Digital Forensics: A Floppy Pilot Project @ the University of Tennessee Libraries

Kris Bronstad & Alesha Shumar

How can an institution with no dedicated technical staff begin to address migration of born-digital materials on legacy media?

Background:
University of Tennessee Special Collections has many collections including electronic media from the last 30+ years. UT Libraries was interested in getting ahead of the problem of preserving information stored on obsolete media through the use of digital forensics.

We are a small department with limited resources/staff to assist with the project.

Given our needs and limitations, we decided on an in-house option to retrieve/migrate media through the Forensic Recovery of Evidence Device or (FRED).

Research Method:
In our pilot project using FRED, we turned our attention to a small set of 45 floppy disks documenting the work of poet and literary scholar Peggy Bach (1929-1996). We followed best practices and protocols from our training and utilized archives-specific digital forensics guidelines provided by the AIMS (An Inter-Institutional Model for Scholarship) project; we also studied and modified workflows used by the University of Virginia and other archival institutions. We followed the forensic process, which can be broken down into the following steps:

1. Identify Evidence
2. Make Forensic Duplicate
3. Process Media
4. Report
5. Archive

Questions to ask when beginning small-shop digital forensics practice:

-Proprietary and costly tools (FRED/FTK) v. open source (BitCurator)? What resources do you have? How much time will you need from tech people?

-Who will be responsible for the maintenance of any equipment/software?

-The workflow details: To photograph media or not? When and what to scan for viruses ?

-Keeping the disk image is essential. But what else should you keep in terms of files and metadata? How will you organize it?

-Do you have a system or workflow in place to ingest, manage and disseminate digital files?

-Do you have digital preservation policies/road map in place?

IDENTIFY EVIDENCE — **1**

Survey

The Peggy Bach Floppy Disk Collection contains 44 5.25 disks and one 3.5 disk, with documents created in the early 1990s. Here's what we needed to identify our evidence.

HARDWARE
-External 3.5 drive
-External 5.25 drive
(TEAC to work with controller)
-Drive controller for 5.25 drive
-Dedicated, offline workstation

SOFTWARE
-FTK Imager
-Disk Image & Browse software

METADATA SPREADSHEET
to collect information about the disks, including:
-Author Name (Peggy Bach [default])
-Disk Number (unique number mentioned above)
-Disk Size (Physical size of disk, ex. 5.25 disks and one 3.5 disk)
-Disk Type (Format. This was especially important for 5.25 disks)
-Written Label ("Europe - Poems"; "DM")
-Image Made Date
-Additional notes (any other notation on the media)
-Disk Errors (rate of)

MAKE FORENSIC DUPLICATE — **2**

Accession

-Use drive, FC5025 drive controller, to access 5.25 disks.
-Write-block by switching notch on disk.
-Used Disk Image and Browse to create .img files of the 5.25 disks (.img being the only image format available via DI&B).
-Use external 3.5 drive and FTK Imager to make image from 3.5 disk.
-Open image files in FTK imager to examine and verify image.
-Enter information into spreadsheet.

PROCESS MEDIA — **3**

Use FTK to process files

-Open disk images as a "case" in FTK
-Run pattern searches for Personally Identifiable Information - eg., social security or credit card numbers. Label any files containing these strings "restricted"
-Compile and label undesirable (slack space, sensitive) files
-Label "access" files
-Export desired files and file list to folder in same directory as accession

REPORT — **4**

Generate report in FTK based on chosen criteria

Name: INT-VIEW
Modified Date: 10/14/1994 11:32:46 AM (1994-10-14 15:32:46 UTC)
Path ms2012017D5.E01/NONAME [FAT12]/[root]/INT-VIEW
Accessed Date: 8/19/2001
File Type: WordPerfect 5.1
Duplicate File: 3
MD5 Hash: f6f8166585a5c725fad935399edb2fd6
SHA256 Hash: 141f648f66080ce6af723b6ad6.....
Actual File: True
Container: False
Object Type: File
Label: Access
Export: HTML

ARCHIVE — **5**

Ensure preservation and access

LibreOffice to open/export documents as PDF/A.
-Store exported PDF/As along with disk images and report (from step 4) on secure server.
-Run checksums on server.
-Create finding aid.
-Associate with appropriate records in DAMS.

Results:
Migrated 74 unique documents on floppies from WordStar and WordPerfect formats to PDF/As in archival storage

Generated report with checksum and provenance information, with varying degrees of success in accurate metadata on original files (e.g., date information). Kept report with disk image on secure server

Our experience proved we could migrate the files on these disks to more stable formats, and in this instance bit rot was not a problem.

But…Tools are only part of the solution:
We are still left with many unanswered questions about appraisal, privacy, security and long-term management..

THE UNIVERSITY of TENNESSEE KNOXVILLE

WORKS CONSULTED

The following is a rough list of sources we consulted while trying to get a handle on digital forensics in our institution. We also received invaluable help and documents from Gretchen Gueguen, formerly with the University of Virginia. Also of help were folks at CurateCamp 2013 (http://wiki.curatecamp.org/index.php/CURATEcamp_SAA_2013_Notes)

Aims Working Group. *AIMS Born-Digital Collections: An Inter-Institutional Model for Stewardship*, 2012.

"Born-Digital Archives Program: Forensics Workflow Documentation." Accessed August 18, 2014. https://sites.google.com/site/workflowdocumentation/.

Becker, Devin, and Collier Nogues. "Saving-Over, Over-Saving, and the Future Mess of Writers' Digital Archives: A Survey Report on the Personal Digital Archiving Practices of Emerging Writers." *American Archivist* 75, no. 2 (2012): 482–513.

Carroll, Laura, Erika Farr, Peter Hornsby, and Ben Ranker. "A Comprehensive Approach to Born-Digital Archives." *Archivaria* 72, no. 72 (2011). http://journals.sfu.ca/archivar/index.php/archivaria/article/viewArticle/13360.

Chassanoff, Alexandra, Kam Woods, and Christopher A. Lee. "MAPPING DIGITAL FORENSICS METADATA TO PRESERVATION EVENTS USING BITCURATOR." Poster, n.d. http://www.bitcurator.net/wp-content/uploads/2013/08/SAA-researchforum-chassanoffwoodslee.pdf.

Douglas, Jennifer Lynn. "Archiving Authors: Rethinking the Analysis and Representation of Personal Archives," 2013. https://tspace.library.utoronto.ca/handle/1807/35808.

Duranti, Luciana. "From Digital Diplomatics to Digital Records Forensics." *Archivaria* 68, no. 68 (2010). http://journals.sfu.ca/archivar/index.php/archivaria/article/viewArticle/13229.

Forstrom, Michael. "Managing Electronic Records in Manuscript Collections: A Case Study from the Beinecke Rare Book and Manuscript Library." *American Archivist* 72, no. 2 (2009): 460–77.

Fox, Robert. "Forensics of Digital Librarianship." *OCLC Systems & Services* 27, no. 4 (2011): 264–71.

Garfinkel, Simson. "Digital Forensics XML and the DFXML Toolset." *Digital Investigation* 8, no. 3 (2012): 161–74.

Gengenbach, Martin J. "'THE WAY WE DO IT HERE': MAPPING DIGITAL FORENSICS WORKFLOWS IN COLLECTING INSTITUTIONS," 2012. http://digitalcurationexchange.org/system/files/gengenbach-forensic-workflows-2012.pdf.

Glisson, William Bradley. “Use of Computer Forensics in the Digital Curation of Removable Media.” *DigCCurr*, 2009, 110–11.

Guerrero, Miriely. “Digital Curation-Papers-Removable Media and the Use of Digital Forensics,” 2012. http://deepblue.lib.umich.edu/handle/2027.42/96441.

Idiot’s Guide to FTK Imager. University of Hull Born-Digital Archives, n.d. http://www.hullhistorycentre.org.uk/discover/pdf/Idiot%27s%20Guide%203%20-%20FTK%20Imager.pdf.

John, Jeremy Leighton. “Adapting Existing Technologies for Digitally Archiving Personal Lives.” In *Ponencia Presentada En Las Actas de Fifth International Conference on Preservation of Digital Objects*, 48–55, 2008. http://www.bl.uk/ipres2008/ipres2008-proceedings.pdf#page=57.

———. “Digital Forensics and Preservation.” *Digital Preservation Coalition*, 2012. http://www.dpconline.org/component/docman/doc_download/810-dpctw12-03.pdf.

Kiehne, Thomas, Vivian Spoliansky, and Catherine Stollar. “From Floppies to Repository: A Transition of Bits,” May 2005. http://thomas.kiehnefamily.us/from_floppies_to_repository_a_transition_of_bits.

Kirschenbaum, Matthew G., Richard Ovenden, Gabriela Redwine, and Rachel Donahue. *Digital Forensics and Born-Digital Content in Cultural Heritage Collections*. Council on Library and Information Resources, 2010. http://mith.umd.edu/wpcontent/uploads/whitepaper_borndigital.pdf.

Kirschenbaum, Matthew, Christopher A. Lee, Kam Woods, and Alexandra Chassanoff. “From Bitstreams to Heritage: Putting Digital Forensics into Practice in Collecting Institutions,” 2013. http://drum.lib.umd.edu/handle/1903/14736.

Knight, Gareth. “The Forensic Curator: Digital Forensics as a Solution to Addressing the Curatorial Challenges Posed by Personal Digital Archives.” *International Journal of Digital Curation* 7, no. 2 (October 23, 2012): 40–63. doi:10.2218/ijdc.v7i2.228.

Lee, Christopher A., and Kam Woods. “Digital Acquisition Learning Laboratory: A White Paper.” *School of Information and Library Science, University of North Carolina at Chapel Hill. Http://www. Ils. Unc. Edu/callee/dall-White-Paper. Pdf*, 2011. http://www.digpres.com/publications/dall-white-paper.pdf.

Levi, Charles. “Five Hundred 5.25-Inch Discs and One (Finicky) Machine: A Report on a Legacy E-Records Pilot Project at the Archives of Ontario.” *Archivaria* 72, no. 72 (2011). http://journals.sfu.ca/archivar/index.php/archivaria/article/viewArticle/13365.

Ross, Seamus, and Ann Gow. “Digital Archaeology: Rescuing Neglected and Damaged Data Resources: A JISC/NPO Study within the Electronic Libraries (eLib) Programme on the Preservation of Electronic Materials,” 1999. http://opus.bath.ac.uk/35447/.

“Stanford FTK to Hypatia Object Mapping.” Wiki, n.d. https://wiki.duraspace.org/display/HYPAT/Stanford+FTK+to+Hypatia+object+mapping.

Sternfeld, Joshua. “Archival Theory and Digital Historiography: Selection, Search, and Metadata as Archival Processes for Assessing Historical Contextualization.” *American Archivist* 74, no. 2 (2011): 544–75.

Thomas, S., R. Gittens, J. Martin, and F. Baker. *Paradigm: Workbook on Personal Digital Archives*. Bodleian Library, University of Oxford: Oxford, UK, 2007.

Thomas, Susan, and Janette Martin. “Using the Papers of Contemporary British Politicians as a Testbed for the Preservation of Digital Personal Archives.” *Journal of the Society of Archivists* 27, no. 1 (2006): 29–56.

Weber, Nicholas M., Carole L. Palmer, and Tiffany C. Chao. “Current Trends and Future Directions in Data Curation Research and Education.” *Journal of Web Librarianship* 6, no. 4 (2012): 305–20.

Wilsey, Laura, Rebecca Skirvin, Peter Chan, and Glynn Edwards. “Capturing and Processing Born-Digital Files in the STOP AIDS Project Records: A Case Study.” *Journal of Western Archives* 4, no. 1 (2013): 1.

Woods, Kam, Alexandra Chassanoff, and Christopher A. Lee. “Managing and Transforming Digital Forensics Metadata for Digital Collections.” Accessed January 15, 2014. http://purl.pt/24107/1/iPres2013_PDF/Managing%20and%20Transforming%20Digital%20Forensics%20Metadata%20for%20Digital%20Collections.pdf.

Xie, Sherry L. “Building Foundations for Digital Records Forensics: A Comparative Study of the Concept of Reproduction in Digital Records Management and Digital Forensics.” *American Archivist* 74, no. 2 (2011): 576–99.

.