Appendix G: Technical Evaluation and Use

1. AccessData FTK3.3

Purpose

AccessData FTK (Forensic ToolKit) generates summary information on a collection (single floppy disk or a collection with floppy, zip, CD and hard disks) of files and provides different views of files, sophisticated search, bookmarking and labeling functions.

Use of Software

This software can be used in the accessioning, arrangement and description phases of the AIMS framework for born digital material.

Key Functionality

1. Summary information of a collection (single floppy disk or a collection with floppy, zip, CD and hard disks) by file extension, file category, file status and Email message.
    a. Summarizes files by their extensions, such as .TXT, .JPG, and .DOC and lists them in a tree view.
    b. Summarizes files by type, such as a word processing document, graphic, email, executable (program file), or folder, and lists them in a tree view.
    c. Summarizes files by status such as deleted files, duplicate items, and encrypted files, etc. and lists them in a tree view.
    d. Provides message counts of Emails in AOL DBX, PST (Outlook email), NSF (Lotus Notes email), MBOX (Thunderbird, Netscape, Eudora, etc. email) formats.
2. Different views of files, including explorer tree, file list, file content and thumbnail.
    a. Explorer Tree View lists directory structure of disks/folders, similar to the way one would view directory structure in Windows Explorer in original order.
    b. File List View displays files and pertinent information about files, such as filename, file path, file type, file formats (identified by FTK) and checksums (generated by FTK), etc.
    c. File Content View displays files as Hex (hexadecimal representation), Text (in different character encoding scheme such as ASCII, Chinese Traditional (Plane 1), EBCDIC (37 United States), Mac OS Roman, Windows 1252 (Latin I), etc.), Filtered (file's text created during indexing), and

Natural (file's contents as it would appear normally) formats. The "Natural" format uses the Oracle Stellent INSO filters for viewing hundreds of file formats without the native application being installed.

d. Thumbnail View displays graphics files in thumbnails in photo-album style.

3. Index Search, Pattern Search and Fuzzy Hashing

   a. Index search compares search terms to an index file containing discrete words or number strings found in a collection. Index search options include: "Stemming Words" that contain the same root, such as raise and raising, "Phonic Words" that sound the same, such as raise and raze, "Synonym Words" that have similar meanings, such as raise and lift, "Fuzzy Words" that have similar spellings, such as raise and raize.

   b. Pattern Search includes many predefined regular expressions for searching, including the following: U.S. Social Security Numbers, IP Addresses, U.S. Phone Numbers, Visa and MasterCard Numbers, U.K. Phone Numbers, and Computer Hardware MAC Addresses, etc. Users can also create their own pattern.

   c. Fuzzy Hashing is a tool which provides the ability to compare two distinctly different files and determine a fundamental level of similarity._Traditional cryptographic hashes (MD5, SHA-1, SHA-256, etc.) are useful to quickly identify known data, to indicate which files are identical. However, these types of hashes cannot indicate how closely two non-identical files match. Fuzzy hashing identifies similarity by a score from 0-100. A score of 100 would indicate that the files are close to identical. Alternatively a score of 0 would indicate no meaningful common sequence of data between the two files.

4. Provide Labeling and Bookmarking

   a. Labels give you a method of grouping files in a completely user defined way.

   b. A bookmark is a group of files that users want to reference. These are user-created and the list is stored for later reference, and for use in the report output. Users can create as many bookmarks as needed. The main difference labels and bookmarks is that bookmarks can be nested within other bookmarks and labels do not have such feature. This makes bookmark a good choice for representing "series" and "subseries". Install

Verdict

FTK is the only software I know to perform all the functionalities mentioned above in a totally integrated environment.

Further Information

http://accessdata.com/products/computer-forensics/ftk

Questions:

Contact Peter Chan, digital archivist, Stanford University Libraries, at pchan3@stanford.edu.

2. AccessData FTK Imager 3.0

Purpose of software

FTK Imager is a data preview and imaging tool.

Use of Software

The software can be used to create forensic or logical (deleted files, unallocated space not included) images of local hard drives, floppy diskettes, Zip disks, CDs, and DVDs, entire folders, or individual files from various places within the media in the accessioning phase of the AIMS framework.

Key Functionality

FTK Imager is a data preview and imaging tool created by AccessData Corp. With FTK Imager, you can:

- Create forensic images of local hard drives, floppy diskettes, Zip disks, CDs, and DVDs, entire folders, or individual files from various places within the media.
- Create logical images of the contents of folders. The image created will include only logical files. It will not include deleted files, unallocated space, etc. It does not store sector information.
- Preview files and folders on local hard drives, network drives, floppy diskettes, Zip disks, CDs, and DVDs.
- Preview the contents of forensic images stored on the local machine or on a network drive.
- Mount an image for a read-only view that leverages Windows Explorer to see the content of the image exactly as the user saw it on the original drive.
- Export files and folders from forensic images.
- See and recover files that have been deleted from the Recycle Bin, but have not yet been overwritten on the drive.
- Create hashes of files using either of the two hash functions available in FTK Imager: Message Digest 5 (MD5) and Secure Hash Algorithm (SHA-1).
- Generate hash reports for regular files and disk images (including files inside disk images) that you can later use as a benchmark to prove the integrity of your case evidence. When a full drive is imaged, a hash generated by FTK Imager can be used to verify that the image hash and the drive hash match after the image is created, and that the image has remained unchanged since acquisition.
- Encrypt data during export to an image.

Identified and Analyzed File Systems

- Microsoft: FAT 12, FAT 16, FAT 32, NTFS, exFAT
- Apple: HFS, HFS+

- Linux: Ext2FS, Ext3FS, Ext4FS
- Others: ReiserFS 3, VXFS, CDFS

Identified and Analyzed CD and DVD File Systems and Formats

Alcohol (*.mds), IsoBuster CUE, PlexTools (*.pxi), CloneCD (*.ccd), Nero (*.nrg), Roxio (*.cif), ISO, Pinnacle (*.pdi), Virtual CD (*.vc4), CD-RW, VCD, CD-ROM, DVD+MRW, DVCD, DVD-RW, DVD-VFR, DVD+RW Dual Layer, DVD-VR, BD-R SRM-POW, BD-R DL, BD-R SRM, CloneCD (*.ccd), HD DVD-R, HD DVD-RW DL, SVCD, HD DVD, HD DVD-RW, DVD-RAM, CD-ROM XA, CD-MRW, DVD+VR, DVD+R, DVD+R Dual Layer, BD-RE, DVD-VRW, BD-ROM, HD DVD-R DL, BD-R RRM, BDAV, Pinnacle (*.pdi), HD DVD-RAM, ISO, CD-R, Virtual CD (*.vc4), SACD, DVD+RW, DVD-ROM, VD-R, DVD-VM, DVD-R Dual Layer, DVD+VRW, BD-R SRM +POW

Verdict:

The ability to create logical image is extremely important when a bit-by-bit forensic image is not allowed.

Further information

FTK is a proprietary software but is free and can be downloaded at http://www.accessdata.com/downloads.html.

Questions:

Contact Peter Chan, digital archivist, Stanford University Libraries, at pchan3@stanford.edu.

3. Comparison of 5.25" Floppy Disk Drive Solutions

Purpose

Most modern computers do not have the hardware needed to read 5.25 floppy diskettes. This review compare 4 solutions to connect a 5.25 floppy drive to your existing computers or a new one built from the motherboard suggested. Catweasel is an expansion card to be inserted in the PCI slot of your existing PC. Both KryoFlux and FC5025 are bare circuit board with an USB interface for connecting to the USB port of your existing PCs. Gigabyte GA-880GA-UD3H is a motherboard with a floppy disk controller which allows you to connect your 5.25 inch floppy drive.

Key features for each solution:

	Catweasel	KryoFlux	FC5025 USB 5.25" floppy controller	Gigabyte GA-880GA-UD3H
Hardware	PCI expansion card	Printed circuit board with USB interface	Printed circuit board with USB interface	Motherboard
Included Software	IMAGE (GUI) Command Line Tools	DTC (command line) GUI	Disk Image and Browse (GUI) Command Line Tools	Nil (the following is based on FTK Imager)
Operating System requirement	Windows XP; works on Linux with additional software	Windows XP,Vista (32-bit) Windows 7 (32/64-bit) Mac OS X Linux	Linux x86 2.6.24 Mac OS X PPC 10.4.11 Mac OS X Intel 10.6.4 Windows XP SP3 32-bit Windows 7 (32/64-bit)	Linux x86 Windows XP,Vista (32-bit) Windows 7 (32/64-bit)

	Catweasel	KryoFlux	FC5025 USB 5.25" floppy controller	Gigabyte GA-880GA-UD3H
Supported disk type / File system	PC-formats (180K up to 1440K) Amiga DD and HD (also 5,25" formats) Atari 9, 10 and 11 sector disks Macintosh 720K, 800K, 1440K (DD, GCR, HD) Commodore 1541, 1571, 1581 (C64, C128 and 3.5" C-64 disks) XTRA High density with 2380KByte per disk Nintendo backup station 1600KB format Atari 800XL (all MFM formats, FM under development) Apple IIe disks (Apple DOS 3.3 and up)	KryoFlux supports dumping any floppy disk to "stream files", which contain the low level flux transition information present on a disk. It also supports output of a range of common "sector dumps" to allow you to use your dumped images right away in your favorite emulator. The currently supported disk image formats are: KryoFlux stream files CT Raw image, 84 tracks, DS, DD, 300, MFM FM sector image, 40/80+ tracks, SS/DS, DD/HD, 300, FM FM XFD, Atari 8-bit MFM sector image, 40/80+ tracks, SS/DS, DD/HD, 300, MFM MFM XFD, Atari 8-bit AmigaDOS sector image, 80+ tracks, DS, DD/HD, 300, MFM CBM DOS sector image, 35+ tracks, SS, DD, 300, GCR Apple DOS 3.2 sector image, 35+ tracks, SS, DD, 300, GCR Apple DOS 3.3+ sector image, 35+ tracks, SS, DD, 300, GCR DSK, DOS 3.3 interleave Apple DOS 400K/800K sector image, 80+ tracks, SS/DS, DD, CLV, GCR	Apple DOS 3.2 (13-sector) Apple DOS 3.3 (16-sector) Apple ProDOS Atari 810 Calcomp Vistagraphics 4500 Commodore 1541 Kaypro 2 CP/M 2.2 Kaypro 4 CP/M 2.2 MS-DOS North Star MDS-A-D TI-99/4A	Microsoft: FAT 12, FAT 16, FAT 32, NTFS, exFAT Apple: HFS, HFS+ Linux: Ext2FS, Ext3FS, Ext4FS Others: ReiserFS 3, VXFS, CDFS
Disk image output format	Raw (plain, .bin, .d64, .d71, .d81, .adf, .xfd), .d64 with error information, .atr	Raw	Raw (.d64, .img, .po, .do, .dsk)	Raw (dd), SMART, E01, AFF
Directory listings of all files in the image	No	No	No (only browse)	Yes
Log file (date, time, checksums, actions, results)	No	No	Partial (only success/failure and bad sectors)	Yes (checksum of both original disk and the disk image)
Filesystem browse (appraisal)	No	No	ProDOS, MS-DOS and Kaypro disks	MS-DOS
Integrate with QuickView Plus (appraisal)	No	No	No; can import disk images into FTK Imager	Yes
Cost	USD120	USD3,000 (non-personal edition price) Euro94.95 (Personal Edition Advanced)	USD55.25 (additional USD48 for disk drive external enclosure and power supply)	USD120

Verdict

Building a pc based on the Gigabyte GA-880GA-UD3H motherboard is only solution mentioned above to allow you to use QuickView Plus and your antivirus software to see/scan the files in the floppy diskette without creating a disk image. The other 3 solutions require the user to create a disk image of the diskette and extract the disk image in order to see the files using QuickView Plus or to scan the files with your antivirus software in a floppy diskette. Anyway, all four solutions provide unique features and users have to match the solutions to their problems.

Further information:

FC5025: http://www.deviceside.com/fc5025.html

Catweasel: http://www.jschoenfeld.com/products/catweasel_e.htm

Gigabyte GA-880GA-UD3H: http://www.gigabyte.us/products/product-page.aspx?pid=3758#ov

Kryoflux: http://www.kryoflux.com/

Contact Peter Chan, Digital Archivist at Stanford University Libraries, at pchan3@stanford.edu for questions.

I would like to thank Mark Matienzo for supplying the information on FC5025 and commenting this review.

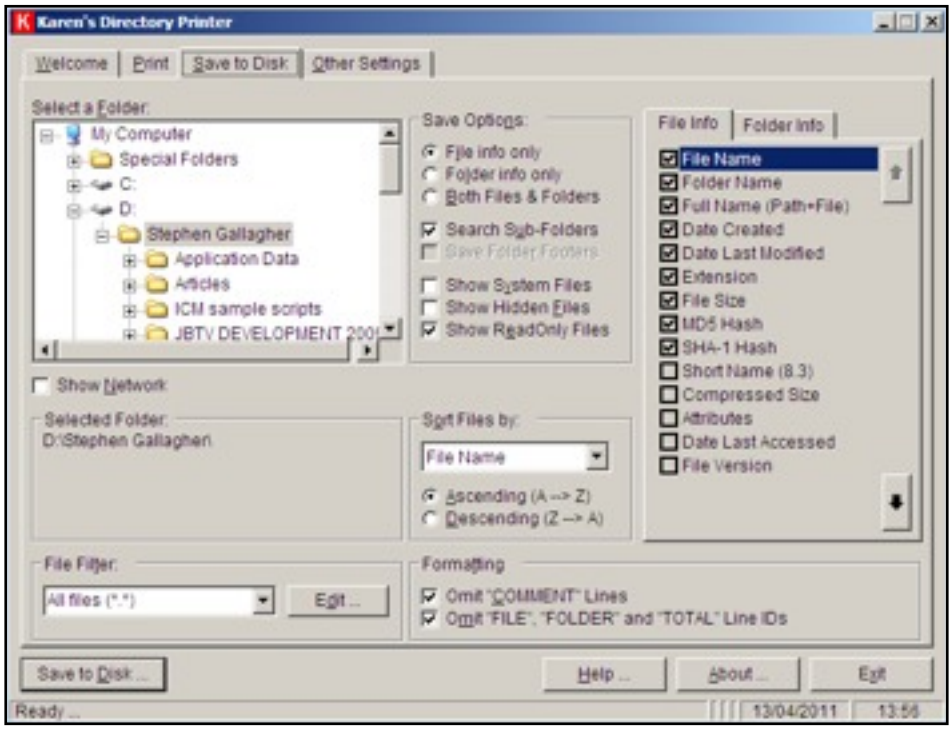
4. Karen's Directory Printer (v.5.3.2)

Purpose

Karen's Directory Printer is a freeware tool that can capture key details about each file and folder in an accession. It cannot be used to capture details from disc images.

Use of Software

We were first drawn to the potential of this software by accounts of its use by West Yorkshire Archives Service. This software can be used to create a manifest of the files that have been transferred to us before we undertake any processing, that is with-in the accession phase of the AIMS framework, and can be used in conjunction with write-blockers.



The information can include file and folder name, the full path, the size of the file (in kb), the date the file was created and last modified. For folders it can record the folder name, the number of files, the number of sub-folders and total size of the folder. This data can be saved as a text file, using .csv format, that can be easily imported into MS Excel and then manipulated in a number of ways including identifying duplicate items –where the checksums match – irrespective of the filenames.

It is possible to create file and folder information at the same time, but having two separate manifests makes data analysis and the potential for re-use easier. The software remembers the settings between uses which make subsequent re-use easier and quicker.

Key Functionality

One of the most useful features of this software is that it can create both MD5 and SHA1 checksums and these can be compared with checksums generated through other tools like FTK Imager. The ability to capture file extension also provides an indication of possible file types to be encountered - this can then be verified through the use of DROID.

The ability to view key information about the folders - and in particular the number of files and its size. This information can be used to provide a useful perspective of the entire collection and may suggest particular folders for appraisal and this can be documented in the processing plan.

Verdict

It is possible to create file and folder information at the same time, but having two separate manifests makes using the data in further tasks easier. Indeed this is one factor that has meant we have decided to keep using this software despite similar functionality being offered by FTK Imager.

Although its use involves another piece of software in our workflow we felt the tool was simple and easy to use and feel confident in suggesting its use by depositors who may wish to create a list of files that they intend to transfer.

Further Information

http://www.karenware.com/powertools/ptdirprn.asp

5. Curator's Workbench

Purpose

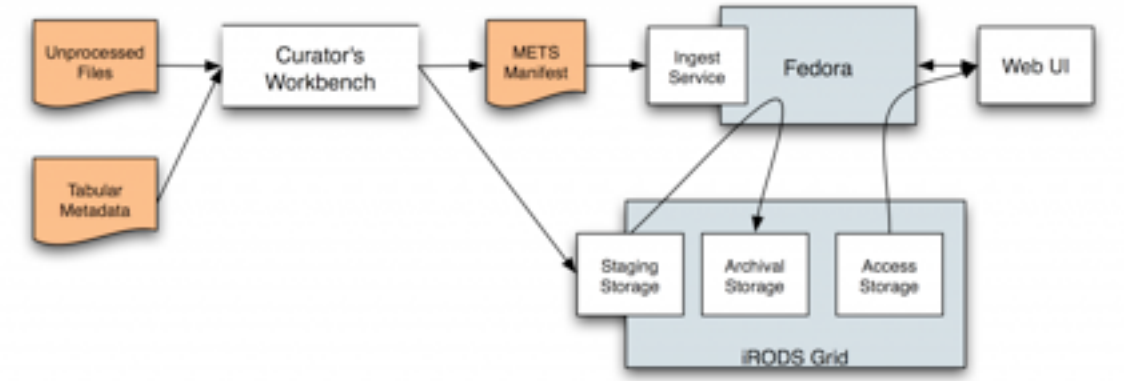
The Curators Workbench is an open source tool designed to assist with the accession, arrangement, description and staging of digital objects. The tool was developed as part of the Carolina Digital Repository over the summer of 2010 by developer Greg Jensen and Erin O'Meara Electronic Records Archivist.

The tool is still in active development with version 3.0 being due for release in early October 2011. It has been deliberately designed to have a modular framework to allow other institutions to use and extend the tool according to their particular institutional requirements. In the summer of 2011 Greg and Erin hosted a number of workshops in the UK as part of their attempts to establish a user community that can actively contribute to the development process.

Use of Software

The tool creates a METS file documenting the processes that have been applied and can create MD5 checksums and unique IDs for each object (UUIDs). MODS descriptive metadata can be mapped to individual objects and folders using the impressive crosswalk feature.

The software requires each accession to be handled as a distinct project which is useful and each project is then built around the METS manifest which tracks the objects and their metadata and is then exported to form the basis of a submission package prior to ingest.



See http://www.lib.unc.edu/blogs/cdr/index.php/2010/12/01/announcing-the-curators-workbench/

Key Functionality

The crosswalk is one of the distinctive features of Curators' Workbench with the crosswalk editor allowing a user to visually map their data with MODS data elements. At present this only supports tab-separated metadata sources but it is planned to be extended to any delimited file and XML sources. The ability to save and then re-use the crosswalk definition allows a user to generate the MODS records. This re-use saves considerable time and effort and in most cases should avoid the need for custom scripts for each data source.

The editor also allows you to add standard text for example a statement relating to copyright as part of the crosswalk process.

The tool also includes a staging area designed to facilitate the processing and ingest of files to your preservation storage environment, critical considering the sheer number of files contained with-in many born-digital archive accessions. The staging area can be configured to your specific storage environment and it can also be used to identify issues prior to forming the submission package.

The tool does claim that you can add descriptive metadata but it was unclear whether this could be applied in batch mode or even whether this conformed to any descriptive standards and it also allows you to view the properties of each file.

Verdict

The tool looks very professional and very polished and in the most part is easy to use. The crosswalk editor does require getting used to but is worth the investment in time and effort.

It is difficult to be too judgemental for a tool that is in such active development, but aspects I would like to see include / explore further are;

1. How easy it is to create suitable metadata to implement the crosswalk from a position of having a batch of born-digital files, which is how many collections will be received
2. Whether a distinction can be made between the original folders and the staged files – with both containing the same files it is easy to forget "where" you are
3. Clarification whether arranging the files is an intellectual process only, as is proposed with the Hypatia tool, before you start renaming, re-arranging and deleting objects

Further Information

Curator's Workbench at UNC that includes links to manuals, screencasts, etc.:
http://www.lib.unc.edu/blogs/cdr/index.php/about-the-curators-workbench/

Curators Workbench wiki:
https://github.com/UNC-Libraries/Curators-Workbench/wiki