Jump in, Too/Two Report

Society of American Archivists
Manuscript Repositories Section
May 1, 2014

Daniel Weddington
Digital Archivist
Special Collections Research Center
The College of William & Mary

Overview

The Special Collections Research Center at the Swem Library is the repository for university archives, rare books, and special collections at the College of William & Mary. Our library strategic plan calls specifically for us as an organization to "create rich digital collections" and, as Digital Archivist, I play a large role in ensuring that this goal is carried out. I took the *Jump In* initiative as an opportunity to focus our efforts on building the structures needed to provide and care for the digital collections in our the Special Collections Research Center, with an eye toward aiding in the creation of mechanisms for better digital stewardship library-wide. With no digital content inventory in place, no set method for accession and ingest of digital records, and no formal plans for access and preservation of digital content, this was a chance to begin making digital stewardship at the Special Collections Research Center a priority not only in concept, but also in practice.

Parameters:

The first step to this project was to identify a scope. We have more digital content than we could feasibly address in a semester, so we had to establish priorities for the initial *Jump In* inventory.
The decision was made to approach this inventory in three phases, the first of which would be material collected for the *Jump In* initiative. We would start on a small selection of material and build out further with each phase.

First inventory phase (Jump In Initiative):

- *Digital data originating from archives and manuscript collections currently stored only on physical media*

  This phase did not include a large majority of material stored in the University Archives Audio-visual collection, as portions of that collection have been transferred to other storage media and in many cases did not represent unique data (commercial audio CDs, for example). This phase also did not include digital material currently stored on local network drives or hard drives originating from Special Collections. The concentration is on material produced outside Special Collections donated to Special Collections as part of archives and manuscript transfers.

Second inventory phase (Special Collections-wide):

- *All remaining digital materials within Special Collections*

  This phase will include all remaining digital material from the following locations: University Archives Audio-visual collection; local network drives and hard drives; online Special Collections material that exists only on platforms other than the W&M Digital Archive (images in Flickr, videos on Youtube, etc.).

Third inventory phase (Library-wide):

- *Digital material produced by Swem Library*

  This phase will become part of a longer-term plan to have a campus-wide digital asset management system, the beginning discussions of which have already begun. In order to assess campus-wide DAM needs, each stakeholder in the project -Swem Library being one of several- will need to provide an overview of their digital storage and access needs in order to make decisions on acquiring an appropriate DAM.
  Swem Library, Special Collections notwithstanding, produces a great deal of digital material (Swem Media Center and our External Relations, for example, produce a lot of audio, video, and image data). While much of the more ephemeral material would be best served by the DAM, the material of lasting value would eventually come to the Special Collections Research Center. While Special Collections will not directly inventory material this material for the whole library, we do intend to provide inventory templates from the previous two phases to aid in the process.

Process:

The set-up procedures were simple enough. I designated one of our work-study students to assist directly with the project. We set aside two sets of shelves in the Digital Archivist office: one for items not inventoried, the other for items in the inventory. Once we had a space, we then worked with the University Archives Specialist to identify material according that met our initial parameters and moved them to the pre-inventory shelves. Luckily, a great deal of our digital material had already been separated out from paper collections upon accession, so there was minimal time spent scouring our Archon database to locate digital materials in the records (many thanks due to our University Archives Specialist, Steve Bookman). Once we had identified our materials I then created a spreadsheet in Google Drive for the work-study and me to work from and we began surveying the materials.
Expanding on the templates provided by SAA, the inventory headings from the spreadsheet are as follows:

- Accession ID
- Itemized Accession ID
- Media Markings (Title)
- Original Collection
- Transferred Collection
- Physical Location
- Media Type
- Label Information
- Estimated total storage capacity
- On physical medium?
- On network drive?
- In DSpace?
- Description
- Additional notes
- Physical media photo
- Donor Info
- Possible action

We took the best care possible to allow for further information as the inventory moved from a cursory overview of materials to a more in-depth and exhaustive investigation.

The final component of the process involved setting up a camera stand to capture images of the physical media as it was inventoried. Eventually, once we begin transferring files to our digital archive and begin arrangement and description, we will attach these images to our object metadata. This will provide for us a level of provenance even after data moves off the physical media and the physical media is discarded or repurposed.

Findings:

From our initial inventory:

- 15 Collections
- 194 individual items
- 9 Unique Physical Storage Formats:
    - 3 1/2-inch HD diskette
    - 5 1/4-in floppy disk Double-Sided, Double-Density
    - 5 1/4-in floppy disk (type unknown)
    - CD-R
    - DVD-R
    - DVD-RW
    - Enhanced IDE Hard Drive
    - Flash drive
    - Zip-1
- 351.6904 GB estimated total storage

Next Steps:

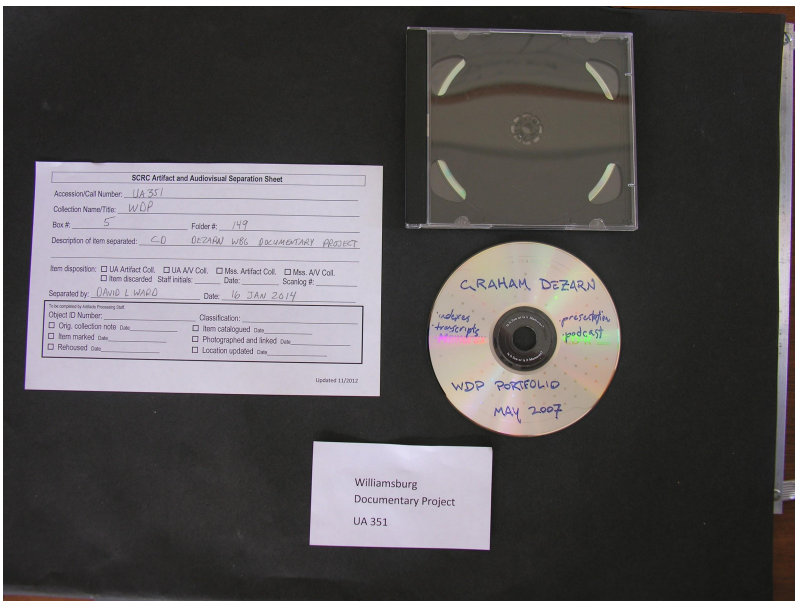
Now that we have an initial inventory, we will begin pulling together resources and establishing workflows for the next steps in the digital archives process. Our inventory gave us a quantifiable justification for resource allocations, so we will now be getting the equipment to begin transferring the materials off portable physical storage, properly accessioning, and begin arrangement and description. I will work with our University Archives Specialist and Metadata/Cataloging Librarian to create systems for accessioning, transferring, ingesting, and making our materials accessible in the W&M Digital Archive. This work will be done in conjunction with our move to the second inventory phase as well as with the creation of our first digital preservation policy. The eventual goal is to establish comprehensive policies and procedures guiding each stage of the digital lifecycle: from intake, to transfer, to validation, to access, to preservation and migration.

Conclusions:

The first and perhaps most obvious take-away from this project would be that digital material quickly becomes overwhelming. Even as we were doing the inventory we constantly had new digital material coming in. Even leaving aside the enormous amount of digital material we have left in our University Audio-visual collection, on our network drives, or on social media platforms, the amount of digital content not currently being dealt with is plentiful and ever-growing. It was an undertaking just to get a clear inventory of the materials we had, let alone to move to the next steps of transfer, arrangement, description, and preservation.

With the quick pace at which digital collections grow, and the inherent fragility of the digital materials in those collections, it is difficult sometimes to know where to begin. However, this fast growth and fragility are the very reasons why there is a need for not just action, but immediate attention. This is why I see the *Jump In* initiative as an incredibly important exercise for any institution. In my experience, and from what I have gathered from the experiences of others in the field, in terms of priority setting in archives digital content work has a tendency to get pushed to the periphery. It becomes easier to overlook a few GBs worth of material on a 1 inch flash drive than it is 30 cubic ft. boxes full of unprocessed paper records. However, the risk of quick, permanent loss is far greater for those digital files than that of the papers. There is a cognitive dissonance guiding how we respond to our digital stewardship responsibilities. What the *Jump In* initiative forces on an organization is expediency where it is needed most. We tend to put off the work that needs to be done with digital content simply because it does seem overwhelming, because we may not have all of the answers, or because we feel we do not have the resources we need to do what needs to be done. What this project has provided is at least a starting point. And now, knowing what digital material we have, we can make a more accurate assessment of our needs for better decisions as we proceed. Even if we do not have all the fine details sorted out for how we intend to go forward, at the very least we know we have begun.

Media Examples



Physical Setup

Shelving after project completion

Camera area for capturing storage media images