bergfeld explained that they cannot provide electronic access to these materials for security reasons. Materials are accessed through Freedom of Information Act requests, at which point he searches the repository for

---

<sup>137</sup> Adam Bergfeld, call with author, 11 March 2013; Sarah Ticer, email message to author, 15 March 2013.

relevant records (both electronic and analog) and provides paper copies to the researcher.<sup>138</sup> However, Sarah Ticer at the George W. Bush Presidential Library explained that their goal is to make processed born-digital records available through NARA's Online Public Access Catalog. So even though there will not be direct links from the online finding aids hosted on the Bush Library web site, there will be a mechanism for online delivery of records.

### ***3. Feedback from Jump In Initiative Participants***

Five of the manuscript repositories that provided feedback are housed in universities. Another is a historical society, and one is a religious organization.<sup>139</sup>

There were numerous explanations of their motivations to participate in the Jump In Initiative. Respondents A and G both mentioned the importance of knowing this is a community of other archivists working through the same issues at the same time who can be looked to for guidance and support. Respondent A elaborated to say that "now that I've said I'll do this and I am part of this group, I feel obligated to finish." Respondent B asserted that having the SAA sponsor this activity gave an aura of "credibility/authenticity/authority" that helped her sell the project to her superiors. Respondent C suggested that the requirements of the initiative were simple enough that "there is little to lose and easily something to gain." Respondents B and D both acknowledged the possibility of winning tuition to a Digital Archives Specialist (DAS) class helped seal

---

<sup>138</sup> For more information on FOIA, the Clinton Library has a web page that explains its applications to presidential records. See <http://www.clintonlibrary.gov/foia.html>.

<sup>139</sup> Given the fact that these responses are more reflective of being new to handling born-digital materials than they are informative about a particular repository, the author chose to anonymise these results.

their decisions to participate. Respondents A and E mentioned that having a deadline can be helpful when confronting a difficult task. Respondents B, E and F all indicated they recognized they needed to conduct a survey of the electronic media in their collections, and this initiative gave them the incentive to do so.

The early results of the surveys of computer media were also wide-ranging. Respondent A actually found fewer computer media than she had presumed. Respondent B, on the other hand, found many more media than expected, but she was relieved to find that more of them are of the CD and DVD variety rather than more difficult to access 3.5-inch floppy disks. Several respondents found fugitive media in collections that had not been properly identified in finding aids or other accessioning materials. Respondent C admitted that “I’m afraid that we have in the past adopted the ‘file it and forget it’ approach to the problem; we’re in for nothing but surprises in earlier acquisitions.” Along with media that were not counted in the finding aids, Respondent F found instances where born-digital materials were printed at the time of donation, and the physical media were never deposited. Respondent G indicated that simple searches of the finding aids for terms like “CD” or “computer” were not sufficient to find all of the computer media in their collections. Respondent D found a plethora of CDs and DVDs, many of which are “commercial appearing disks,” which raises copyright issues.

More of the respondents (A, B, C, D, and F) indicated that they are mostly doing the work first before designing overarching policies. A few of the institutions already have some relevant policies in place; for instance, a retention schedule governs the accessions of Respondents A and E. Respondent G indicated that her repository is “looking more at the big picture and working our way down,” so they have incorporated

language about born-digital items into their donor agreement and have developed a digital strategy that outlines their mechanisms for establishing a trustworthy repository.

Respondent F mentioned that they have experimented with using the Duke Data Accessioner and are working to construct their own “forensic and quarantine machine to use as a point of ingest.” Finally, although the policy piece is not necessarily the first piece of the puzzle for Respondent E, she does anticipate that her work for this initiative will be helpful in developing a protocol for accessions. She has in her collection some legacy electronic media about which no decision was ever made whether or not they were record materials; but now in the future, she will have a framework to help determine which digital objects are truly worthy of being kept, thereby limiting the electronic records on which she needs to perform preservation measures. Several repositories deal with born-digital materials on an as-requested basis; for Respondent D this means if there are no requests, there likely will be no preservation steps taken, and if there are requests for immediate use, this may entail serving content from the original disks. Respondent C indicated that her repository has a “standing practice of having surrogates made only when readers request access to material on obsolete media,” although they do “proactively create surrogates” for some “very high-use collections.”

## Discussion

---

*“Any genuine case of thinking starts, for example, with considerations which as they stand are fragmentary and discrepant. Thinking then has the task of effecting unification in a single coherent whole. In this sense the goal of all thinking is the attaining of unity.”*<sup>140</sup>

---

Anne Gilliland-Swetland best describes the feelings of many archivists when confronted with born-digital materials: “confused, anachronistic, insecure, even stupid. Like a rabbit out of its burrow on a dark night, many an archivist, faced with venturing into the realm of electronic records, has found herself or himself frozen in the lights of oncoming traffic, unable to move either forward or backward, doomed to be roadkill on the information superhighway.”<sup>141</sup>

Ben Goldman describes the quest to resolve the issues surrounding born-digital materials as a “Quixotic one,” with archivists waiting “for that one perfect, affordable, all-encompassing solution for electronic records.”<sup>142</sup> In his musings on the value and values of archivists, Mark Greene interjects a thought that has interesting application to born-digital materials. He suggests that archivists “tend to focus too much on our

---

<sup>140</sup> John Dewey, “Context and Thought,” *University of California Publications in Philosophy* 12, no. 3 (1931): 209.

<sup>141</sup> Anne Gilliland-Swetland, “Digital Communications: Documentary Opportunities Not to Be Missed,” in *American Archival Studies: Readings in Theory and Practice*, ed. Randall C. Jimerson (Chicago: Society of American Archivists, 2000), 589-90.

<sup>142</sup> Goldman, 11.



processes and not enough on our purpose.”<sup>143</sup> An earlier article by Greene provides a broad answer to the question of what purpose archives serve: “the archival mission is about meaning.”<sup>144</sup>

The time for passivity has elapsed. Even writing in 1994, Adrian Cunningham suggests the metaphor of a ticking time bomb with regards to electronic personal records in a precustodial environment, but he concludes that “the approach has been to ignore it in the hope that by the time the suspect device is offered for transfer someone will have discovered an easier way of defusing it than is currently available. This approach may be tantamount to the reckless endangerment of both the records themselves and to the very future of those institutions which collect personal records.”<sup>145</sup> Patrick Cullom adds an anecdote from the visual materials realm, suggesting that the archives profession tends to be wary of moving too fast with change because they have been burned in the past with decisions, such as switching from nitrate to safety film.<sup>146</sup> But the luxury of a wait-and-see attitude has long since passed.

In his 2009 article, Adrian Cunningham concludes with a simple to-do list for the archival profession: “conduct more research into the dynamics of personal record keeping, the societal warrants for personal record keeping, and the functional

---

<sup>143</sup> Mark. A. Greene, “The Power of Archives: Archivists’ Values and Value in the Postmodern Age,” *American Archivist* 72 (Spring/Summer 2009): 18.

<sup>144</sup> Mark. A. Greene, “The Power of Meaning: The Archival Mission in the Postmodern Age,” *American Archivist* 65 (Spring/Summer 2002): 50.

<sup>145</sup> Cunningham, “Archival Management of Personal Records,” 99. Even more colorfully, Cunningham explains in a later article that the Australian phrase for postponing a decision is “putting them on the never-never.” Adrian Cunningham, “Waiting for the Ghost Train,” 56.

<sup>146</sup> Patrick Cullom, interview by author, 4 March 2013.

requirements for evidence in personal record keeping.”<sup>147</sup> Yet these directives are probably still a little too heavily tilted toward theory than practice. My research returned sharp differences between those repositories engaged in handling born-digital materials and those repositories yet to enter the realm of born-digital archivy; for example, the repositories not currently working with born-digital materials indicated that they think policies should be in place before processing records, while those repositories already engaged in the work find it more effective to divine appropriate policies only after understanding the various attendant issues of handling born-digital records. Given these differences, it seems imperative that the archival literature begin to reflect more of the common sense approaches developed by those in the trenches. Just as importantly, there must be more research into the users of born-digital materials. As Ian Anderson concludes, “if archives are to maintain their high standards of service in the digital age, it is fundamental that these are based on a thorough understanding of users’ information-seeking behaviour and requirements.”<sup>148</sup>

An invitational symposium at the University of Maryland in May 2010 entitled *Computer Forensics and Cultural Heritage* prompted the generation of this list of recommended next steps:

- “1. Develop policy frameworks and best-practice agreements for donor relations, liability, workflows, and researcher access.”
- “2. Develop regional networks for collaboration.”
- “3. Define requirements for and develop new tools.”
- “4. Aid in articulating a scholarly research agenda.”
- “5. Collect more stories and case studies.”
- “6. Facilitate training.”

---

<sup>147</sup> Cunningham, “Waiting for the Ghost Train,” 63.

<sup>148</sup> Anderson, 83.

“7. Encourage cross-publication of research literature and cross-promotion of professional events.”

“8. Pursue terminology mapping.”<sup>149</sup>

Three years later, some work has been accomplished on these steps, but there is still much to do. One example of a project that is attempting to develop new tools is the BitCurator Project. The BitCurator Project aims to help libraries, archives, and museums (LAMs) in “(1) integrating digital forensics tools and methods into the workflows and collection management environments of LAMs and (2) supporting properly mediated public access to forensically acquired data.”<sup>150</sup> Margaret Hedstrom has been advocating since the 1990s for the field testing of the theoretical models of how to manage electronic records.<sup>151</sup> And while the SAA is maintaining a collection of case studies related to born-digital materials in campus archives, most of these do not directly relate to the types of records commonly collected by manuscript repositories.<sup>152</sup> Noticeably absent from the focus of each of these efforts is attention on the users of electronic records.

Writing in 1998, Philip Bantin identifies the “new skills” that will help archivists handle electronic records: “a basic knowledge of how automated systems are created and work; a more detailed knowledge of data and information management principles and techniques; experience implementing functional decomposition and business process modeling methodologies; and knowledge of computer-based information systems,

---

<sup>149</sup> Kirschenbaum, Ovenden, and Redwine, 62-64.

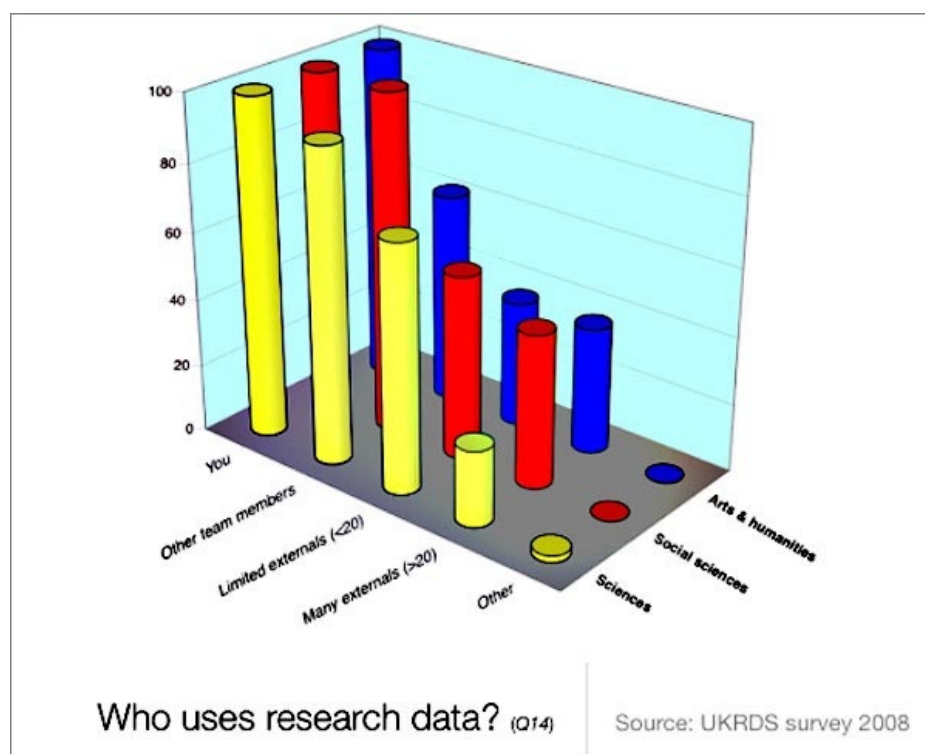
<sup>150</sup> Christopher A. Lee, et al., “BitCurator: Tools and Techniques for Digital Forensics in Collecting Institutions,” *D-Lib Magazine* 18 (May/June 2012), <http://www.dlib.org/dlib/may12/lee/05lee.html>.

<sup>151</sup> Bantin, 29.

<sup>152</sup> “Campus Case Studies,” Society of American Archivists, accessed 28 February 2013, <http://www2.archivists.org/publications/epubs/Campus-Case-Studies>.

particularly metadata systems, such as data dictionaries and information resource dictionary systems.”<sup>153</sup> While there has been much written on these topics in the intervening years, and there have been some efforts to address the educational needs of digital archivists through programs like SAA’s DAS certificate or the DigCCurr program at UNC-CH, these skills remain outside of the grasp of most current archivists in manuscript repositories.

The scientific community has already begun addressing many of the issues surrounding the preservation of born-digital materials. For instance, many grant funders now require data sets to be made public.<sup>154</sup>



<sup>153</sup> Bantin, 30.

<sup>154</sup> For example, the Data Archiving Policy of the National Science Foundation can be viewed at <http://www.nsf.gov/sbe/ses/common/archive.jsp>. The journal *Nature* also requires that data and materials be made public before articles will be published. See <http://www.nature.com/authors/policies/availability.html>.

Unfortunately, according to this graph from the 2008 UK Research Data Service feasibility study presented by Neil Beagrie, Robert Beagrie and Ian Rowlands, the arts and humanities field tends to re-use research data in a manner that differs from the sciences<sup>155</sup>, so the models established by repositories of scientific data may not directly translate to manuscript repositories. Nonetheless, the principle of engendering cooperation among records creators, publishers, other organizations, and data repositories bodes well for the long-term preservation of digital objects. Perhaps archivists could initiate an alliance with the writers' guild and discuss what sorts of drafts and correspondence should be preserved.

In his presidential address to the 2006 annual meeting of the Society of American Archivists, Richard Pearce-Moses identifies Janus, the Roman god who looks both forward and backward, as “the perfect patron of archivists.”<sup>156</sup> In challenging archivists to consider the future of the digital era, Pearce-Moses provides his own definition of a worst-case future for the profession, one in which records are lost or so disorganized that they cannot be discovered and used. He continues with a vision of how this world would look: “We will have lost our social memory. I believe that society entrusts archivists with preserving the cultural record and our documentary heritage. If we fail to adapt to the digital era, we will necessarily fail that mandate.”<sup>157</sup> But rather than leaving an impression of the futility of the situation, he suggests certain attitudes that are crucial to

---

<sup>155</sup> Neil Beagrie, Robert Beagrie, and Ian Rowlands, “Research Data Preservation and Access: The Views of Researchers,” *Ariadne* issue 60 (2009), <http://www.ariadne.ac.uk/issue60/beagrie-et-al>.

<sup>156</sup> Richard Pearce-Moses, “Janus in Cyberspace: Archives on the Threshold of a Digital Era,” *American Archivist* 70 (Spring/Summer 2007): 13.

<sup>157</sup> *Ibid.*, 16.

enable archivists to become more comfortable in the digital world. He believes that the profession needs early adopters, risk takers, problem solvers, creative thinkers, and those with the initiative and drive to dive into this problematic arena.<sup>158</sup> Where others have cautioned patience until technologies have improved and policies are in place, Pearce-Moses encourages action, stating: “We don’t have to have everything figured out and planned before we start. The essence of strategic thinking is about direction, not about steps. We have a vision of where to go, and we figure out the path as we go along. . . . We cannot let the perfect be the enemy of the possible.”<sup>159</sup> He concludes with an image of archivists as pioneers on the digital frontier, taking risks in order to preserve our documentary heritage.<sup>160</sup>

Given this inspiring image, the question that remains is whether the archival community agrees on this vision of where we need to go. If the archival profession is dedicated to the long-term preservation of born-digital materials, the literature convincingly identifies these practical issues that need to be addressed in a coherent, unified manner:

- determine a best practice for acquiring born-digital materials (e.g., by transfer of physical media or by disk image created by repository staff)
- determine a method for protecting digital objects, including documentation that can be used for authenticity
- determine a method for the appraisal and acquisition of electronic records,

---

<sup>158</sup> Ibid., 19.

<sup>159</sup> Ibid., 21.

<sup>160</sup> Ibid., 22.

including whether these should occur at regular intervals or once the records become inactive

- determine a mechanism for interacting with potential donors of born-digital materials, including written guidelines of preferred formats and suggested file naming conventions
- determine how to handle the interpretation and application of copyright protection to born-digital objects
- write new acquisition and appraisal policies and donor agreements that incorporate issues unique to born-digital materials

Many resources already exist that can help resolve these issues – it is merely a matter of summoning the collective will to make the decisions that will ultimately benefit all constituencies of manuscript repositories. For example, the 2011 “Managing and Sharing Data” report by the UK Data Archive provides a useful one-page data management checklist.<sup>161</sup> The Digital Curation Centre has been collecting and creating resources for a decade, and their web site includes briefing papers, how-to guides, and a data management planning online tool, among other resources.<sup>162</sup> The Consultative Committee for Space Data Systems has produced extensive specifications for the Open Archival Information System (OAIS), and with its acceptance as ISO 14721:2012 (International Organization for Standardization), this document provides a common framework and terminology for archives that are providing for the long-term preservation

---

<sup>161</sup> Veerle Van den Eynden et al., “Managing and Sharing Data,” Colchester, UK Data Archive, 2011, 35, <http://www.data-archive.ac.uk/media/2894/managingsharing.pdf>.

<sup>162</sup> Digital Curation Centre, “Resources for digital curators,” accessed 31 March 2013, <http://www.dcc.ac.uk/resources>.

of digital objects.<sup>163</sup> The Section 108 study group that was convened by the National Digital Information Infrastructure and Preservation (NDIIP) program of the Library of Congress and by the U.S. Copyright Office has provided a useful commentary on the application of copyright to digital objects.<sup>164</sup> OCLC and the Center for Research Libraries developed criteria and a checklist for measuring trustworthy repositories.<sup>165</sup> Finally, later this year a set of Getting Started Guides will be released under the auspices of the Closing the Digital Curation Gap, an IMLS-funded grant project operated out of the School of Information and Library Science at the University of North Carolina at Chapel Hill that has developed online guides for small- and medium-sized repositories that are curating digital objects.<sup>166</sup>

Although it can be tedious and time-consuming work, the policy piece of this problem is actually the easiest to remedy. Just as occurred with the adoption of worldwide description standards, there is much to be gained from an approach that can be embraced by all sizes and types of repositories. The two issues raised in my research that do not yet have reliable solutions are how to provide access to born-digital materials and how to engender administrative support for the work. Perhaps the key to resolving these issues is to recognize that they are intertwined. Cal Lee argues one approach to preserving the layers of meaning held in digital materials is to make the information they

---

<sup>163</sup> Reference Model for an Open Archival Information System (OAIS), June 2012, <http://public.ccsds.org/publications/archive/650x0m2.pdf>.

<sup>164</sup> The Section 108 Study Group Report, March 2008, <http://www.section108.gov/docs/Sec108StudyGroupReport.pdf>.

<sup>165</sup> *Trustworthy Repositories: Audit and Certification*, February 2007, [http://www.crl.edu/sites/default/files/attachments/pages/trac\\_0.pdf](http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf).

<sup>166</sup> Information about the project can be found at <http://digitalcurationexchange.org/cdcg/?q=about>.



possess useful<sup>167</sup>; but this utility is not easily measured when the born-digital materials are either undiscoverable or inaccessible. And in a time when both public and private funding sources are increasingly limited, it is imperative to demonstrate the positive impact that a resource can have on vital constituents in order for that resource – such as a manuscript repository – to be guaranteed the ongoing administrative support necessary for its long-term health.

---

<sup>167</sup> Cal Lee, “A Talk on Digital Preservation,” Information in Life Digital Video Series, University of North Carolina at Chapel Hill, YouTube, 44:41, May 31, 2007, posted by UNCChapelHill, October 7, 2007, [http://www.youtube.com/watch?v=RH9y9CW\\_vMp4](http://www.youtube.com/watch?v=RH9y9CW_vMp4).

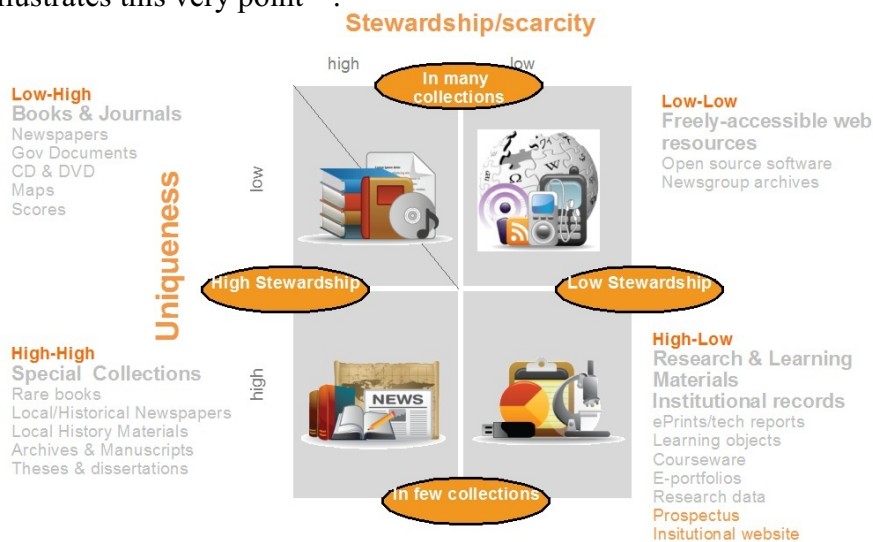
## Conclusion

---

*“Digital information is only as permanent as the hardware and software that gives it intelligibility. A significant percentage of organizational information is born digital, lives, and dies digital without ever being made manifest in the form of paper.”<sup>168</sup>*

---

Manuscript repositories do not play well with others.<sup>169</sup> This is an understandable phenomenon, due to the fact that the guiding purpose of manuscript repositories is to preserve unique papers. Lorcan Dempsey of OCLC Research has developed a collections grid that illustrates this very point<sup>170</sup>:



## COLLECTIONS GRID

(Lorcan Dempsey and Eric Childress, OCLC Research)

---

<sup>168</sup> Stephens, 234.

<sup>169</sup> Actually, this issue is not unique to manuscript repositories. Susan Davis points out that the work that “Camp Pitt” did for government electronic records programs from the late 1980s to the late 1990s did not translate into other types of repositories. Davis, 172.

<sup>170</sup> Lorcan Dempsey, “Reconfiguring Library Boundaries,” (paper presented at RLUK pre-conference, London, 24 November 2011), 17, <http://www.oclc.org/resources/research/presentations/dempsey/rluk2011.pptx>.

Manuscript repositories lie clearly in the high stewardship, high scarcity quadrant of this graphic. The results of the 2012 ARL survey also found that “few of the solutions developed to date have been transferable between institutions.”<sup>171</sup> However, the problems of handling born-digital materials are not unique, and the profession would be well served by finding a space for collaboration to solve these thorny issues. In the words of Don Mennerich, digital archivist for Manuscripts and Archives at the New York Public Library, “local practice is the enemy.”<sup>172</sup>

Writing in 1993, Margaret Hedstrom warns that “archivists should avoid becoming attached to a model or a formula, because the state of the technological evolution and the nascent response by archivists do not yet permit conclusive answers.”<sup>173</sup> But as already admitted, these sorts of acknowledgments of the fluidity and complexity of handling electronic records have only served to sanction a dereliction of duty when it comes to putting systems in place that can perpetuate born-digital materials into the future. Adrian Cunningham cites the imagery of Australian archivist Roger Jones, “who argues that collecting institutions have to grasp the nettle and commence the preservation of records in electronic form.”<sup>174</sup> Although Jones wrote this statement in 1993, his challenge has not yet been embraced twenty years later. As many from Richard Pearce-Moses to Mark Greene<sup>175</sup> have counseled, archivists cannot allow a striving for the

---

<sup>171</sup> Nelson et al., 19.

<sup>172</sup> Interview by author, 14 February 2013.

<sup>173</sup> Hedstrom, “Electronic Records Program Strategies,” 6.

<sup>174</sup> Cunningham, “Archival Management of Personal Records,” 97.

<sup>175</sup> Pearce-Moses, 21; Greene, “The Power of Archives,” 30.

perfect to eclipse the good that can possibly be accomplished in the short term.

Sometimes it has been verbalized while other times it has been implicit in interviews and case studies, but it seems clear that many manuscript repositories are trying to approach the handling of born-digital materials in a manner similar to the way they have digitized analog materials. For a multitude of reasons, this is not an appropriate approach. As Liz Bishoff points out, “digital preservation is an ongoing process rather than an event-driven process.”<sup>176</sup> But more often than not, digitization has been rolled out in a project manner, often through grant funding for “boutique” projects that highlighted a special holding of a repository as a means of calling attention to the collection. Mark Greene suggests that the tendency to approach born-digital collections in the same manner that digitized collections have been handled will lead to paralysis and serves as evidence that archivists as a profession are slow learners, for even in digitization there already should have been a move away from the boutique model.<sup>177</sup> There apparently exists between digitized materials and born-digital materials an unhealthy competition for resource allocation; in the section of the 2012 ARL survey dedicated to access and discovery challenges, a respondent indicated that “we often focus on digitizing collections and providing access to those before we can work with the born-digital content.”<sup>178</sup> Yet there exists one dramatic difference between the common approach to digitization and the

---

<sup>176</sup> Liz Bishoff, “Digital Preservation Assessment: Ready Cultural Heritage Institutions for Digital Preservation” (paper, DigCCurr2007: An International Symposium in Digital Curation at the School of Information and Library Science at the University of North Carolina at Chapel Hill, April 20, 2007), 2, [http://ils.unc.edu/digccurr2007/papers/bishoff\\_paper\\_8-3.pdf](http://ils.unc.edu/digccurr2007/papers/bishoff_paper_8-3.pdf).

<sup>177</sup> Greene, interview by author, 18 February 2013.

<sup>178</sup> Nelson et al., 77.

common approach to born-digital materials: where the primary focus of the former efforts was to provide access to unique materials, even to patrons who might not be able to visit in person, and in so doing to raise the notoriety and esteem of the manuscript repository, the primary focus of handling born-digital materials remains mired in a basic level focus on how to preserve the bits. Until the access piece can be determined, manuscript repositories run the risk of devaluing the content held in electronic records.

Does the lack of tangibility and other sensory inputs make born-digital records harder to handle – or does it at least generate less of a visual trigger that there is pressing work to be accomplished? Does the sheer quantity of electronic records make them seem less in line with the mission of a manuscript repository and more appropriate for oversight by an institutional repository? Are there ways in which manuscript repositories can work together to solve some of the problems of born-digital materials, whether by setting up SWAT sites or sharing policies and workflows that can form the backbone of best practices? And most importantly, will the focus of the archival community about born-digital materials ever shift from preservation to access? Obviously, there is still much research that needs to be accomplished in this arena. Perhaps manuscript repositories should look outside of their usual realm to the work being done with repositories of scientific data or to the types of uses of archival records being designed by digital humanists. Based on the positive feedback that the Jump In initiative has generated by creating a sense of community among those repositories that are trying to begin working with born-digital materials, one simple solution would be to perpetuate this feeling of communal responsibility by establishing mentoring partnerships between

manuscript repositories more comfortable with handling born-digital records and those less practiced. Even if standards are developed, the realm of electronic records is one that will constantly be in flux due to changes in technology, so having a support system in place with other archivists who are facing similar challenges could be a valuable means of preventing the obstacles from appearing insurmountable. Most importantly, if the archival community could embrace a vision of our responsibility to provide to users both analog and born-digital materials, it will be easier to figure out the path that we should take.

## Bibliography

- AIMS Work Group. *AIMS Born-Digital Collections: An Inter-Institutional Model for Stewardship*. January 2012.  
[http://www2.lib.virginia.edu/aims/whitepaper/AIMS\\_final.pdf](http://www2.lib.virginia.edu/aims/whitepaper/AIMS_final.pdf).
- Anderson, Ian G. "Are You Being Served? Historians and the Search for Primary Sources." *Archivaria* 58 (Fall 2004): 81-129.
- Bantin, Philip C. "Strategies for Managing Electronic Records: A New Archival Paradigm? An Affirmation of Our Archival Traditions?" *Archival Issues* 23, no. 1 (1998): 17-34.
- Barry, Rick. "Opinion Piece – Electronic Records: Now and Then." *Records Management Journal* 20, no. 2 (2010): 157-71.
- Beagrie, Neil, Robert Beagrie, and Ian Rowlands. "Research Data Preservation and Access: The Views of Researchers." *Ariadne* issue 60 (2009).  
<http://www.ariadne.ac.uk/issue60/beagrie-et-al>.
- Bishoff, Liz. "Digital Preservation Assessment: Readyng Cultural Heritage Institutions for Digital Preservation." Paper presented at DigCCurr2007: An International Symposium in Digital Curation at the School of Information and Library Science at the University of North Carolina at Chapel Hill, April 20, 2007.  
[http://ils.unc.edu/digccurr2007/papers/bishoff\\_paper\\_8-3.pdf](http://ils.unc.edu/digccurr2007/papers/bishoff_paper_8-3.pdf).
- Blue Ribbon Task Force on Sustainable Digital Preservation and Access. *Sustainable Economics for a Digital Planet: Ensuring Long-term Access to Digital Information*. La Jolla, CA: Blue Ribbon Task Force on Sustainable Digital Preservation and Access, 2010.  
[http://brtf.sdsc.edu/biblio/BRTF\\_Final\\_Report.pdf](http://brtf.sdsc.edu/biblio/BRTF_Final_Report.pdf).
- Carroll, Laura, Erika Farr, Peter Hornsby, and Ben Ranker. "A Comprehensive Approach to Born-Digital Archives." *Archivaria* 72 (Fall 2011): 61-92.
- Cook, Terry. "Byte-ing Off What You Can Chew: Electronic Records Strategies for Small Archival Institutions." *Archifacts* (April 2004): 1-20.
- Cunningham, Adrian. "The Archival Management of Personal Records in Electronic Form: Some Suggestions." *Archives and Manuscripts* 22 (May 1994): 94-105.
- . "Journey to the End of the Night: custody and the dawning of a new era on the archival threshold." *Archives and Manuscripts* 24 (November 1996): 312-21.
- . "Waiting for the Ghost Train: Strategies for Managing Electronic Personal Records Before It Is Too Late." *Archival Issues* 24, no. 1 (1999): 55-64.
- Davis, Susan E. "Electronic Records Planning in 'Collecting' Repositories." *American Archivist* 71 (Spring/Summer 2008): 167-89.
- Dewey, John. "Context and Thought." *University of California Publications in Philosophy* 12, no. 3 (1931): 203-24.
- Dollar, Charles M. *Archival Theory and Information Technologies: The Impact of*

- Information Technologies on Archival Principles and Methods*. Macerata, Italy: University of Macerata, 1992.
- Dooley, Jackie M. and Katherine Luce. "Taking Our Pulse: The OCLC Research Survey of Special Collections and Archives." *OCLC Research*, October 2010.  
<http://www.oclc.org/resources/research/publications/library/2010/2010-11.pdf>.
- Dow, Elizabeth H. *Electronic Records in the Manuscript Repository*. Lanham, MD: Scarecrow Press, 2009.
- Duff, Wendy. "Understanding the Information-Seeking Behaviour of Archival Researchers in a Digital Age: Paths, Processes and Preferences." In *Proceedings of the DLM-Forum 2002*, 331-39. Luxembourg: Office for Official Publications of the European Communities, 2002.  
[http://www.dlmforum.eu/index.php?option=com\\_jotloader&section=files&task=download&cid=171\\_bd67f32c493b7a29e9271897f06d92e1&Itemid=105&lang=en](http://www.dlmforum.eu/index.php?option=com_jotloader&section=files&task=download&cid=171_bd67f32c493b7a29e9271897f06d92e1&Itemid=105&lang=en)
- Duff, Wendy. "Will Metadata Replace Archival Description: A Commentary." *Archivaria* 39 (Spring 1995): 33-38.
- Duff, Wendy, Barbara Craig and Joan Cherry. "Historians' Use of Archival Sources: Promises and Pitfalls of the Digital Age." *Public Historian* 26, no. 2 (Spring 2004): 7-22. doi:10.1525/tph.2004.26.2.7.
- Duff, Wendy M. and Catherine A. Johnson. "Accidentally Found on Purpose: Information-Seeking Behavior of Historians in Archives." *Library Quarterly* 72, no. 4 (Oct. 2002): 472-96.
- Erway, Ricky. *Defining "Born-Digital"*. OCLC Research, November 2010.  
<http://www.oclc.org/resources/research/activities/hiddencollections/borndigital.pdf>.
- . *Swatting the Long Tail of Digital Media: A Call for Collaboration*. Dublin, OH: OCLC Research, 2012.  
<http://www.oclc.org/content/dam/research/publications/library/2012/2012-08.pdf>.
- . *You've Got to Walk Before You Can Run: First Steps for Managing Born-Digital Content Received on Physical Media*. Dublin, OH: OCLC Research, 2012.  
<http://www.oclc.org/content/dam/research/publications/library/2012/2012-06.pdf>.
- Forstrom, Michael. "Managing Electronic Records in Manuscript Collections: A Case Study from the Beinecke Rare Book and Manuscript Library." *American Archivist* 72 (Fall/Winter 2009): 460-77.
- Goldman, Ben. "Bridging the Gap: Taking Practical Steps Toward Managing Born-Digital Collections in Manuscript Repositories." *RBM: Journal of Rare Books, Manuscripts, and Cultural Heritage* 12, no. 1 (2011): 11-24.
- Greene, Mark. A. "MPLP: It's Not Just for Processing Anymore." *American Archivist* 73 (Spring/Summer 2010): 175-203.
- . "The Power of Archives: Archivists' Values and Value in the Postmodern Age." *American Archivist* 72 (Spring/Summer 2009): 17-41.
- . "The Power of Meaning: The Archival Mission in the Postmodern Age." *American Archivist* 65 (Spring/Summer 2002): 42-55.
- Grimard, Jacques. "Managing the Long-term Preservation of Electronic Archives or Preserving the Medium and the Message." *Archivaria* 59 (Spring 2005): 153-67.
- Hedstrom, Margaret. "Descriptive Practices for Electronic Records: Deciding What is



- Essential and Imagining What is Possible.” *Archivaria* 36 (Fall 1993): 53-63.
- , ed. *Electronic Records Management Program Strategies*. Pittsburgh: Archives & Museum Informatics, 1993.
- Henry, Linda J. “Schellenberg in Cyberspace.” *American Archivist* 61 (Fall 1998): 309-27.
- Hurford, Amanda A. and Carolyn F. Runyon. “New Workflows for Born-Digital Assets: Managing Charles E. Bracker’s Orchid Photographs Collection.” *Computers in Libraries* 31, no. 1 (2011): 6-10, 40.
- Hyry, Tom and Rachel Onuf. “The Personality of Electronic Records: The Impact of New Information Technology on Personal Papers.” *Archival Issues* 22, no. 1 (1997): 37-44.
- Jimerson, Randall C., ed. *American Archival Studies: Readings in Theory and Practice*. Chicago: Society of American Archivists, 2000.
- John, Jeremy Leighton. “Adapting Existing Technologies for Digitally Archiving Personal Lives: Digital Forensics, Ancestral Computing, and Evolutionary Perspectives and Tools.” iPRES 2008: The Fifth International Conference on Preservation of Digital Objects (29-30 September 2008).  
[http://www.bl.uk/ipres2008/presentations\\_day1/09\\_John.pdf](http://www.bl.uk/ipres2008/presentations_day1/09_John.pdf).
- John, Jeremy Leighton, Ian Rowlands, Peter Williams, and Katrina Dean. “Digital Lives: Personal Digital Archives for the 21st Century >> an Initial Synthesis.” 2010.  
<http://britishlibrary.typepad.co.uk/files/digital-lives-synthesis02-1.pdf>.
- Ketelaar, Eric. “Archives in the Digital Age: New Uses for an Old Science.” *Archives & Social Studies* 1, no. 0 (March 2007): 167-91.  
[http://archivo.cartagena.es/files/36-168-DOC\\_FICHERO1/10-ketelaar\\_archives.pdf](http://archivo.cartagena.es/files/36-168-DOC_FICHERO1/10-ketelaar_archives.pdf).
- Kirschenbaum, Matthew G., Richard Ovenden, and Gabriela Redwine. *Digital Forensics and Born-Digital Content in Cultural Heritage Collections*. Washington, DC: Council on Library and Information Resources, 2010.  
<http://clir.org/pubs/reports/pub149/pub149.pdf>.
- Lavoie, Brian F. *The Incentives to Preserve Digital Materials: Roles, Scenarios, and Economic Decision-Making*. Dublin, Ohio: OCLC Research, 2003.  
<http://www.oclc.org/research/projects/digipres/incentives-dp.pdf>.
- Lee, Cal. “A Talk on Digital Preservation.” Information in Life Digital Video Series, University of North Carolina at Chapel Hill. YouTube, 44:41, May 31, 2007. Uploaded by UNCCChapelHill, October 7, 2007.  
[http://www.youtube.com/watch?v=RH9y9CW\\_vMp4](http://www.youtube.com/watch?v=RH9y9CW_vMp4).
- Lee, Christopher A., Matthew Kirschenbaum, Alexandra Chassanoff, Porter Olsen, and Kam Woods. “BitCurator: Tools and Techniques for Digital Forensics in Collecting Institutions.” *D-Lib Magazine* 18 (May/June 2012).  
doi:10.1045/may2012-lee. <http://www.dlib.org/dlib/may12/lee/05lee.html>.
- Lee, Christopher A. “Guerilla Electronic Records Management: Records Management Lessons Learned.” *Records and Information Management Report* 18, no. 5 (May 2002): 1-13.
- , ed., *I, Digital: Personal Collections in the Digital Era*. Chicago: Society of American Archivists, 2011.

- Losee, Robert M., Jr., and Karen A. Worley. *Research and Evaluation for Information Professionals*. San Diego: Academic Press, Inc., 1993.
- MacNeil, Heather. "Metadata Strategies and Archival Description: Comparing Apples to Oranges." *Archivaria* 39 (Spring 1995): 22-32.
- Mellor, Phil, Paul Wheatley, and Derek Sergeant. "Migration on Request, a Practical Technique for Preservation." *The Sixth European Conference on Research and Advanced Technology for Digital Libraries*, 2002.
- Millar, Laura A. *Archives: Principles and Practices*. New York: Neal-Schuman Publishers, 2010.
- Nelson, Naomi L., Seth Shaw, Nancy Deromedi, Michael Shallcross, Cynthia Ghering, Lisa Schmidt, Michelle Belden, Jackie R. Esposito, Ben Goldman, and Tim Pyatt. *Managing Born-Digital Special Collections and Archival Materials*. SPEC Kit #329. Washington, DC: Association of Research Libraries, 2012.
- Pearce-Moses, Richard. "Janus in Cyberspace: Archives on the Threshold of a Digital Era." *American Archivist* 70 (Spring/Summer 2007): 13-22.
- Peters, Catherine Stollar. "When Not All Papers Are Paper: A Case Study in Digital Archivy." *Provenance* 24 (2006): 22-34.
- Prom, Christopher J. and Thomas J. Frusciano, eds. *Archival Arrangement and Description*. Chicago: Society of American Archivists, 2013.
- Roe, Kathleen D. *Arranging & Describing Archives & Manuscripts*. Chicago: Society of American Archivists, 2005.
- Rothenberg, Jeff. "Ensuring the Longevity of Digital Documents." *Scientific American* 272 (January 1995): 42-47.
- Rothenberg, Jeff. *Ensuring the Longevity of Digital Information*. Washington, DC: Council on Library and Information Resources, 1999.  
<http://www.clir.org/pubs/archives/ensuring.pdf>.
- Salo, Dorothea. "Innkeeper at the Roach Motel." *Library Trends* 57 (Fall 2008): 98-123.  
<http://digital.library.wisc.edu/1793/22088>.
- Snow, Kellie, Bart Ballaux, Birte Christensen-Dalsgaard, Hans Hofman, Jens Hofman Hansen, Perla Innocenti, michael Poltorak Nielsen, Seamus Ross, and Jørn Thøgersen. "Considering the User Perspective: Research into Usage and Communication of Digital Information." *D-Lib Magazine* 14, no. 5/6 (May/June 2008). doi:10.1045/may2008-ross. <http://dlib.org/dlib/may08/ross/05ross.html>.
- Stephens, David O. *Records Management: Making the Transition from Paper to Electronic*. Overland Park, KS: ARMA International, 2007.
- Thibodeau, Kenneth. "Preserving Digital Memory at the National Archives and Records Administration of the U.S." Paper presented at the Workshop on Conservation of Digital Memories, Second National Conference on Archives, Bologna, November 20, 2009.  
<http://www.mybestdocs.com/thibodeau-k-preser-dig-memNARA091120.pdf>.
- Thomas, Jeff. "What Do I Do With The Black Box? Processing the Electronic Records in the Deborah Pryce Papers." *Congressional Papers Roundtable Newsletter*. Forthcoming.
- Tibbo, Helen R. "Primarily History in America: How U.S. Historians Search for Primary Materials at the Dawn of the Digital Age." *American Archivist* 66, no. 1 (Spring -

Summer, 2003): 9-50.

Van den Eynden, Veerle, Louise Corti, Matthew Woollard, Libby Bishop and Laurence Horton. "Managing and Sharing Data." Colchester, UK Data Archive, 2011.

<http://www.data-archive.ac.uk/media/2894/managingsharing.pdf>.

Wallace, David A. "Managing the Present: Metadata as Archival Description."

*Archivaria* 39 (Spring 1995): 11-21.

## Appendix A

### 3 February 2013 message to the SAA Manuscript Repositories section list:

I am a student at the School of Information and Library Science at the University of North Carolina at Chapel Hill, pursuing a Master of Science in Library Science degree and an Archives and Records Management concentration. I am writing my master's paper about the efforts of manuscript repositories to begin accessioning and processing born-digital records, and I would like to request your input.

If your repository has not yet begun to process born-digital records, you are invited to answer two questions about what factors have limited the ability of your repository to process born-digital records and what factors would facilitate your ability to do so. Participating in this survey should take only about five minutes and will significantly contribute to the writing of my master's paper. To participate in this survey, please follow this link: [https://qtrial.qualtrics.com/SE/?SID=SV\\_6yAhPV24ZlbFAH3](https://qtrial.qualtrics.com/SE/?SID=SV_6yAhPV24ZlbFAH3). The survey will be available for one month. I thank you for your participation.

If your repository already processes born-digital records, I would like to conduct a short interview with you, at your convenience, to discuss your workflow and the lessons you've learned. We could conduct this interview via email or via phone, whichever you prefer. Please respond to this email to indicate your willingness to participate along with your preference for method of contact, and I will be in touch shortly to confirm.

Thank you for your support in this research endeavor.

Sincerely,

Courtney Bailey  
School of Information and Library Science  
University of North Carolina at Chapel Hill  
Advisor: Jackie Dean

## Appendix B

### Qualtrics survey of repositories not currently processing born-digital materials:

What factors have limited the ability of your repository to process born-digital records?

	Not a factor	A small consideration	A significant consideration	A monumental obstacle
Training to know how to handle born-digital records	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Cost of handling born-digital records	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Concerns about providing patrons access to born-digital records	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Time required to process born-digital records alongside backlog of paper records	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Inadequate administrative support	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

What would facilitate the processing of born-digital records by your manuscript repository?

	Not an aid	A small aid	A significant aid	An extremely important solution
Being able to bring in a consultant to provide assistance in establishing a workflow for processing born-digital records	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Being able to study established workflows from other similarly sized repositories to learn from their processes	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Being able to acquire grant funding to help cover costs	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Being able to provide training for reference staff in how to facilitate patron access to born-digital records	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Being able to provide training for patrons in how to access born-digital records	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Being able to hire an electronic records archivist who could focus only on the processing and access of born-digital records	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Being able to develop an institutional policy governing the acquisition, preservation, and access of born-digital materials	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Please supply information about the repository described by these responses.

Name of repository:

Address of repository:

Type of repository:

- ☐ College/university special collections/manuscripts
- ☐ Independent research library
- ☐ Museum
- ☐ Historical society
- ☐ National institution
- ☐ Governmental institution
- ☐ Public library
- ☐ Consortium
- ☐ Other (please elaborate)

## Appendix C

### interview questions posed to archivists at repositories already processing born-digital materials:

1. How long has your repository been processing born-digital records and what prompted you to begin?
2. What workflow have you established for the processing of born-digital records and what prompted the creation of this workflow?
3. What training do you provide your staff in the processing of born-digital records?
4. How have you found the costs of processing born-digital records compare to the costs of processing paper records?
5. How have you found the time for processing born-digital records compares to the time of processing paper records and have you discussed how the principles of original order and MPLP apply to born-digital records?
6. What training do you provide your staff about providing access to born-digital records?
7. What training do you provide to patrons about accessing born-digital records and where can these records be accessed?
8. Has your processing of born-digital records caused you to amend any of your appraisal and acquisition policies?
9. Are there collaborations within or without your institution that have made this work more successful?
10. What, if any, improvements do you wish you could make to how your repository handles born-digital records?

## Appendix D

### 26 February 2013 email sent to participants in the Jump In Initiative:

First of all, I want to congratulate you on your decision to participate in the Jump In Initiative that is being sponsored by the SAA Manuscript Repositories Section.

Secondly, I believe Chris Burns mentioned that I would be contacting you as a part of my research for a master's paper. I am a student at the School of Information and Library Science at the University of North Carolina at Chapel Hill, pursuing a Master of Science in Library Science degree and an Archives and Records Management concentration. I am writing my master's paper about the efforts of manuscript repositories to begin accessioning and processing born-digital records, and I would like to request your input. I have a short list of questions, and I would greatly appreciate your responses. If you would prefer to talk with me via phone instead of responding to this email, please let me know your phone number and a time that it would be convenient to call you, and I will gladly follow up that way.

1. What prompted your repository to participate in this Jump In Initiative?
2. While I realize your survey may not yet be complete, I'm curious whether your preliminary inventory of your collection has been surprising in any way (e.g., a greater/lesser quantity of computer media than presumed; more/fewer file formats than presumed)?
3. What order of procedure has your repository established – are you first trying to establish a workflow for handling born-digital materials and then will address related policy issues, or have you already established policies for the handling of born-digital materials and are now addressing the workflow issues? If you already have policies (e.g., regarding acquisition of or access to born-digital materials) and can point me to those on the Web or send me a copy, I would appreciate being able to review them.

Thank you for your support in this research endeavor, and good luck with your survey and essay.

Sincerely,

Courtney Bailey  
School of Information and Library Science  
University of North Carolina at Chapel Hill  
Advisor: Jackie Dean