Steven Duckworth

Records Management Intern

Drexel University

Jump In, Too/Two Initiative Report

May 2014

The Drexel University Archives and Special Collections houses roughly 2,000 linear feet of records that document the founding, growth, and functions of the University. The Archives also serves as an educational resource, encouraging the use of the records for researchers and the general public. Prior to the addition of a Records Management Archivist to the staff in 2011, the Archives did not consistently describe digital holdings in accession records. After 2011, accession records include basic description of electronic records and digital media. The Jump In, Too/Two project marks the start of a longer-term project to improve preservation of and access to digital files accessioned by the Archives.

We chose to survey the entire holdings of the Archives during this project. The survey process began by looking at accession records in Archivists’ Toolkit (AT) to gain an overall picture of the holdings. After this review, we decided to only examine materials originating after 1990. We believe that little digital material would be found in pre-1990 collections. Additionally, the resources required to preserve digital material from this long ago would likely prove cost-prohibitive to the archives. An exception was made for the papers and collections related to the Microcomputing Program of the 1980s, which required all Drexel students to purchase Macintosh computers and resulted in the creation of original software; these collections were more likely to contain digital storage materials and are also of high research value.

The survey was conducted by Drexel University Records Management Intern Steven Duckworth. He created an online survey to capture the required information and identify the collections that would be surveyed. This survey worked well as anyone with an internet connected device (computer, iPad, smartphone, etc.) could easily fill in the requested information. This allowed additional Archives staff to enter information on materials discovered when processing their collections and avoided redundancy in multiple passes through collections. This data was automatically saved in a Google Docs spreadsheet that was later exported as an Excel file for ease in analysis. The Records Management Archivist provided guidance throughout the process.

The entire survey process took about 120 hours over the course of two months. The initial stages of background research, survey creation, and collection designation took roughly 50-60 hours. The bulk of the physical survey of collections took another 50-60 hours, with additional materials being added as they are discovered.

Overall, we found more data and a wider number and variety of formats than anticipated. The variety of digital tapes was the most surprising and also adds significantly to the total size of the materials. All told, approximately 6.89 TB of data was discovered through the process. Formats discovered and quantities of each are listed in the chart below.

|  |  |
| --- | --- |
| **Media Type** | **Quantity Found** |
| 3½” disk | 202 |
| 5¼” disk | 1 |
| Zip disk | 30 |
| CD | 1,553 |
| DVD | 476 |
| Mini (“pocket”) CD | 44 |
| USB flash drive | 5 |
| Digital Audio Tape (DAT) | 1 |
| Digital Video Tapes (5 different formats) | 108 |

The vast majority of the media found were labeled in some fashion and many are filed with related material (e.g., proof sheets for photographs, directory printouts, published materials). However, some of the label annotations will likely prove too vague to be of much assistance for additional metadata. The materials date from 1986 to 2013.

One challenge discovered was noting the annotations written on each physical item. Due to time constraints and the proposed next-steps of the project, we decided that the notations of every label were not needed. This allowed the collections to be surveyed more quickly (e.g., a box with 219 CDs and 29 DVDs of student theses and dissertations was listed as two entries, rather than 248, by only noting that the majority of discs contained the student’s name and occasionally a paper title or file format designation). Specific annotations and metadata will be added during further phases of the project. Additionally, the specific location of digital media was noted in the survey, but no items were separated from their collections during this phase.

One takeaway from the project is the need to accurately document digital media at the moment of transfer. The current archivists have made advances in this area and newer collections are more accurately described. Another finding of note is the sheer volume of data in digital video formats. This is an important aspect to bear in mind for other digital initiatives. The future of this project will include transferring files to our server, processing collections, and, eventually, providing access to many digital files through Drexel’s institutional repository. We are currently working out guidelines to prioritize the order in which items will be processed as this project continues.