Society of American Archivists' Manuscript Repositories Section
Jump-In Initiative: Electronic Records Survey
The Westminster Schools Archives
Pamela Nye, CA
May 1, 2013

Scope of Survey

I limited the scope of this survey to the physical, removable media in one collection. This survey does not take into account any electronic records that are not yet accessioned into the archives or are in other collections that are unprocessed. This particular collection is one that is separated by format, not by content, and consists of eleven boxes (7.75"W x 12"L x 5.75"H), containing 87 CDs and 273 DVDs, or 360 removable media total. The range of content included school-wide assemblies, class projects, special events, and performing arts. The previous archivist had carefully accessioned these electronic records by copying them onto the CDs and DVDs and listing the content, but she had not noted the size of the files or the type of files included. It is a great, discrete collection to begin this foray.

Procedures

As per the survey instructions, I did the following:

- Located computer media in any physical form.
- Recorded the location, inventory number, type of physical medium, and any identifying information found on labels or media such as creator, title, description of contents, and dates. If no identifying information existed, I indicated this as "unknown."
- Recorded anything that was known about the hardware, operating systems, and software used to create the files.
- Counted the number of each media type, calculated the total maximum amount of data stored in each medium, and then calculated the overall total for the collection.

The information went into a structured spreadsheet where I am also inventorying all AV materials, and will become part of the larger survey of the collection. The fields I created were:

- Institution: TWS (The Westminster Schools); this field will eventually include WS (Washington Seminary) and NAPS (North Avenue Presbyterian School), our predecessor schools.
- Collection: The name of the collection; in this case, the Audio-Visual Department.
- RG: The Record Group number; in this case, RG 28.05 (AV Department-CD/DVD)
- Subgroup: The subgroup name: CD/DVD.
- Series: Some of the Record Groups do have a series. This particular collection does not have one.
- Box: Box number
- Item: Item number (unique)

- Description: The information found on the front of the CD or DVD.
- Date of event: If known, the date of the actual event, typically found on the front of the medium.
- Date of file: This is the date that is on the directory of files on the medium.
- Location: The shelf location of the box in the archives.
- Accession: The accession number of the item.
- Actual File Size (GB): The total, actual size of the files on the medium in GB.
- Media Size (GB): The total amount that the medium *could* hold, in GB, if known. Most of the media were plain or generic on the front, giving no indication as to the capacity.
- Format: The format of the medium; in this case, it was either CD or DVD (or some variation such as DVD+R or CD-R).
- Manufacturer: The manufacturer of the medium, if known.
- Software: The software used to create the file, if known.
- Type of Files: The extension of the files located on the media. Since most were DVDs, I had a preponderance of .bup, .ifo, and .vob files. I also had .jpg and .ppt, and a few .cda on the CDs.
- School: This will contain the words "Lower" (Elementary), "Middle" (Junior High), and "Upper" (High School), for extra metadata in the future. Most of the collection surveyed contained Lower School materials.
- Notes: This is where I could put any additional information that did not fit in the other fields. I usually noted when I had a physical program of the event attached to the case.
- Date of Survey: I put in the date that the medium was inventoried.
- Photograph of medium: This is where I will put the photograph of the individual medium, as suggested by the OCLC research report, "You've Got to Walk Before You Can Run: First Steps for Managing Born-Digital Content Received on Physical Media" (August 2012) to use as placeholders or separator sheets within the proper content collection of the item.

For each item, I removed it from the box and placed an alkaline bookmark in its place. None of these items have a unique number, and I will be reprocessing many collections in the next couple of years. As a temporary measure, I cut up a few blank, alkaline papers into small rectangles, where I noted the record group, the box number, and a new item number in pencil. This small label, simply placed inside the case, will stay with the item until it is reprocessed and given a permanent number.

After creating the temporary label, I placed it beside the medium and took a photograph of both. I then removed the CD or DVD from its case, and placed it into my computer. I took care not to start the files, and simply reviewed the directory for the date of the files and the type of files. I then put it back in its case with its new temporary label inside (but not affixed), and placed it back in its proper place in the box. I timed several of these transactions, and found that on average, I spent about two minutes with

each. I imagine that the more files on the item, the longer I might spend on each for the inventory, but having a minimum two-minute mark will help with estimating how long it might take to finish this and other collection inventories.

Findings

I was only able to spend ten hours on this project as I am in the middle of packing up the entire archives collection for a major renovation to the library and archives. The vast majority of my day-to-day duties are in support of that activity. In this short time, however, I was able to survey five boxes of CDs and DVDs.

Box Number	**No. of CDs**	***Av. CD Size***	**No. of DVDs**	***Av. DVD Size***	**Total Media**	**Total GB in Box**	***Average Media Size***
1	17	*.04*	24	*1.46*	41	35.85	*.85 GB*
2	4	*.05*	30	*1.08*	34	32.64	*.96 GB*
3	3	*.45*	27	*1.85*	30	51.29	*1.71 GB*
4	4	*.25*	28	*1.63*	32	46.66	*1.46 GB*
5	1	*.28*	32	*1.71*	33	54.99	*1.67 GB*
Totals	**29**	***.21***	**141**	***1.55***	**170**	**221.43**	***1.33 GB***

In this collection there are a total of eleven boxes of CDs and DVDs. Extrapolating from this, I counted the number of removable media in each box, and estimated the total GB in each box. I will be curious as to how close these numbers are when I finish the inventory:

Box Number	**No. of CDs**	***Av. CD Size***	**No. of DVDs**	***Av. DVD Size***	**Total Media**	***Total GB in Box***	***Average Media Size***
6	5	*.21*	28	*1.55*	33	*44.45*	*1.35 GB*
7	10	*.21*	25	*1.55*	35	*40.85*	*1.17 GB*
8	7	*.21*	27	*1.55*	34	*43.32*	*1.27 GB*
9	21	*.21*	19	*1.55*	40	*33.86*	*.85 GB*
10	15	*.21*	23	*1.55*	38	*38.80*	*1.02 GB*
11	0	*.21*	10	*1.55*	10	*15.50*	*1.55 GB*
Totals	**58**		**132**		**190**	***216.78***	***1.44 GB***

Observations

As a Lone Arranger, I found that the most challenging part of the inventory was trying to find time to conduct the actual inventory in-between all the other activities that I undertake. However, at the same time, I am now appreciative of having a set of procedures and a spreadsheet in place to capture the information about the content and size of the media, which I will be inserting into our soon-to-be-installed Archives Management System. The first part of the inventory took longer per box as I was working out the kinks in the procedures, but by box five there was a clear routine. I am sure the next six boxes will not take as long to inventory.

Plans for the Future

I plan to take these records with me to the new temporary space and can thus finish the inventory of this particular collection, and hopefully other accessioned electronic records, by the end of the 2013-2014 school year. The electronic format is becoming a very important part of our collections. One major example is that I have been in discussions with the faculty member who serves as an advisor to our on-campus television station. A quick survey of files on his three (!) computers come to a total of 406.48 GB, and the file formats included .mov, .flv, .m4, .m4v, and .wav. This covers two academic years, and on average there are 200 events on campus. Thus, we can estimate that for each event, we will need 1 GB of storage space for the videos. We are working with the IT Department to get this backed up, and then we will work together to accession a copy of each event (both the raw footage and the edited version) into the archives.

The hardware that I will need to obtain for the upcoming year includes a write-blocker and a clean, dedicated computer, and I will need more experience in creating checksums and imaging hard disks. In the new permanent space, I will create a section in the Processing Room for ingesting electronic records. Eventually, I would like to move to a more automated system and will work towards that end.

Conclusion

I am glad that I took part in this initiative, as it did give me a “jump start” into more carefully reviewing the electronic files that I currently have in the collection. This has given me more of a formal working structure to use within the school archives, and one I feel confident in using it moving forward.