Waking the [Digital] Dead: A Continuum Approach to Digital Initiatives

Melanie Griffin, Kyna Herzinger, and Patricia Sasser

Introduction
Over the last two decades, cultural heritage institutions have elevated online access to collections from the status of an atypical luxury to that of a professional imperative. Even small libraries, archives, and museums see digital projects as a necessary component of their work while larger institutions regularly develop not only new initiatives but original methods of access and interaction, such as interactive mobile applications like Murder at the Met or participatory, user-driven digital archives like the Occupy Archive and the Grateful Dead Archive Online. Most institutions, whether large or small, are engaged in ongoing collaborative projects and it would not be an exaggeration to say that discrete projects debut nearly weekly.

From the beginning, digital initiatives have required not only a diverse set of intellectual and technical skills, but they have also involved a large number of stakeholders. These include not only the full-time staff who manage these initiatives but also, depending on the circumstances, colleagues from other university departments or institutions, donors, vendors, student workers, and volunteers. In addition, the allocation of institutional resources, both financial and otherwise, constitutes another kind of stake in these digital projects that supersedes any one person or group of people and their work.

The imperative to produce digital projects and the significant investment involved has been well documented by the professional community with most attention paid to the initial, or creative, stages of these projects. This has been both appropriate and useful, highlighting the real need for practical and theoretical discussions about where to begin. Increasingly, the professional community has also considered use and user-behavior, charting the interactions between digital projects and their constituencies, working out where these initiatives fit into the digital landscape as a whole and into the life of the library or archive in particular. Perhaps most challenging, however, has been the inevitable “end” of these digital projects. Like all technologies, digital projects have an inherent obsolescence; innovation becomes commonplace, trends change, priorities shift. Many projects, once proudly and prominently featured on institutional homepages, slip gradually into the background, experiencing a quiet but literal atrophy as links die and files decay. But what are institutions supposed to do with them? Are these the digital equivalent of brittle books, not worth the server space they occupy? Or are they part of the cultural record, deserving preservation in every detail, including even the hardware and software that generated the images? Must they be maintained in perpetuity? Or should they be archived in some way, making space for the next interesting endeavor?

As early as 2002, the professional community recognized that the dilemma of digital preservation must be addressed. A CLIR report that surveyed digital
cultural heritage institutions and their sustainability concerns noted that “the process of phasing out these [projects] was thought to be critical, albeit difficult.” While the CLIR report did not offer recommendations for “phasing out” projects in any detail, it is not hard to see the critical nature and the intrinsic difficulty in such a recommendation. On one hand, there is a very practical challenge: the number and proliferation of these digital initiatives means that more and more projects are moving towards or reaching obsolescence. The diversity of the projects and the institutions resists a simple “one-size-fits-all” solution for preservation. On the other hand, there is also a philosophical challenge: the high investment of the many stakeholders in these projects means that it is sometimes psychologically difficult to let go of projects and move on to the next thing. The behavioral scientist Hal Arkes has called this difficulty the “sunk-cost effect”: humans exhibit a “tendency to continue in an endeavor once money, effort, or time has been made” regardless of the real quantifiable benefits of such an endeavor. “Abandoning a project you’ve invested a lot in,” Arkes said more recently, “feels like you’ve wasted everything, and waste is something we’re told to avoid.” Given the precarious and shrinking budgets of many institutions, the cultural heritage community is all too cognizant about avoiding waste. While preserving a digital project is not tantamount to abandoning it, moving projects into the category of “preservation” may appear to indicate something about the current quality or the value of the project itself.

With these many factors influencing digital preservation, it is perhaps no surprise that Humanities Advanced Technology and Information Institute founding director Seamus Ross concluded that few of the organizations that are engaged in digital preservation view their activities with a long-term lens, lacking, in some instances, even a uniform understanding of digital preservation issues. Perhaps this reason, more than any other, is why institutions often let initiatives linger in limbo. A strict cost-benefit analysis might show that a project ought to be (in the words of the CLIR report), “phased out.” But, of course, cultural heritage activity is not solely predicated on cost-benefit analysis. What it is predicated on, among other things, is long-term thinking. Digital projects are made with specific kinds of use and users in mind, but like all other materials, they may be used in the future in new ways and by very different groups of users.

Stated another way, cultural heritage institutions endeavor to pass the world’s informational heritage onto the next generation in a form that is functional, but they do so without address the looming question of digital longevity. Ross responded to this dilemma by recommending that digital curators ply the depths of archival theory in order to identify solutions for digital asset management. This paper is a response to Ross’s call. Drawing from archival theory’s notion of the records continuum, this paper recasts digital preservation as a long-term professional responsibility and digital initiatives as sustainable assets rather than one-time projects destined to die. By laying a conceptual foundation on which to build a discussion of digital preservation, this paper argues that digital curators must continue to develop, maintain, and evaluate digital initiatives after projects are “finished.” Rather than dreading obsolescence, institutions should expect and anticipate the progression of digital projects. Drawing from frank analysis of five case studies, this paper further offers practical applications to “wake” digital projects gone cold.

Continuum Theory
Initiated by the National Archives of Australia and articulated by archivist Ian Maclean, the records continuum, in its most basic form, envisions recorded information as a continuous whole. In contrast, the lifecycle model, which is embraced in the United States, views information in distinct stages—ranging from record “birth” to “death”—with only a select few records that are “resurrected” as archives. The life-cycle view creates a sharp distinction between current and historic recordkeeping; continuum theory, on the other hand, views records as both current and historic from the moment they are created, a relatively radical assumption that necessitates an archivist’s involvement before a record is “born.” Continuum theory expands the boundaries of an archivist’s sphere of responsibility by anticipating record use before the event that occasions its recording, and in the process, continuum theory dismantles the conceptual barriers between records management and archival administration, providing records managers and archivists with a mechanism to articulate a unified mission. More importantly, however, continuum theory attempts to solve the age-old archival question, “What do we save?,” and is a theoretical model of archival appraisal.
As an intellectual and conceptual activity, appraisal—the act of selecting materials of enduring value for long-term retention—is the moment in which an event or activity is memorialized and indefinitely added to the historical record. The Society of American Archivists defines appraisal more pragmatically as “the process of determining the length of time records should be retained, based on [...] their current and potential usefulness.” Though this definition is somewhat limited, it does begin to offer a rubric by which to view records, even those records that are simply digital surrogates created by archivists for access. As such, continuum theory challenges digital curators to think about digital objects from the moment of their creation.  

Continuum theory, to be sure, has been debated in the context of archives management since it was first introduced. Opponents, for instance, have argued that professional dichotomies between record managers and archivists are more “apparent than real.” Others have expressed concern that archivists are placing too much faith in themselves to identify what should be saved. And those outside of the U.S., would argue that record creators—not archivists—should be the ones to make appraisal decisions.  

Regardless of one’s attitude toward this theory, the records continuum provides a conceptual foundation on which to build a discussion of digital preservation. In particular, this model asserts that the appraisal, capture, and preservation of digital objects occur at the moment they are created—an action that is incredibly proactive. It argues that saving, managing, preserving, representing, and re-presenting digital heritage can and should be viewed as a continuum of seamless and ongoing responsibilities. As created records, digital surrogates and any corresponding metadata should be appraised at the time of their creation, keeping future use in consideration.  

Continuum theory also emphasizes integrated processes that are firmly rooted in an institution’s mandate and workflows. In much the same way, digital initiatives must support an institution’s mission and therefore demand a role in that institution’s strategic planning. In order to sustain digital initiatives over time, responsible institutions must secure staff time, monetary resources, and administrative support in order to perform the work required. Rose Sherman of the Minnesota Historical Society poignantly noted that libraries, archives, and museums that have no plan for their digital initiatives are left directionless. “[You] won’t know where you’re headed, how much it will cost you to get there, and what it will cost to maintain your technology environment.” Given the cost and complexity of digital initiatives, as well as the power and potential to exploit the rich sets of relationships across collections, institutions, and social boundaries, it is imperative that digital initiatives become not simply an appendage or afterthought, but that they become fully integrated into an institution’s essential functions.  

Finally, continuum theory is “postcustodial,” focusing not on preserving manuscripts, books, ephemera, photos, and the other “stuff” of special collections, but on bits and bytes that have no discernable physical substance. Nearly twenty years ago, archivist Terry Cook argued that “we must stop being custodians of things” and start focusing on the concepts and contexts that accompany digital records. Indeed, because digital data are “fragile,” and do not survive like their analog counterparts, digital objects simply cannot be left to benign neglect. Digital preservation must be a priority and must capture the processes of digital object creation. Successful digital curation, then, involves a range of activities that lies outside of merely technical, institutional, and juridical boundaries.  

Where, then, does this leave us, the curators of digital objects? Many repositories now have over a decade’s worth of digital objects and their accompanying projects, a sizable number of which were created without serious thought to long-term sustainability or, indeed, to the desirability of long-term sustainability. The following case studies attempt to understand digital projects not merely as “special initiatives,” as mechanisms for gaining public support, or even as an end to ensure preservation, but as a key element in the special collections and archives continuum. As such, digital projects deserve attention not only in annual budgets and daily staff activities, but also in the more theoretical aspects of collections management and maintenance that our centuries’ worth of paper materials routinely receive. We offer these case studies to facilitate a discussion of the ways in which continuum theory can be concretely applied to digital initiatives. While inspired by the authors’ experiences with digital projects at a range of academic institutions, identifying information has intentionally been removed in order to focus on practical applications rather than pinpointing past failures and successes.
Case Study #1: Format-Driven Migration

Summary: In 2006, a large research university began a collaborative project between the university library and a faculty member from an academic department on campus to create an image-based collection of digital surrogates. The images were scanned at 1200 ppi on a scanner that, at the time that it was purchased, was state of the art; it was aging in 2006 and, by 2013 standards, is the functional equivalent of a dinosaur. These surrogates, while created by the library and housed in the library’s content management system, are accessed via an external website hosted and created by the partner academic department.

Over the last six years, this website and the metadata associated with the digital objects have undergone a series of radical changes. First, the initiative was successfully targeted at other discipline-specific researchers, with the hopes that the website would be the first step in creating a multi-institutional digital corpus of materials for advanced study. Later, when attempts at raising grant funding to continue the creation of the digital corpus were unsuccessful and the initial flurry of excitement (and accompanying collection uses), the project was reconceived as a teaching tool, offering digital materials for both undergraduate and high school classrooms that would otherwise be unavailable. More recently, site and download analytics indicate that the project is receiving primarily internal use by the academic department which conceived of the project in 2006. It is, in essence, saving researchers in the home department a trip to the library.

Challenges: While the project was originally conceived with much excitement, there was not a great deal of attention paid to what would happen if enthusiasm for the collection waned and one day the initiative proved more costly than beneficial. As with many early attempts at large-scale digital humanities initiatives, there was an implicit belief that the project would never die, that the surrogates would remain safe on the library’s servers forever, and that the initiative would be greater than the sum of its individual objects. As a result, when the project’s first iteration began to die a natural death, both the parent department and the libraries engaged in a series of reincarnation attempts.

Recommendations: Continuum theory provides a way out of this never-ending loop with its emphasis on the continuous need to appraise records throughout their entire life cycles. The digital surrogates here, once the bedrock of a successful initiative, have come to resemble a faculty research or clippings collection. In the paper-based world of archives and special collections repositories, clippings collections are subject to extensive weeding or even deaccessioning. Why should the same principals not apply to the digital instance of a collection with limited value? The physical collection from which the surrogates were derived is in excellent condition, and the digital surrogates were never intended to serve as long-term stand-ins for the originals. Instead, they were intended to accomplish specific tasks that the physical collection could not. With these tasks accomplished or forgotten and the digital dust of benign neglect gathering, the collection requires a return to appraisal and, perhaps, the acknowledgement that declaring a project completed is not to declare the entire effort a waste of time or money. Hindsight is, as we all know, perfect, but in this instance it provides excellent fodder for considering not only when to cut one’s losses but also for planning new digital initiatives and evaluating others that are floundering.

Case Study #2: Digital Collection Building

Summary: Another large research library, located at a private doctoral-granting university, has been engaged in systematic digital collection building for the past eight years. While it began as a subunit of special collections, Digital Collections is now a standalone department in the library, with three fulltime staff members and five part-time student workers. Most of this library’s digital collection building activities have been driven by institutional priorities, not those of a single faculty member in an academic department. Collections include format-driven migration projects, ranging from audio/visual materials for which the physical instance no longer exists in a stable, usable form, to images that were created for convenience and to encourage wider use of somewhat hidden physical collections; other digital collections are topical and are built around areas of distinction, including digital surrogates as well as born-digital materials. Like many institutions, this digital collection building initiative initially focused on what the library identified as the most valuable materials in its collection, followed by the most heavily used. Since these digital surrogates were created at the beginning, they were scanned according to then-current best practices.
Challenges: Some of the collections were conceived in a continuum theory model, giving thought to appraisal and preservation along with capture at the moment that the digital surrogates were created. Others, like the collection discussed in the first case study, focused exclusively on capture, assuming that appraisal and preservation would occur later. Like many institutions, the library has been playing catch-up in the field of digital preservation, but has for a number of years had a sound, behind-the-scenes digital preservation program in place. All of the library’s digital files could then be saved in perpetuity if the library wished to do so. Born-digital materials, however, are proving particularly problematic. Currently, the library has receives these materials well after the point of creation, with the department’s first point of contact occurring at the point of deposit, long after digital decay has begun to occur.

Recommendations: Just because the library can save all of its digital objects and accompanying digital projects, initiatives, interfaces, and platforms, does it really want to? As we have already seen in the previous case study, the user interface created for a body of digital surrogates can change radically due to shifting priorities. Linked open data provides an opportunity to integrate sound collection management practices with a more fluid approach to interface design and adoption. Instead of continuously re-creating, re-loading, or otherwise proliferating copies of digital resources, linked open data calls for creating links to the master resource (or a presentation derivative thereof). By linking digital objects instead of endlessly re-creating them, digital projects are able to make use of existing resources, including both preservation infrastructures as well as descriptive information, while continuously avoiding the temptation to consider a collection “finished.”

An application of continuum theory principals would encourage digital collections staff to consider proactively the proposed longevity of their digital objects. In traditional, format-driven migration projects, this might be as simple as considering at the outset what the hoped for goals of the project are and acknowledging that the collection will require additional assessment as time passes. It might also, however, include drafting a digital records retention and disposition schedule for the collection.

Collections comprised of born-digital materials, however, are somewhat trickier to fold into the continuum theory model given current library practices. The library’s digital collections include a sizable component of born-digital materials, ranging from recent documents in the University Archives to hybrid collections of paper and digital personal papers. Like many institutions, the first wave of the library’s born-digital materials came in long past the point of initial creation, resulting in a wide range of digital preservation and presentation challenges. In order to proactively encourage different practice that will result in more stable and useful born digital materials, the digital collections staff could work directly with departments on campus that deposit digital papers into the University Archives, providing them with basic training in best practices for document creation, encouraging departments to think about using software and file types with a wide user base (.doc or .rtf files, for example, instead of .pages). Staff could also work proactively with authors, politicians, and artists who have deposit agreements with the library and who have, over the last two decades, shifted from creating paper files to digital ones. Integrating consideration for the digital life cycle of the documents at the point of their creation offers the potential of radically reducing preservation challenges down the road, as well as encouraging stakeholders to consider factors such as versioning, retaining versions of important documents, and depositing those versions.

Case Study #3: Consortial Projects

Summary: While there are inevitable challenges associated with a single institution changing its workflows and data curation practices, these challenges are magnified when they occur in a consortial environment due to the increased number of stakeholders and varying degrees of technological fluency among those stakeholders. Consider this consortial situation: A statewide consortium, comprised of the libraries located at all the state universities and funded through the Library Services and Technology Act (LSTA), has created two parallel digital initiatives: one provides access to finding aids, the other provides access to digital files and file types with a wide user base (.doc or .rtf files, for example, instead of .pages). Staff could also work proactively with authors, politicians, and artists who have deposit agreements with the library and who have, over the last two decades, shifted from creating paper files to digital ones. Integrating consideration for the digital life cycle of the documents at the point of their creation offers the potential of radically reducing preservation challenges down the road, as well as encouraging stakeholders to consider factors such as versioning, retaining versions of important documents, and depositing those versions.
vested, possibly because they have been working with smaller staff with less specialized technical expertise.

Challenges: In the early days of the consortium, the pilot library spent much time and energy training all member libraries on image and metadata creation according to a set of statewide best practices, resulting in a large number of digital surrogates that were generated with long-term preservation goals for all. What was not considered at the project’s outset, however, was what might happen when the interest of project partners waned, either due to shifting institutional priorities or budget cuts. A consideration of the project’s potential pitfalls at the outset would have been ideal, allowing as it would have for creating tiered levels of preservation and participation needs. While retrospective project planning is not possible, a reevaluation of the project that takes into account the partners’ differing participation and preservation needs is, as is educating all partners about the project’s lifecycle.

Recommendations: Although it would require time and effort, continuum theory suggests that a second educational initiative be undertaken to discuss and implement preservation goals. Since this is a large project with many stakeholders, additional LSTA funding could be petitioned to allow adequate time for discussion, appraisal, and planning. While this might seem to be an overwhelming task involving certain compromises, the gains could be correspondingly significant. The primary achievement would be a statewide plan reflecting the different needs of the finding aids and the digital collections (for example, the consortium might agree to archive all iterations of the style sheets used to transform the finding aids but allow libraries to make local decisions about access copies for the digital collections). Secondary but no less important would be a higher level of awareness among all the libraries about digital preservation practices. This secondary gain might, in the long term, prove to be the more important. A preservation plan like this could be adapted for use at the local level in the smaller universities, giving staff both guidance and confidence.

Case Study #4: Projects with Ownership and Content Concerns

Summary: In 2012, a medium-large American research university’s library began a formal partnership with an overseas not-for-profit organization devoted to documenting 20th century genocides. According to the partnership agreement, the overseas organization will retain physical custody over the original materials at all times, while the American library will be responsible for all “digital” aspects of the project, ranging from capture of digital images, image processing, metadata, digital preservation, and digital access. An important component of the project is digital preservation, given the physical items’ proximity to a potentially unstable environment, both physically and politically. At the same time, however, digitization activities must take place onsite at the partner organization’s overseas headquarters, requiring the use of flexible, transportable equipment rather than a permanent, state-of-the-art digitization lab.

Challenges: Much of the material in this collection is extremely sensitive. Many documents include the names and even the home addresses of the perpetrators of a very recent genocide; similarly, the documents also reveal personally identifying information of survivors of the same very recent genocide. For the personal safety of all concerned, the project plans to provide access to redacted copies, with the redactions being determined by the overseas partner institution. Knowing this, the project plan includes the creation of two sets of documents: archival masters and redacted access copies. The project plan also includes a date on which the need for redaction should be visited, with the knowledge that, ultimately, the redacted access copies will one day be deleted.

Recommendations: This project provides an ideal exemplar for considering digital collection building through the lens of continuum theory: it is in the early stages of the planning process, allowing for the records continuum to be considered first rather than years down the road, after significant challenges have arisen. Continuum theory’s call for integrated processes is particularly important, for without knowledge of the records’ context, the digital collection becomes useless; without the inclusion of knowledgeable digital collections librarians and digital archivists involved from the point of digital record creation through the entire project’s life cycle, the contextualized records would be highly vulnerable to digital decay.

Case Study #5: Donors and Digital Initiatives

Summary: In 2004, a special library within a public university received a large donation of materials relating to American popular culture in the early twentieth century, including a collection of sheet music. The donor, an active member of the library’s friends’ group,
stipulated in the gift agreement that the sheet music would be made available online and provided an additional financial gift for this purpose. At the time of the gift, the historical society was in the earliest stages of its own digitization initiatives. For this reason, the society chose to work collaboratively with an outside vendor. The library would generate metadata and digital surrogates while the vendor would create and host both a content management system as well as an access website. The project was well-received by both the donor and by users, with positive feedback about usability and content. It has also been used by the society’s development officer as a tool for fund-raising and outreach to potential donors, giving it a high secondary value to the society’s administration.

Since the original project was established, library staff have continued to generate both metadata and digital surrogates, expanding the project to include sheet music from other local collections. There are now about four hundred discrete objects in the online collection. Statistics provided by the vendor allow the staff to track user behavior within the collection and to identify the most frequently viewed materials. In the nearly ten years since it was initiated, however, several significant changes have taken place. First, the library now has several other digital collections available, all of which are hosted locally and accessed via their webpages. These collections are of complementary interest to the sheet music. Second, with the natural progression of staff turnover and the evolution of best practices, the access website, the search functions and the metadata for the sheet music project are all in need of technical and descriptive updates. Finally, while the donor has passed away, one of the donor’s children retains an active interest in the donor’s contributions to the library.

Challenges: The library has identified several difficulties with the collection. The cost of updating both the website and the content through the vendor is prohibitive. While the library has all the digital surrogates and metadata readily accessible, migrating the project to its own content management system would require a very significant investment of staff time. The library believes, however, that users would find searching and discovery easier if the sheet music were integrated with the other digital collections. Some staff members have proposed ending the contract with the vendor and uploading only the most heavily used sheet music to the local digital collections, but this action would necessarily leave some potentially useful collections out of a central, searchable system and would ultimately work against the inherent benefit of searching across multiple collections within a single repository. From a public relations standpoint, the society also wishes to proceed with some caution, continuing to honor the spirit of the original gift agreement while still acknowledging the practical needs of the collection.

Recommendations: Although a plan for the vendor-hosted website was not implemented when the new content management system was adopted, a time that would have been a natural transition period, these challenges offer an opportunity for the library to re-appraise the sheet music collection. Continuum theory provides the theoretical foundation on which to build a serious analysis of the primary goals for both the collection and the context in which the collection was originally presented, in turn helping to establish a digital strategy that would simultaneously document the appraisal decision and create guidelines by which to make a unified content management system possible. The library’s digital strategy would then serve to integrate the sheet music migration project into existing budgets and workflows, thereby providing ongoing support within the institution’s priority functions, rather than perpetuate short-range planning.

Conclusion
Archival theory in general and continuum theory in specific tend to have a reputation as academic exercises, providing interesting fodder for thought but not much help solving quotidian problems in the archives (or the digital initiatives department). Digital projects management, on the other hand, is built on practicals, focusing on solutions for everyday problems with less emphasis on abstract conceptualization. This paper provides a bridge between the two different worlds, finding practical applications for theoretical ideas and looking to theoretical ideas for new solutions to a long-standing set of issues that have plagued digital projects since their inception. As the case studies show, these applications are not special but rather intended to move the conversation about digital projects away from novelty, boutique initiatives that exist to harness emerging technologies without a foundation for replication. Instead, they seek to demonstrate that digital projects are essential components of cultural heritage institutions that must and can be sustained.
through and across professional trends and during economic shortfalls. Where periodic maintenance has failed, continuum theory provides a method for considering digital projects through a new lens and breathing new life into projects that might otherwise be declared dead.

Notes
5. The authors follow the thinking of Christopher Lee and Helen Tibbo who have advocated the use of a common term, “digital curation.” Lee and Tibbo argue that it “has recently come into use, reflecting the increasing confluence of previously distinct communities [i.e. archives, libraries, museums, computer science, data engineering, etc.]” and follow other thinkers in applying the term to all aspects of maintaining digital materials—from creation to management to preservation. See: Christopher Lee and Helen Tibbo, “Where’s the Archivist in Digital Curation? Exploring the Possibilities through a Matrix of knowledge and Skills,” Archivaria 72 (2011): 24.
13. The American archival tradition typically ascribes to the philosophy of Theodore Schellenberg who believed that archivists should be involved in shaping the historic record by making selection decisions. British archival theorist, Sir Hilary Jenkinson, argued, on the other hand, that archivists should be passive recipients of records that have been saved by record creators. For an excellent analysis of these two lines of thought, see: Reto Tschan, “A Comparison of Jenkinson and Schellenberg on Appraisal,” American Archivist 65, no. 2 (2002): 176-85.