

Augustana College Special Collections
Jump In Three/Third Time's the Charm
Report submitted by Samantha Crisp, Special Collections Librarian
7 April 2015

Background

Special Collections is a department of the Thomas Tredway Library at Augustana College, a small liberal arts college in Rock Island, IL. Our repository contains Augustana's collections of rare books and manuscripts and the college archives. The primary mission of Special Collections is to collect, preserve, and make available materials that serve the teaching and research needs of Augustana College and the local community. Our manuscript and archival collections focus on Augustana College offices and services, student groups, alumni, and faculty; Rock Island and Quad Cities history; and local civic organizations and businesses. Our holdings include approximately 2,500 linear feet of material arranged into just under 400 collections. Materials date from the 17th century to the present, but the bulk of our collection hails from the 19th and early 20th centuries.

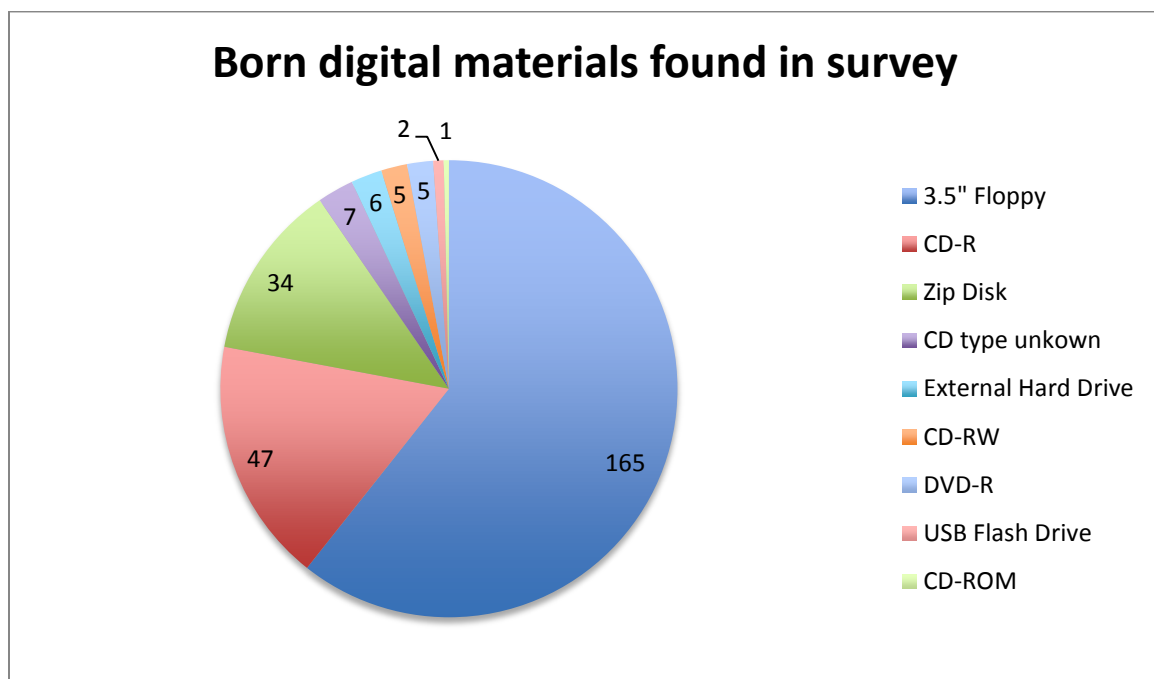
Augustana College Special Collections has previously had no mechanism for addressing born digital materials before participating in the Jump In Initiative. I (the Special Collections Librarian) was the only staff member working on this project. I am the only full time employee of the department, and I just took over the position in June. No record of any previous electronic records management program exists, beyond some basic introductory research on the subject.

Survey

For this project I attempted to conduct as complete a survey of our archival holdings as possible. Most of our holdings are older, paper-based collections, and because I anticipated finding very few born digital items, it seemed best to try and knock everything out at once. I did not include audiovisual materials (sound or video recordings on removable media) because we are currently in the process of conducting a separate survey of those holdings.

Roughly 60% of our collections have finding aids, so I began by performing a keyword search of existing finding aids to see if any of them mentioned electronic media. For those collections without finding aids, I used accession records and collection-level descriptions to determine whether they could conceivably contain any born digital materials. I eliminated from consideration any collection whose materials predated the 1970s and was left with only about a dozen collections to check. A quick check of the boxes revealed four collections with born digital content, plus a handful of other unlabeled items and departmental files. I then physically arranged and rehoused the media in each collection, assigned each disk or unit a unique identifier, and created a spreadsheet with each item's description, URI, file size, and any other pertinent information. I used this spreadsheet to prioritize each collection for reformatting. The entire process took about two months to complete, but I was only dedicating about 2-4 hours per week to the project during this time.

I found that, as suspected, our collections contain very little born digital media. The few items that are there are largely our own departmental records and data, including a number of items in our Thomas Tredway Library records (MSS 109). However, three of our manuscript collections do contain born digital content, largely on 3.5" floppy disks and zip disks. The survey found approximately 270 unique born digital carriers (disks or external drives), with a maximum potential file size of about 4 TB. I expect the actual amount of data to be much smaller than that, as this number reflects the total storage capacity of the carrier rather than the data itself. Most content was identifiable and hailed from the late 1990s-early 2000s, with the exception of a few unlabeled external hard drives (I expect that these contain backups of our department's digital files and collections).



Total amount of data by medium									
Format	CD-R	CD-RW	CD type unknown	CD-ROM	DVD-R	Zip disk	3.5" floppy	External hard drive	Flash drive
Number of items	47	5	7	1	5	34	165	47	5
Storage capacity	32800 MB	3350 MB	?	?	23.5 GB	5300 MB	237.6 MB	4+ TB	?
Total	4.065 TB+								

The project went surprisingly smoothly and I encountered no real setbacks or challenges, beyond a couple items with no labels to indicate their content or storage capacity. I was somewhat surprised that it went as quickly as it did. One of the biggest benefits to the survey was the discovery of a few dozen disks containing images and mock-ups for our college yearbook, the *Rockety-I*, dated after the last yearbook was printed in 2004. I suspect these contain the designs for an additional yearbook project that never made it off the ground, and thus the content of the files is likely of high value to our constituents (the *Rockety-I* is one of our most highly consulted resources).

Next Steps

I plan to prioritize the *Rockety-I* images described above for reformatting as soon as possible. Our ITS department has equipment capable of reading floppy disks and zip disks, so I hope to use their equipment to generate preservation and use copies of the files rather than sending them to a vendor for reformatting, provided the ITS equipment will not physically harm the media. I plan to wait until the parent collections for the remaining items are processed to consider reformatting those items, and I have no established timeframe for that work.

Although we don't currently have much born digital media in our collection, we will undoubtedly take in more of it as time goes on. The spreadsheet that resulted from this project will become a living document that will be added to and modified as we add new electronic records to the collection. Without participating in this project, I would not have been able to develop such a comprehensive approach to establishing initial physical and intellectual control of born digital content that comes into our collection. I am excited to see the project through to its next stage and develop preservation workflows for the content I discovered.