Literature Review:
Digital Preservation and Archives

Kelly Guris
INFO 560: Intro to Archives I
April 27, 2010
Digital preservation, especially of “born-digital” content, is a hot button issue in the archival community. The quickly expanding digital content created daily, lack of a standardized approach to digital preservation, and limited economic resources available to collecting institutions are several of the pressing concerns to the archival community, and, truly, society at large. Valuable digital records are in peril of being lost forever – clearly an alarming problem with far reaching consequences to more than just archivists: individuals, businesses, and governments are all affected. This paper analyzes three scholarly articles in an attempt to explore the concerns and implications surrounding digital archives. All of the articles were written this decade, spanning from 2003 to 2007 and still mostly relevant to the current debate. Each article focuses on a different aspect of digital preservation giving a comprehensive view of the topic.

“Defining digital sustainability,” written by Kevin Bradley of the National Australian Library, introduces the technology-centered concept of digital sustainability in relation to digital preservation. Bradley views digital preservation as an “integral part” of digital sustainability, an approach that examines the “overall lifecycle, technical, and socio-technical issues” presented by digital content, its creation, and management (p. 151). The search for technologies to secure and preserve data is only part of the suggested approach to digital preservation. Rather than defining digital “sustainability,” Bradley seems to be redefining what digital “preservation” encompasses: it’s not just a “simple preservation model” (p.148) but “building an economically viable infrastructure, both social and technical, for maintaining valuable data without significant loss or degradation” (p. 157).
Basically, we cannot just save everything in databases without providing the ability to retrieve and meaningfully use the data in the future. Digital content with its complex copyright contracts and licensing agreements, its ease of duplication and dissemination, and its ever-changing, advancing, and impermanent state make this issue too messy to wrap up with a simple digital repository.

Bradley sufficiently presents a brief technical background on digital preservation including the various past solutions proposed such as migration and emulation (p. 153), before outlining a true sustainable approach. He essentially gives meaning to the term digital sustainability through enumerating its critical functions. These functions include maintaining a “source of reliable funding,” selecting collections that will be valuable to future users, “drawing on highly skilled or informed experts” for risk assessment, and providing “stable technical support” (Bradley 2007, pp. 159-160).

In addition, Bradley briefly introduces two terms that “support the sustainability approach: curation and stewardship” (p. 161). Curation involves maintaining access to constantly changing digital content, and stewardship means, “caring for,” “honoring,” and “preserving” information and our “cultural heritage for the benefit of future generations” (Bradley 2007, p. 161). Although indicating earlier in his paper that digital preservation is a part of digital sustainability, Bradley later applies a Venn diagram to the approach using sustainability, stewardship, and curation as components comprising digital preservation (p.161). Despite this small inconsistency, Bradley thoroughly examines the issues a
successful digital preservation approach must address to be considered such from a technical perspective.

In Lavoie and Dempsey’s “Thirteen ways to look at... digital preservation,” the authors provide a broader perspective. Instead of “focusing on the technical obstacles,” Lavoie and Dempsey present the “view that digital preservation is [...] one component of a broad interconnected services, policies, and stakeholders which together constitute a digital information environment” (Introduction, n.p.). Lightheartedly using Wallace Steven’s poem “Thirteen Ways to Look at a Blackbird” as a template, Lavoie and Dempsey list thirteen maxims of what digital preservation is, and they are: “an ongoing activity,” “a set of agreed outcomes,” “an understood responsibility,” “a selection process,” “an economically sustainable activity,” “a cooperative effort,” “an innocuous activity,” “an aggregated or disaggregated service,” “a complement to other library systems,” “a well-understood process,” “an arm’s length transaction,” “one of many options” of preservation activities, and “a public good.” Most points are self-explanatory and provide the reader with a good overview of the terms and issues associated with digital preservation and archives. If anything, Lavoie and Dempsey’s thirteen perspectives are in fact thirteen objectives of successful digital preservation. The concluding remarks give a brief informative synopsis of the main facets of digital preservation – it’s actually more of a great introduction to this topic, touching on the technical, social, cultural, economic, and legal processes involved in carrying out digital preservation projects (Lavoie & Dempsey 2004, n.p.).
In “Scarcity or abundance? Preserving the past in the digital era,” Roy Rosenzweig of George Mason University’s History department gives the historian’s perspective on digital preservation and archives. He begins the discussion with an interesting story about a website called *Bert Is Evil*, a good example of potential downfalls from unresolved digital preservation conflicts. “Although some historians might object that the *Bert Is Evil* web site is of little historical significance, even traditional historians should worry about what the digital era might mean for the historical record (Rosenzweig 2003, p. 736). Rosenzweig presents two historical concerns resulting from digital preservation and archives: the potential scarcity a digital historic record, due in part to the nature of its form, and the ever-increasing and overwhelming abundance of information generated as a result of digital resources and technologies. Surprisingly, historians, according Rosenzweig, “have almost entirely ignored” the idea of “thinking simultaneously about how to research, write, and teach in a world of unheard-of historical abundance and how to avoid a future of record scarcity” (p. 738). According to the author, historians see the responsibility of handling these issues as falling largely on archivists. Additionally, this “detachment” can be attributed to assumptions that the topic is strictly a “technical” one (Rosenzweig 2003, p. 738). “Yet the more important and difficult issues about digital preservation are social, cultural, economic, political, and legal – issues that humanists should excel at” (Rosenzweig 2003, p. 738). Most debate centers on who is responsible and what should be saved. Of interesting note is the author’s question “will abundance bring better or more thoughtful history,” a
concept I otherwise would not have considered. Unfortunately, Rosenzweig does not expand on that topic much further.

He, does, however, provide an eloquent explanation of various technical issues and gives insight into copyright and intellectual property concerns. Of these latter concerns, Rosenzweig asks a question fundamental to the intellectual property debate: “if libraries don’t own digital content, how can they preserve it?” (p. 744). Subscription fees and licensing agreements do not equate to legal ownership. Without this ownership, preservationists’ hands are tied from manipulating records for archival purposes without consent from the copyright owner. The real friction here lies in dealing with publishers, whose main objective of profit can be a detriment to preservation efforts. Further discussion covers government records, which are increasingly born-digital material. Concerning proposed resolutions to the digital preservation issues discussed, Rosenzweig’s best advice is to avoid “all-or-nothing, magic-bullet approaches” (p. 747).

Despite emphasis that digital preservation is not just a technical issue, most of the promising projects the author highlights are based on technical challenges to digital preservation. We need more discussion and resolution of broken copyright law. Or how about examples outside of the U.S. of working solutions and processes? Rosenzweig does touch on government-run digital preservation projects in Europe that seem to show positive results. Further analysis of these initiatives would be beneficial. Overall, Rosenzweig points out deficiencies but fails to offer much in way of a solution other than to rightly urge historians to participate in “making decisions about priorities in digital preservation” (p. 759).
The literature discussing digital preservation is abundant in some scholarly and professional arenas, like computer science and database management, and lacking in others, like history. Much of the discussion concerning technical solutions and approaches can be dense and difficult to translate into laymen’s terms. Bradley’s explanation of specific technical issues associated with a digital preservation project lacked accessibility to those outside the scope of computer science and Internet technology. I theorize that bridging the gap between the technical matters and scholarly, legal, and economic issues resulting from digital archives will encourage further discussion and a larger awareness of the societal impact of digital preservation. Rosenzweig makes a good attempt at stimulating interest in historians and other social scientists unaware of their role in affecting positive results and what they have at stake to lose. Ultimately, Lavoie and Dempsey offer the most accessible and comprehensive view of digital preservation, and thus, a great jumping off point for further and more specific research on the challenges and benefits of sustainable digital archives.

References

