University of Vermont Libraries, Special Collections
Jump In Survey
Chris Burns
Curator of Manuscripts

UVM Special Collections does not have a formal program in place to manage born-digital materials. While we do not have a large volume of born-digital materials in our holdings, we do have some and our rate of acquiring these materials is increasing. Just in the past year, we have acquired the records on a poet's laptop and the email of a poetry publisher. In addition, we are talking with two politicians about their collections, both of which are likely to contain significant born-digital holdings. For this survey, I chose to focus not on our entire holdings but just on the recently acquired records from a poet's laptop. I chose to focus on just this material as it also met an immediate need. The poet's widow was willing to donate the records on the computer, but wanted the laptop returned in a timely fashion. In order to meet this immediate need, I strayed from the survey steps outlined in Ricky Erway's report, looking more in-depth at the contents of one particular set of files as opposed to looking at a higher level at all of our born-digital media.

For the purpose of this survey, the name of the poet is not important. What is important was that the poet used the computer in the final years of his life and that he was a novice computer user. Furthermore, the poet died in 2008 and his widow and her grandchildren also used the computer, during his lifetime and after. What was on the computer was not organized in a systematic fashion, had several file naming issues, was of mixed authorship, and the time/date stamp was set incorrectly. In addition, the email account featured an inbox with only emails received after the poet's death and with only a few subfolders with relevant correspondence between the poet and others. Because we are only adding the records of the poet to the collection, there is some folder and item level appraisal that will have to happen. While this work is not yet complete, we do know what is actually the records of the poet is not a large number of files and does not take up a large amount of disk space.

In capturing the files, we decided to copy them off of the hard drive to a flash drive and then to a server. We do not have a system in place to capture and store forensic disk images, and I ultimately decided that there were a number of additional reasons not to take that step.
-   We did not have the donor's permission to keep records that were not by or from the poet.

- The actual amount of file space taken up by his materials is probably around 50-60 MB. The amount of space taken up by movies and photographs of and by the grandchildren was far greater. Even if we had permission, it seemed a waste of our storage space to create and keep a disc image for the small number of files that would ultimately be kept.
- The poet is featured in some of the image files, but primarily his work was in text and Word files, and in emails. There was little additional information to capture about the programs on the computer or the file structure that lent weight to an argument for capturing a disc image. In addition, the time/date stamp issues also had already comprised some of the "authenticity" of the records.

Here is a breakdown of some of the information about the computer as well as the files that were copied off of the computer.

MacBookPro running MAC OS X Panther 10.3.9.
1.5 GB of records – 2,376 items
I used two software programs to analyze the contents of the files, Disk Inventory X (freely available) and a trial version of DaisyDisk.

Here is a screen shot from Disk Inventory X.
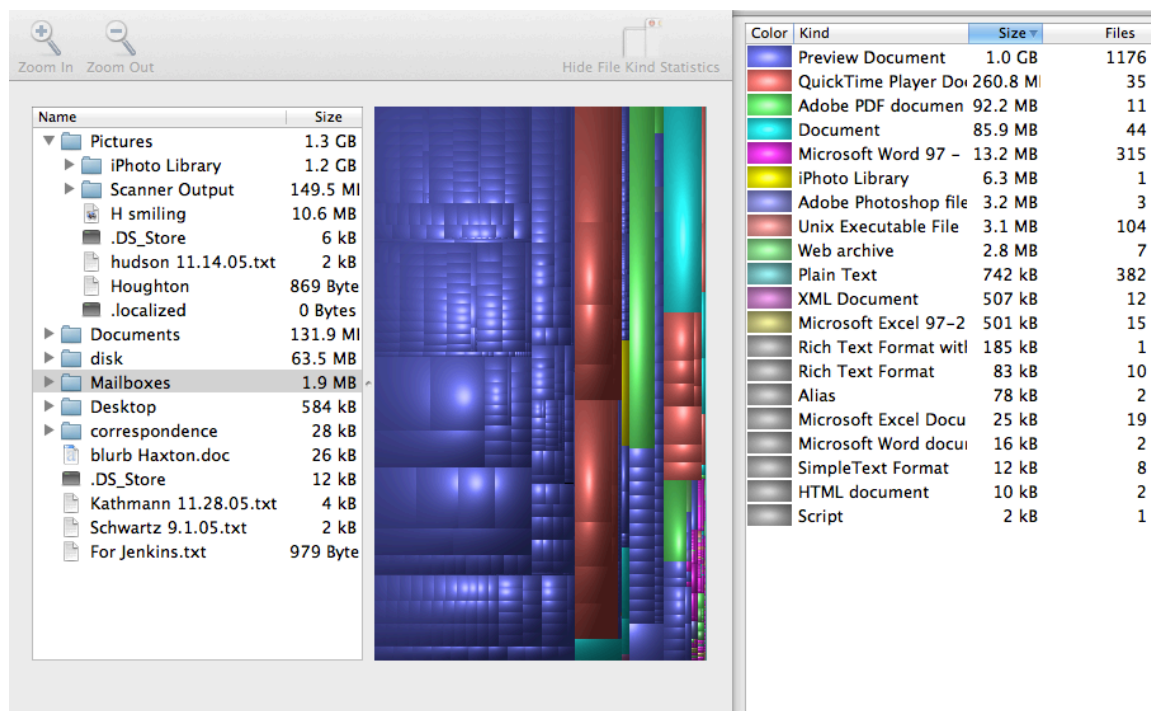
While the frame on the right sorts files by the type of program that my computer wants to open them with, you can get a pretty good general sense of what's there.

Fore example, the 1176 Preview Documents at the top of the list that take up 1.0 GB of disk space are image files. The Quick Time files are moving image files. The useful files of the poet are generally the Word, Text, and Excel files, with some pdf and image files. The email having been pulled out of an Apple Mail program is predominantly unix executable files.

Next Steps
1. Institutionally, we need to investigate putting a repository infrastructure in place so that we have a more trustworthy method of storing and managing these digital objects.
2. We also need to work out an access system for these born-digital objects. For this collection, there may be certain files where it makes sense to print them out. The email files may be a good candidate for this approach, as they are small in number and access to them seems less straightforward than word, text, excel, or image files. For the latter, it may suffice to make a use copy on a portable storage medium and store it with the physical collection.
3. The files from the poet's laptop need folder, in some cases item, level appraisal before they can be made available to researchers.
4. We need to put in place standard language in our deed of gift for dealing with born-digital materials in manuscript acquisitions. We have begun work on this step, working from existing language already created at Duke University.