

When I took my position as the first “official” University Archivist at Ferris State University in 2002, one of the challenges and benefits of my job was that I was not inheriting a lot of legacy policies and procedures. Among my early projects was the implementation of a records management program which gave me the mandate to collect historical records. Because of my situation where records were not mandated to be turned over in earlier years, I do not have some of the problems of older physical media such as 5.25 disks. This played out in the digital records survey which I did.

The Jump In Survey provided me with the motivation in a small shop to prioritize a project which had been on the list for the last several years. I began considering my survey in January with the listserv message about whether they were considering legacy formats such as tape. I chose to focus solely on digital content. I was less strict about whether material was born-digital. I was greatly aided by the Library of Congress’s Digital Preservation Outreach and Education program (<http://www.digitalpreservation.gov/education/>) I am also indebted to Kevin Driedger of the Library of Michigan for his webinar series entitled “Introduction to Digital Preservation: Managing Content Over Time.” As one of the DPOE trainers, he provided helpful concrete steps in the detail required within survey fields.

By April 29, 2013, I had identified over 150 groups of records which ranged from one donation with 4 boxes filled with cds (623 cds) to a single tiff photo file on a flash drive. I separated out the 15 survey items which were digitization projects of other formats. I defined physical media to include: 3.5 disks, cds, dvds, flash drives, portable hard drives, zip disks, and my own hard drive where some materials emailed to me end up. The initial survey identified materials in one of two ways- those tagged as electronic records in the accessions process and then reviewed, and a full review of cds/dvds/drives which are isolated into a separate storage area with better climate control. Then the data was de-duplicated. Some of the items I had identified in larger groups from Accessions (such as 6 cds 1998-2005, Advancement events), ended up being smaller line items on the survey. This is because a cd may have run under a different operating system, or may have been a completely different format (.jpg vs. dvd format). An added bonus to this part of the survey was a greater control of what was in the separate area where patrons may be looking for a very specific item. And I now have a concrete set of data. For example, I can also now cluster the items and more easily quantify what types of space it would take to migrate these items to a server environment.

Perhaps the most interesting case for me of the project was the University Photographer series, which I received in 3 different ways. Between 2002 and 2004, I received annual donations of records on a cd. In 2009, I received a portable hard drive which all of the 2001-2008 files from the cds. Then in 2011 when the offices were redone in 2011, they donated all their copies of cds. I was surprised to learn that the numbers were different. In 2002, the Archives received 38 cds of photos in raw formats. In the 2009 donation, there were 43 cds for that same year. A few in the first donation had been split into an AB and CD cd which would add two additional cds. But where are the other 3? The folder structure on the cds were by week and no weeks appear to be missing. In

2003, there was one additional cd in in the 2009 donation and in 2004 they were the same. The drive donation was harder to compare (it was 2.4 terabytes and over 100,000 photos- but separate data per year was not available). Additionally, the box of additional files included 35 cds of greatest hits, which is an additional collection to be processed- and a misc. collection of 100 cds of individual events (sporting events, commencements, receptions, etc.) which will need to be compared to the weekly structure to see if it is duplicative. I've learned that repeat donations at different times do not lead to identical donations and further investigation is required. This fits in with my general acquisitions which is to encourage donations, even if they turn out to be duplicative.

Additionally, I thought I had a fairly good grasp of file structures and formats but I learned something new from this project. Many of the dvds which were video were in a format unable to be identified easily. They were not segmented into files, but rather into a chapter format. This was a format known as dvd particularly common on the campus from 2003-2008. They are not files in the traditional sense and will require separate migration procedures, which has led to a preliminary discussion with a campus unit.

What I accomplished from this project was a sense of accomplishment that I now have some baseline data from which to chart a future course. I've developed a better sense of useful information to include with initial accession records to add to my data set. And I have identified some issues and discrepancies which have led me to review procedures and particular groups of digital records.