Collaboration for a 21st Century Archives: Connecting University Archives with the Library's Information Technology Professionals

Erin Lawrimore (elawrim@uncg.edu)
University of North Carolina at Greensboro

Abstract

As communication technologies change, so do the records being produced and acquired by the archival repositories tasked with documenting society. This article, written from the perspective of a University Archivist, discusses the need for collaboration between archivists and information technology professionals in a university library in order to manage the university's born-digital archival records. Using specific examples of collaborative projects of University Archives and the Electronic Resources and Information Technology (ERIT) department in the University Libraries of The University of North Carolina at Greensboro, the article makes specific recommendations for overcoming challenges related to professional jargon and work practices shared by archivists and information technologists to produce a successful collaboration.

Keywords: Archives; Born-digital Records; Information Technology

Introduction

While the mythical "paperless office" is nowhere near a reality, more and more records are created in a digital format – both in an office environment as well as in the personal realm. These born-digital records can range from a Word document to digital photographs to websites and other complex objects. The basic modes of communication in the modern world have changed. As such, the records being produced and the information recorded have changed in turn.[1]

Archival repositories are tasked with acquiring and managing records that best reflect the work or output of certain segments of society. Sometimes this segment is the administrative workings of a college campus. Sometimes it is local civic organizations or experts in a certain subject area. Regardless, in today's electronic age, born-digital records play a significant part in documenting the modern world. As a result, archivists are working to develop tools and best practices for acquiring, preserving, and providing access to these complicated formats in an effort to ensure that modern society does not lose a large swath of its documentary heritage.[2]

The change in communication methods also affects archival researchers. Increasingly greater numbers of researchers approach the archives through its digital portal using a website to navigate finding aids or catalog records and then contact the repository via email to begin the reference transaction. Additionally, these researchers are more frequently using online digital collections consisting of digitized images in a repository's archival holdings to see copies of the records without needing to contact the archives itself. To meet these needs and accommodate the typical archival researcher's workflow, archivists are responding through increased digitization efforts and by providing greater access to collections through digital portals.[3]

While many of these topics have correlations to traditional archival practices, many, if not most, archivists do not have the technical skills or resources necessary to build this 21st century archives based on born-digital records with the needed access to collections through digital means. While archivists may have a basic understanding of advances in communications technologies, they rarely have the programming skills needed to create a tool to effectively man-

age these responsibilities. This is where they must find partners who can provide the technical know-how necessary to take on and support these important tasks.

The unit tasked with managing information technology development within the library is, of course, the most obvious place to find the needed talent and resources. But collaboration between archives and information technology is not always a smooth road or an easy one to begin or to navigate.

Challenges in an Archives/IT Collaboration

Are We Speaking the Same Language?

The first challenge the archivist faces in working with information technology professionals relates to language. In fact, the term "archives" holds a completely different meaning for the two. The Society of American Archivists' (SAA) *Glossary of Archival Records and Terminology* defines "archives" as "materials created or received by a person, family, or organization, public or private, in the conduct of their affairs and preserved because of the enduring value contained in the information they contain or as evidence of the functions and responsibilities of their creator, especially those materials maintained using the principles of provenance, original order, and collective control; permanent records."[4] These are organically-created collections that are purposefully selected and maintained for their ability to document the work or life of the collection's creator. Many archivists insist on the word being used with the "s" (as opposed to "archive"), and many refrain from using "archive" as a verb.[5]

To an information technology professional, however, "archive" carries a different meaning. The *Oxford English Dictionary* denotes that in computing terminology, "to archive" means "to transfer to a store containing infrequently used files, or to a lower level in the hierarchy of memories, esp. from disc to tape."[6] Additionally, many tools used in relation to digital records use "archive" in an even more expansive way. Google's Gmail, for instance, lets you archive e-mail messages and "tidy up your inbox by moving messages from your inbox into your All Mail

label, so you don't have to delete anything."[7] In these uses, "archive" loses the sense of purposeful selection and organization inherent in the archivist's use of the term.

Similarly, to an archivist, "context" means, "the organizational, functional, and operational circumstances surrounding materials' creation, receipt, storage, or use, and its relationship to other materials" and is a fundamental aspect of a record. Where does a record fit within a larger whole, and how might that context affect the researcher's understanding of the record itself? The key archival principles of provenance and original order exist to ensure that contextual information about a record is maintained. Archivists arrange and describe on the collection level instead of on the individual document level in order to make certain that the important contextual details are retained.

Information technologists, on the other hand, may not see the value in retaining or maintaining context or the value of thinking in terms of archival collections as opposed to individual records. They may view the information contained within the records as the primary piece of the archival puzzle and not think of the importance of the record's structure, context, or relationship to other records in providing evidence of a creator's work or life. Without an understanding of historical research methodology, the information technology professional would have no background for understanding how an archival researcher might approach or use archival collections online or in person.

While the archivist and the information technology professional may work within the same organization, they may not be speaking the same language. The archivist must have a basic understanding of technology in order to identify these key areas for misunderstanding. But, perhaps more importantly, the archivist must understand why and how he or she uses these particular terms. What do these words mean, why are these principles in place, and how might they be articulated to the non-archivists? As well, the archivist must be willing to speak up when the information technologist uses unfamiliar terminology, and vice versa.

Long-term versus Adaptability

As professionals, archivists have a developed a history of attention to detail. Newly-acquired collections often are given complex arrangement and extensive description with carefully-labeled folders typically arranged within acid-free boxes. Their contents then are described in a comprehensive finding aid. As a result of this detailed level of work, a 2003-2004 survey of archival repositories showed that 60% had at least a third of their total holdings unprocessed (no level of arrangement or publicly-available description). Moreover, 34% of repositories had more than half of their holding unprocessed. While the minimal description mindset espoused in Mark A. Greene and Dennis Meissner's 2005 article, "More Product, Less Process: Revamping Traditional Archival Processing," (also known as MPLP), has resulted in a reduction of backlogs at a number of archival repositories, still more archivists maintain traditional, artisanal processing styles and insist that the attention to detail is the only way to ensure preservation and access in the long term. By processing a collection "right" the first time, the archivist creates a tool that he or she assumes will not need revision and thus provides preventative preservation measures to help ensure the collection's survival for the foreseeable future.[8]

Information technology professionals, on the other hand, do not have the lengthy professional traditions of archival practice. This field typically requires that professionals adopt a more flexible, project-based mindset. A specific need is identified, tasks are laid out, deadlines are set, and progress towards the stated goal is monitored. Rapidity and adaptability, not longevity, are favored—since it is understood that a tool created today may be considered out-of-date in the very near future. As Larry Oberg and Keiko Pitter noted as far back as 1994, information technologists "have been defined from their inception by flexibility, innovation, and responsiveness to change."[9]

As with impediments brought about by professional jargon, challenges of differing work styles must be addressed before a successful collaboration can occur. First, both parties should understand that, at the core, similarities exist. Archivists, like IT professionals, have always held a project management mindset in that each individual collection processed constituted a single project that must be planned and managed accordingly. Likewise, information technologists, like archivists, should have a keen attention to detail in order to produce useful tools and programs. In order to create a productive partnership, the two must meet in the middle. The archivist must develop projects that fit within the scope and capabilities of an IT department, but also fit into the larger scope of the archives' mission and goals. The archivist and information technologist must work in tandem to create a timeline with structured markers for judging progress towards the end goal, with room for change as development progresses. In doing so, the archivist must abandon the mindset of "getting it right the first time" and be more responsive to quick, on-the-fly changes and a sense of agile development. Rigidity in workflow and process can result in a product with a very small range of application and time of use, and might quickly cause an IT project to grind to a halt.

How Much Risk Is too much Risk?

Archivists as professionals are tasked with preserving records with enduring value. As a result, they often think of and work to prevent possible preservation concerns long before they actually occur. As noted in Greene and Meissner's "More Product, Less Process," the removal of staples, rubber bands, and other fasteners, and the re-boxing and re-folding of newly-acquired archival collections are tedious tasks that may not be necessary with modern records, but which continue to be practiced by many processing archivists in the name of long-term preservation. The controversy over reappraisal and de-accessioning of archival collections demonstrates the profession's notion of permanence of records deemed to be "archival," and implies that decisions made at one time are not to be questioned or reversed at another.[10] With this focus on permanence comes a tendency to avoid risk, as risk might endanger the longevity of the records.

On the other hand, information technology professionals constantly work in a changing field.

Technological obsolescence can occur rapidly and products may become obsolete soon after they are created (if not before). Agile development as a mantra provides information technologists with a mindset for efficient and effective project management. As stated in one of the four key points of "The Manifesto for Agile Software Development," these information technology professionals value "*responding to change* over following a plan."[11] As tools develop and technologies change, so must the workflow need to change. Each piece of the puzzle carries an inherent risk that it might be quickly changed or even be discarded outright as project needs evolve and the development process proceeds.

To promote effective collaborations and to work effectively in today's constantly changing world of communications technologies, archivists must be willing to adopt both an MPLP mindset (in which work done now need only be the baseline, not the ideal end goal) as well as an agile development mindset (in which change is expected and work can quickly adapt when needed). In doing so, the archivist must be willing to accept some level of risk, acknowledging that the product being developed might not be "perfect" and might quickly become out-of-date due to evolving technologies, workflows, or needs. The information technology professional, in turn, must acknowledge that preservation is a key component of archival practice. Certain risks that might result in total loss of critical information or violation of copyright laws are not acceptable and must be avoided in order to ensure the continued usability and accessibility of the records.

Archives/IT Collaboration at UNCG

At The University of North Carolina at Greensboro's (UNCG) University Libraries, staff in the Electronic Resources and Information Technology (ERIT) unit frequently work with colleagues in University Archives to better acquire, preserve, and provide access to the University's records of enduring value. Examples of collaborations taking place since January 2012 include:

- **Digitization:** The Digital Projects Librarian, housed within ERIT, is tasked with digitizing records that are unique to UNCG and specifically the UNCG Libraries. The majority of these unique resources reside within University Archives and its parent department (the Martha Blakeney Hodges Special Collections and University Archives, or SCUA). Archivists work in tandem with the Digital Projects Librarian to select records and provide descriptive metadata for digitization projects.

- **Campus Map**: Using a tool developed by ERIT, University Archives staff populated the information bubbles on the campus's Google-based map (http://map.uncg.edu) and created mobile-friendly web pages that provide additional historical details and images about campus buildings. Tying these tools together will allow staff to create mobile historic walking tours of campus.

- **Online Exhibits:** ERIT web developers scripted a template webpage using a slideshow plug-in to create a quick and simple way for archivists to produce basic online exhibits. This template allows the archivists to develop individualized exhibits for donors or classes, to create a web version of physical exhibits housed in the library, and to promote more readily its holdings to a broader constituency.[12]

Perhaps the most in-depth collaboration between the two areas, however, has come with the development of a tool aimed at acquiring, processing, and providing access to born-digital records. This project began with a "simple" need from University Archives – a way to maintain the University's records of enduring value that were created and maintained in a digital format. A number of University records that provided valuable historical insight into the University's work had no analog counterpart, and therefore needed to be acquired by University Archives in their native digital format or they would risk being lost. The project included handling newer resources and records like the University's homepage as well as those that had evolved from print to a digital format, such as the annual *Course Bulletin*.

While advocacy for the management of born-digital records began when the University Archivist arrived at UNCG in July 2011, formal discussions between ERIT and SCUA about

methods for managing these materials did not begin until early 2012. Conversations brought together the head of SCUA, the head of ERIT, and the Dean of the Libraries to discuss the importance of developing a means for acquiring and managing these records in a timely fashion. The University Archivist and the lead developer for the Libraries' Web Applications Development Team (Web Developer) were asked to outline specific product needs, investigate existing tools, and report on existing practices at the other institutions in the University of North Carolina system.

At the beginning of the Fall 2012 semester, the two reported that no existing products fully met the needs of both SCUA and ERIT. Products either lacked some of the functionally required by University Archives, were not compatible to the Libraries' supported development platform (ASP.NET), or both. Therefore, it was proposed that the Web Applications Development Team work with SCUA to create a custom tool, meeting requirements of both departments in acquiring and managing born-digital records. This tool, named BDRM (born-digital records management), is projected to be a flexible, integrated tool that will allow donors to upload digital records and archivists to manage digital records using a single product.

Goals for BDRM were established at the outset by both ERIT and SCUA staff members. Specifically, the previously-conducted survey of existing tools yielded a solid framework for discussion of needs and possibilities. Based on this, the University Archivist and Web Developer created a model for a tool at UNCG supporting an appropriate workflow and set of capabilities. This model was then presented to the other members of the BDRM team, along with upper-level library administrators.

To reflect the importance of this tool for both University Archives and the University Libraries, its development was made a priority goal for each department as well as for the Libraries itself for the 2013 fiscal year. A product development timeline was established, with the overall project broken into three year-long phases:

- **Phase One: Accessioning** (Fiscal Year 2013): In this initial phase, a web-based tool for acquiring born-digital content from University and non-University donors would be created, tested, and implemented. Storage needs would be addressed, and initial metadata concerns would be managed.
- **Phase Two: Processing** (Fiscal Year 2014): This second phase would support the processing (arrangement and description) of the archival records ingested using the tool developed in Phase One. Included would be the ability to weed records, to describe records, and to apply necessary levels of access restriction.
- **Phase Three: Access** (Fiscal Year 2015): The final phase would develop a tool for providing access to the records acquired and processed in the earlier phases. This would include the ability to limit access to the records as needed (based on copyright or other restrictions).

Work on Phase One began in earnest in September 2012. A timeline was established with clear goals for progress agreed to by both departments. The University Archivist and the Web Developer worked collaboratively to determine workflow and metadata needs, in part based on other available tools that were unable to be supported by ERIT. The University Archivist used MODS (Metadata Object Description Schema) to create a schema for describing the newly-acquired, born-digital archival records, while the Web Developer used PREMIS (Preservation Metadata Implementation Strategies) to guide the creation of preservation metadata for the records. Throughout the development process, the Web Developer and ERIT's Web Applications Development Team worked to build the supporting database and the interface, while the University Archivist managed the text pages of the application and the descriptive information gathered by the tool.

The previously addressed challenges in creating a successful archives and IT collaboration cropped up throughout the first phase of the BDRM project. Initially, all parties needed to ensure that they understood the needs and capabilities of the others. For instance, at the outset, the developers primarily focused their atten-

tion on the process involved in uploading files. The archivists, meanwhile, concentrated on the archival principle of provenance, the need to maintain the records from a single creator as a cohesive unit. Once the archivist recognized this disconnect, she was able to discuss the importance of context with the web developers. They in turn were able to understand better the archives' needs and interpret them in a useful way for the project.

Similarly, ERIT's focus on managing the project in an agile manner meant that many of the archivists involved were forced to move at a pace that was quicker than they were accustomed. Often questions would arise that were applicable to later iterations of the project, but not essential to the phase at hand. The web developers and IT professionals were able to effectively manage the project through frequent communication, incremental development to ensure the project was proceeding along the desired path, and flexible testing that allowed for adaptability. Through this process, the archivists have been challenged to take risks and accept "good enough for now" – mindsets that are frequently at odds with an archives' or archivist's notion of permanent preservation.

Frequent communication as well as updates to the other involved parties throughout the University Libraries has resulted in a smooth development process. The accessioning tool (Phase One) is on track to be publicly available prior to Fall 2013, with development on the processing tool (Phase Two) to begin at the same time.

Lessons Learned (and Still Learning)

While the collaborative project to manage born-digital records is still very much in the development phase, the lessons learned thus far at UNCG highlight ways that archivists can bridge the gap and work collaboratively and effectively with library information technology professionals to create and manage tools for building a 21st century archives.

Be Willing to Learn

No one on staff at the UNCG Libraries had expertise or much prior experience working with

born-digital archival records before beginning the collaboration with ERIT. In the development of the born-digital records management program, the University Archivist attended webinars, workshops, and other training activities to learn more about best practices for managing digital records as well as existing workflows and metadata schemas implemented by other repositories.[13] New technologies were learned, and the University Archivist and Web Developer educated each other on the needs and requirements of each party in developing this new system.

While this type of collaboration requires the archivist to have a baseline understanding of communications technologies, it does not require either party in the partnership to have an *extensive* understanding of the other field. Instead, it simply requires both individuals to be open to learning about the other area and an ability to find common language and work plan for a successful product development.

Make Collaboration a Priority, Not an "Extra"

Simply wanting to work on a collaborative project will never make a project happen – even if both parties *really* want it to happen. Instead, the partners must effectively advocate for their collaboration and get buy-in from supervisors and other administrative leaders for the collaboration. It must be clear that the collaboration is a priority project and that time and support for development of the partnership will be provided.

At UNCG, the University Archivist had limited experience working with born-digital records but emphasized the importance of acquiring and maintaining these records in documenting the work of the University in the 21st century. Presentations were given to the Libraries' administrative leaders and department heads, talking points were created for the SCUA department head to hold conversations with the Dean of the Libraries, and other parties across campus (including the campus's Information Technology Systems unit) were brought into the conversation, all in an effort to gain broad support and buy-in for the development of the born-digital program. It was only after this wide-scale advocacy effort that the development of a system for

acquiring and managing born-digital archival records was made a departmental and University Libraries priority.

Because both departments involved, as well as the Libraries as a whole, committed to making this a priority, these partners could dedicate time and resources to the project with an understanding that it is contributing to a recognized goal. The born-digital records management development has proceeded within the specified timeline in spite of budget cuts and other identified projects simply because all involved have identified it as a priority.

Conclusion

The need to preserve born-digital content and to promote archival finding aids and records in the modern researcher's digital workflow necessitates a strong working relationship between archivists and information technology professionals. Archivists are specialists in their collection areas and in knowing the information needs of their constituents. In turn, information technologists have the programming and development skills necessary to address the technical components of modern archival issues. Only with jargon-free (or jargon-neutral) communication, with clearly delineated project workflows, and with a prioritization of joint projects of both the University Archives and IT can these collaborative projects succeed.

Endnotes

[1] For more on defining born-digital records, see Ricky Erway, "Defining 'Born-Digital'," https://www.oclc.org/content/dam/research/activities/hiddencollections/borndigital.pdf (accessed April 1, 2013.

[2] Examples of tools currently in use or under development include Archivematica (http://www.archivematica.org), BitCurator (http://www.bitcurator.net), and the Duke Data Accessioner (http://www.library.duke.edu/uarchives/about/tools/data-accessioner.html).

[3] See Erway, 2011. *Rapid Capture: Faster Throughput in Digitization of Special Collections*. Dublin,

Ohio: OCLC Research, http://www.oclc.org/content/dam/research/publications/library/2011/2011-04.pdf (accessed April 1, 2013); as well as a number of mass digitization projects (such as the Triangle Research Libraries Network's "Content, Context, and Capacity: A Collaborative Large-Scale Digitization Project on the Long Civil Rights Movement in North Carolina," http://www2.trln.org/ccc/index.htm) and archival web portals (such as the California Digital Library's Online Archive of California, http://www.oac.cdlib.org).

[4] Richard Pearce-Moses, "Archives," in *A Glossary of Archives and Records Terminology* (Chicago: Society of American Archivists, 2012), http://www2.archivists.org/glossary/terms/a/archives (accessed April 1, 2013).

[5] For examples, see "The Increasingly Common Use of "Archive" as a Verb" and ensuing comments at *ArchivesNext*, November 19, 2010, http://www.archivesnext.com/?p=1726 (accessed April 1, 2013).

[6] *Oxford English Dictionary*, s.v. "archive, v." accessed April 1, 2013, http://www.oed.com/view/Entry/10417?isAdvanced=false&result=2&rskey=yOWZWX&.

[7] "Archive mail," http://support.google.com/mail/answer/6576?hl=en (accessed April 1, 2013).

[8] Mark A. Greene and Dennis Meissner, "More Product, Less Process: Revamping Traditional Archival Description." *The American Archivist* 68 (Fall/Winter 2005): 208-263. For a summary of contradictory views, see Dennis Meissner and Mark A. Greene, "More Application while Less Appreciation: The Adopters and Antagonists of MPLP," *Journal of Archival Organization* 8 (2010): 174-226.

[9] Keiko Pitter and Larry Oberg, "Toward a Model of Computer Center-Library Cooperation for the Twenty-First Century" (Paper presented at the annual meeting of CAUSE, Australia, 1994). 2, quoted in Mark Cain, "The Two Cultures? Librarians and Technologists," *The Journal of Academic Librarianship* 29:3 (May

2003): 177-178. See also Karen Patten, Jerry Fjermestad, et al., "Leading IT Flexibility: Anticipation, Agility, and Adaptability," *AMCIS 2005 Proceedings* (Paper 361), http://aisel.aisnet.org/amcis2005/361 (accessed April 1, 2013).

10 For more on the controversy over reappraisal and deaccessioning of archival collections, see Mark A. Greene, "What were we Thinking?: A Call to Embrace Reappraisal and Deaccessioning," *Provenance* 20 (2002): 33-49; and Mark A. Greene, "I've Deaccessioned and Lived to Tell About It: Confessions of an Unrepentant Reappraiser," *Archival Issues* 30:1 (2006): 7-22.

11 "The Agile Manifesto," accessed April 1, 2013, http://www.agilealliance.org/the-alliance/the-agile-manifesto. (Italics mine.)

12 For an example of the online exhibits template, see "Spartan Evolution: A History of Basketball on the UNCG Campus from the 1890s to Today," http://library.uncg.edu/collection/exhibits/History_of_Basketball_at_UNCG.aspx (accessed April 1, 2013). This online exhibit was created to complement a physical exhibit housed in UNCG's Jackson Library from February to April 2013. In addition to providing the information to interested parties who were unable to see the physical exhibit, the online exhibit will remain available indefinitely, thereby extending the life of the exhibit beyond the period of physical display.

13 Examples include the Digital Archives Specialist curriculum and certificate program offered by the Society of American Archivists (http://www2.archivists.org/prof-education/das).