

Lora J. Davis
Assistant Archivist
Special Collections and University Archives
Colgate University Libraries
Hamilton, NY

May 2, 2013

Jump In Initiative Final Report

In December I responded to the call for participants in the “Jump In” initiative by writing to Chris Burns, in part, “Thanks to you and the Section for getting this effort off the ground. I know this will be a very beneficial exercise for us, and this initiative should ‘keep us honest’ and ensure that we tackle this worthy goal that has been on all of our to-do lists for too long.” As I sit here squeaking this essay in under the buzzer, it has become even more evident to me that, in this profession comprised of ever-shifting priorities and demands, a little hand-holding can go a long way. After completing my inventory of born-digital media in the holdings of Colgate University Libraries Special Collections and University Archives (SCUA) I am left with several key observations, numerous questions large and small, a few graphic reminders of the importance of following best practices, and a much better sense of how to craft our born-digital collecting practices in the future.

I arrived at Colgate University Libraries on August 1, 2012, so this initiative nicely coincided with my own process of becoming familiar with the holdings of the repository. At the time that I signed up for the project, I had seen a few lone 3.5” floppy disks, USB thumb drives, and DVDs cross my desk during accessioning, but I had no real sense of our existing holdings of digital content on physical media. However, knowing the volume of work done in the digital environment in a modern institution of higher learning, I assumed our University Archives would contain many items worthy of inclusion in this survey.

Unfortunately, the vast majority of the Archives is unprocessed, and no tradition of accessioning or establishing any kind of intellectual control over our holdings existed prior to the arrival of Sarah Keen, Head of Special Collections and University Archivist, in 2010. Although I had no institutional records or memory to draw on, I had the benefit of conducting this survey concurrent with another much-needed inventory of our holdings in all formats. Though that larger collections inventory is still ongoing, knowledge of the pending due date for this initiative also served as valuable motivation in beginning that task as well.

Ultimately, my Jump In survey revealed a much smaller backlog of born-digital media than I ever anticipated was possible. All told, I discovered 49 pieces of physical media located in 11 discrete collections (comprised of both existing processed collections, unprocessed accessions, and internal departmental records). Only 5 primary physical media types are represented, including 1 external hard

drive, 22 3.5" floppy disks (with at least one identified as Mac formatted), 10 DVDs, 15 CDs, and 1 USB thumb drive. The estimated maximum storage capacity for these items is just 1.09TB.

My initial response to these results was relief. The backlog of physical media was both much smaller than I anticipated, but also much newer than I anticipated. Having previously worked in an institution that held laser discs, 5.25" and 8" floppy disks, and multiple forms of early magnetic media, I was delighted to find that the media currently in our repository are all still within our ability to access. The department has retained one easily quarantined desktop with a 3.5" floppy drive, and, for the Mac-formatted disks, I can resort to using my personal Macintosh G3 Powerbook with SuperDrive – a tactic I employed at my former repository. All-in-all my findings were not as bleak as I expected; the challenges that do exist are likely manageable with our current resources.

Of course, with this good news came a new onslaught of questions. Certainly, my initial premise – that any modern institution of higher learning produces volumes upon volumes of electronic records in the course of its normal business – is true. So, if we don't have these materials (or at least analog surrogates of them), where are they? Though I entered this project hoping to get a better handle on the needs of the electronic media currently in my Archives, I'm leaving it with renewed focus on the records existing on campus at large. The report by OCLC Research that served as the basis of this inventory explains, "With relatively few steps and limited resources it is possible to know with more certainty what your repository holds and to make vulnerable data more secure."¹ Entering this inventory I placed much more emphasis on the second goal articulated above, namely I hoped to identify and inventory vulnerable data within our collections. Now I realize that the inverse of the first goal – knowing what my repository *doesn't* hold – is the far more important outcome of this project for our particular institution. It turns out the most vulnerable data out there is likely the data that is not currently in our physical custody.

After discovering how little physical media we actually hold, I decided to move away from the initiative's singular focus on born-digital media, and included digitized items stored on physical media in this inventory as well. Currently, 8 of the 49 items listed in the inventory are digitized, rather than born-digital, in nature. These items include both those digitized by the department itself, but also donations and transfers of digital surrogates of analog originals still retained by various campus units and individual donors. Ultimately, I found the addition of these digitized items to be a particularly helpful goalpost for this project moving forward. Regardless of the content of the media, I wanted to ensure that the department is keeping a solid handle on all of the physical media currently under our direct purview. By working with these largely non-collection backups and working files I will gain valuable experience on managing the digital products this department creates and maintains during the course of our operations.

¹ Ricky Erway, *You've Got to Walk Before You Can Run: First Steps for Managing Born-Digital Content Received on Physical Media* (Dublin, Ohio: OCLC Research, November 16, 2012), 3, <http://www.oclc.org/research/publications/library/2012/2012-06.pdf>.

These experiences are exactly those that we will need to relay when working with other stakeholders on campus to bring their digital products into our repository.

Another unanticipated outcome of this project was the discovery that our fledgling accessioning program is poised to support future growth of born-digital records holdings. Though our nearly complete lack of accession records, donor files, and acquisition records prior to 2010 makes any sort of retrospective projects impossible, we can still look forward and ensure that all future accessions are managed in a way to make the identification, prioritization, and management of digital content on physical media possible. Just by following existing best practices, we have already set a solid foundation for moving forward. Since I assumed accessioning duties in August 2012, I have routinely applied genre terms culled from Getty's Art & Architecture Thesaurus to our Archivists' Toolkit accession records. As a result, when I began this survey I simply ran a report within Archivists' Toolkit for the terms "DVDs," "Compact discs," and "Electronic reproductions," and within seconds had a complete and accurate listing of all accessions that contained physical media. As we continue to receive new materials, I will have a renewed commitment to identifying via controlled vocabularies the carrier types that bring digital media to our doorstep.

Having used this survey instrument to assess the current state of our digital holdings, identify aspects of our current workflows that are important to electronic records initiatives moving forward, and establish future collecting venues, the department is now much better prepared to articulate specific goals for our future born-digital records program. These goals include:

1. To conduct a campus-wide records survey that places electronic and paper records on an equal footing to both records' holders and the department. (Scheduled Summer 2013)
2. To continue positive changes made in the accessioning workflow, especially in light of the potential growth of collections in the future. (Ongoing)
3. To continue experimenting with the digital preservation system Archivematica², acknowledging that our currently small, yet diverse holdings can serve as a wonderful sample set of initial material to play with. (Anticipated Fall 2013)
4. To implement a web archiving program utilizing Archive-It³. (Fall 2013)
5. To continue to update this records survey instrument as new materials are discovered in the collection or added through donations and transfers. All updates will be saved as new documents following the naming convention currently in use, e.g. starting the file name with the date of the update in yyyyymmdd form. (Ongoing)

² "Archivematica," accessed May 2, 2013, https://www.archivematica.org/wiki/Main_Page.

³ Internet Archive, "Archive-It - Web Archiving Services for Libraries and Archives," accessed May 2, 2013, <http://www.archive-it.org/>.