Digital Curation: Two Paths

Society of California Archivists
April 28, 2012
Ventura, California

Peter Chan, Digital Archivist, Manuscripts
Department of Special Collections & Digital Library System Services
Agenda

• Accessioning

• Interface between AccessData FTK and In-house Repository

• Forensic Lab

• Future
Accessioning

• Computers, floppy disks, external hard drives?

• Electronic files?
Physical: Media
AT Accession Record

Resources Linked to this accession

<table>
<thead>
<tr>
<th>Resource Identifier</th>
<th>Resource Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1839</td>
<td>Rebecca Solnit papers</td>
</tr>
</tbody>
</table>

Resource Type: Papers

Resource Title: Rebecca Solnit papers

Physical Descriptions

Extent Number: 3

Container Summary:
1 Macintosh PowerBook 160
1 Macintosh Powerbook G3
1 Apple PowerBook G4

Physical Detail

Dimensions

Repository: SPECCOLL

Created: Mar 30, 2012 by pchan3 | Modified: Apr 2, 2012 by pchan3 | Record Number
## File Overview

4/2/2012

**Evidence Groups**
- Ungrouped: 604578

**File Items**
- Evidence Items: 2
- Checked Items: 0
- Unchecked Items: 604578

**File Category**
- Archives: 11601
- Databases: 19
- Documents: 117754
- Email: 60435
- Executable: 1608
### AT Accession Record for Files

**Physical Description**

<table>
<thead>
<tr>
<th>Primary Extent</th>
<th>Alternate Extent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extent Number</td>
<td>gigabytes</td>
</tr>
<tr>
<td>68.54</td>
<td></td>
</tr>
</tbody>
</table>

**Container Summary**

- megabytes: 70,459
- number of files: 60,458
- captured from: 2 portable computers
Interface between AccessData FTK and In-house Repository

- XML FTK report is in fact XML-FO which is designed for printing as PDF
- One programmer described it as a “mess”
- Transformation is needed!!
XML Content File
XML-FO from FTK

```xml
<?xml version="1.0" encoding="UTF-8"?>
<root xmlns:xs="http://www.w3.org/2001/XMLSchema" font-family="arial unicode ms" font-size="10pt"
xmlns:fo="http://www.w3.org/1999/XSL/Format">
<layout-master-set>

  <simple-page-master master-name="IOC">
    <region-body region-name="body" margin-top="0.5in" margin-bottom="1in"
    margin-left="0.5in" margin-right="0.5in"/>
  </simple-page-master>

  <simple-page-master master-name="caseInfoPage">
    <region-body region-name="body" margin-top="0.5in" margin-bottom="1in"
    margin-left="0.5in" margin-right="0.5in"/>
  </simple-page-master>

  <simple-page-master master-name="bookmarksPage">
    <region-body region-name="body" margin-top="0.5in" margin-bottom="1in"
    margin-left="0.5in" margin-right="0.5in"/>
  </simple-page-master>

  <simple-page-master master-name="index">
    <region-body region-name="body" margin-top="0.5in" margin-bottom="1in"
    margin-left="0.5in" margin-right="0.5in" column-count="2" column-gap="2pc"/>
  </simple-page-master>

  <sequence master-reference="IOC">
    <layout-master-set>
  </sequence>
</root>
```
XSLT

```xml
<?xml version="1.0" encoding="UTF-8"?>
<xsl:stylesheet xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
    xmlns:fo="http://www.w3.org/1999/XSL/Format" version="2.0">

  <!-- This is the root level template -->
  <xsl:template match="/">
    <ftk_report>
      <!-- xsl:apply-templates -->
    </ftk_report>
  </xsl:template>

  <xsl:template match="/ftk_report">
    <!-- xsl:apply-templates -->
  </xsl:template>

  <!-- This template matches any <for-page-sequence> that contains "Case Information" The following <for-block> contains the "Series Information" -->
    <xsl:value-of select="substring-after(.,' ')"/>
  </xsl:template>

  <!-- This template matches any <for-table-body> that contains "File Comments" This indicates the beginning of a new file description. -->
    <xsl:value-of select="."/>
  </xsl:template>

  <!-- This template matches <for-table-row> that contains "Name" The following <for-table-cell> contains the value of the "Name" field. -->
  <xsl:template match="/for-table-row[for-table-cell[text()='Name']]/for-table-cell[2]/for-block">
    <xsl:value-of select="."/>
  </xsl:template>
</xsl:stylesheet>
```
Forensic Lab, etc.

Hardware
FRED. 6GB RAM, Intel i7 920 2.67GHz, 1x160 GB 10K RPM HD, 1x750GB HD, 1x1.5TB HD, 1xBlu-Ray Drive, 1x5.25 inch floppy, 1xZip drive, 1xCatweasel PCI card.
FRED-L: portable FRED; 4GB RAM, Intel Core2 E8500 3.16 GHz, 1x250GB HD, 1xDVD drive
Custom built capture station: 8GB RAM, AMD Phenom II X6 2.8GHz, 1x40GB SSD, 1x128GB SSD, 2x1TB HD in RAID 0, 1x5.25 inch floppy drive, 1xZip drive, 1xDVD drive
3.5 inch floppy drive with write protect switch
KryoFlux
Canon EOS T1i Camera with 18-55 mm lens
Photo copy stand
1x SATA & IDE Portable Write Blocker
1x SCSI Portable Write Blocker
1x USB 2 Portable Write Blocker
Apple Power Macintosh 7100/66

Software
AccessData FTK 11 licenses
FTK Imager
Emailchemy
EnCase (require dongle)
QuickView Plus
Mac/FTK ImagerCL 2.9.0_Mac.zip supports mac os 10.5 and 10.6x
Clean install of Oracle

Manuals, etc.
AccessData FTK
FTK Imager
FTK Recognized File Types 7-28-08
FTK Fuzzy Hash
AccessData FTK Imager CLI v2.9
Use Guide for the FC5025 Floppy Disk Controller
T9 Forensic FireWire Bridge Users Guide

Get Hard drives out from Mac
PowerBook G4 Aluminum 12" 867 MHz Hard Drive Replacement
Macintosh PowerBook G3 Hard Drive Replacement
Installing Macintosh PowerBook 165c Hard Drive
Baseline FRED Specifications

- **23 3/4" High, 8 3/8" Wide, 25 1/4" Deep - 80 lbs**
- **ATX Server Case:** 12 x 5¼", 1100W Modular Power Supply
- **i7 Motherboard** with Intel X58 / ICH10R Chipset
- Intel i7 960 CPU (Quad Processor), 3.20 Ghz, 8M Cache, 4.80 GT/s Intel® QPI
- 5 x PCI-Express(x16) and 1 x PCI-Express(x1) Slots
- 12 GB DDR3-1600 Triple Channel Memory
- **Nvidia GT430 PCI-Express Video Card (1GB)** (1 DisplayPort, 1 HDMI, and 1 DVI Port)
- Dual 10/100/1000 Mbs Gigabit Ethernet Network Adapters
- 8 Channel High Definition Audio Controller
- 7 Ports Primary 3.0 Gb/s Serial ATA (SATA) Controller (1 Back Mounted)
- 2 Ports Marvell PCIe SATA 6.0 Gb/s Controller
- 1 PS/2 Port (Keyboard / Mouse Combo)
- 6 USB 3.0/2.0 Ports: 3 Back Mounted, 3 Front Mounted
- 11 USB 2.0/1.x Ports: 8 Back Mounted, 3 Front Mounted (1 Write Blocked)
- 1 FireWire IEEE 1394a (400 MB/s) Ports: 1 Back Mounted
- 3 FireWire IEEE 1394b (800 MB/s) Ports: 1 Back Mounted, 2 Front Mounted (1 Write Blocked)
- **Digital Intelligence UltraBay Hardware Write-Blocker:**
  - Integrated IDE Drive Write Blocker
  - Integrated SATA Drive Write Blocker
  - Integrated SCSI Drive Write Blocker
  - Integrated USB Write Blocker
  - Integrated Firewire IEEE 1394b Write Blocker
- **Digital Intelligence Integrated Forensic Media Card Reader** - One Switchable Read-Only/Read-Write (MSC, MS Pro, SMC, CF, MD, XD, SDC, and MMC Memory Card compatible)
- **1 x 300 Gb 10,000 RPM 3.0 Gb/s SATA Hard Drive in Shock-Mounted Tray** – OS Drive
- **1 x 1.5 Tb 7200 RPM 3.0 Gb/s SATA Hard Drive in Shock-Mounted Tray** – Data Drive
- **2 x Shock Mounted SATA Removable Hard Drive Bay** (IDE Capable)
- **4 x HotSwap Shock Mounted Universal Removable Hard Drive Bays** (IDE/SATA Compatible)
- BD-R/BD-RE/DVD ± RW/CD ± RW Blu-ray Burner Dual-Layer Combo Drive
- Extendable/Retractable Imaging Workshelf with Integrated Ventilation
- **USB 3 1/2" Floppy Drive with Write Protect Switch**
- Wireless 103 Key Microsoft Keyboard
- Wireless Microsoft Intellimouse
- **22" Widescreen LCD Monitor** with Built-in Speakers
Home Made Capture Station
AccessData FTK

Dear Professor Lindahl,

The time (11 to 1 o'clock) will work very well for me. I’ll plan to arrive there at 9:25 AM March 2nd, and then to leave at 3:40 PM, which I hope will be agreeable for your other arrangements. Please give my best to Professor Sell.

Sincerely,

Robert Creeley

P.S. I think it would be useful to have xerox copies of a small selection of my poems for distribution to the audience at the talk. Whatever seems interesting would be fine—and then I could read and say something about them, as well as read a few others written since I’ve come to Finland.
Index Search Options

Search Options
- Stemming
- Phonic
- Synonym
- Fuzzy

Result Options
- Max Files to List: 500
- Max Hits Per File: 100
- Max Words to Return: 65536

Files To Search
- *All Files*

- File Name Pattern: *.doc

- Files Saved Between: 9/13/2008 and 9/13/2009

- Files Created Between: 9/13/2008 and 9/13/2009

- File Size Between: (bytes)

- Save as Default

[OK] [Cancel]
Live Pattern Search

The more complex Live Pattern “Regex” style search can be used to create pattern searches, allowing forensics analysts to search through large quantities of text information for repeating strings of data such as:

- Telephone Numbers
- Social Security Numbers
- Computer IP Addresses
- Credit Card Numbers

In the Live Search tab, click the Pattern tab. Each has different options.

The patterns consist of precise character strings formatted as mathematical-style statements that describe a data pattern such as a credit card or social security number. Pattern searches allow the discovery of data items that conform to the pattern described by the expression, rather than a known and explicitly entered string are looking for.
FTK Bookmarking / Labelling
Outsource
Outsource
Apple PowerBook 160 Specs

The Apple Macintosh PowerBook 160 features a 25 MHz 68030 processor, 4 MB of RAM, either a 40 MB, 80, or 120 MB hard drive, and an internal 1.44 MB floppy drive in a compact portable case with a 9.8" grayscale passive-matrix display.

Most notably, although the built-in display on the PowerBook 160 only supports 4-bit grayscale, it supports 8-bit color on an external monitor and was the first PowerBook to offer such capability.

<table>
<thead>
<tr>
<th>PowerBook 160 Specs</th>
<th>Ports &amp; Connectivity</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Introduction Date:</strong></td>
<td>October 19, 1992</td>
</tr>
<tr>
<td><strong>Discontinued Date:</strong></td>
<td>August 15, 1993</td>
</tr>
<tr>
<td><strong>Processor Speed:</strong></td>
<td>25 MHz</td>
</tr>
<tr>
<td><strong>Processor Type:</strong></td>
<td>68030</td>
</tr>
<tr>
<td><strong>Details:</strong></td>
<td>This model has a 32-bit processor and a 32-bit data path.</td>
</tr>
<tr>
<td><strong>Processor Upgrade:</strong></td>
<td>Soldered</td>
</tr>
<tr>
<td><strong>FPU:</strong></td>
<td>None</td>
</tr>
<tr>
<td><strong>System Bus Speed:</strong></td>
<td>25 MHz</td>
</tr>
<tr>
<td><strong>Cache Bus Speed:</strong></td>
<td>N/A</td>
</tr>
<tr>
<td><strong>ROM/Firmware Type:</strong></td>
<td>Macintosh ROM</td>
</tr>
<tr>
<td><strong>ROM/Firmware Size:</strong></td>
<td>1 MB</td>
</tr>
<tr>
<td><strong>L1 Cache:</strong></td>
<td>0.5k</td>
</tr>
<tr>
<td><strong>L2 Cache:</strong></td>
<td>N/A</td>
</tr>
<tr>
<td><strong>RAM Type:</strong></td>
<td>DRAM Card*</td>
</tr>
<tr>
<td><strong>Min. RAM Speed:</strong></td>
<td>85 ns</td>
</tr>
<tr>
<td><strong>Details:</strong></td>
<td>*Supports a single 85 ns DRAM card.</td>
</tr>
<tr>
<td><strong>Standard RAM:</strong></td>
<td>4 MB</td>
</tr>
<tr>
<td><strong>Maximum RAM:</strong></td>
<td>14 MB*</td>
</tr>
<tr>
<td><strong>Details:</strong></td>
<td>*With 4 MB of onboard memory and a 10 MB memory module.</td>
</tr>
<tr>
<td><strong>Motherboard RAM:</strong></td>
<td>4 MB</td>
</tr>
<tr>
<td><strong>RAM Slots:</strong></td>
<td>1*</td>
</tr>
<tr>
<td><strong>Video Card:</strong></td>
<td>Dedicated*</td>
</tr>
<tr>
<td><strong>VRAM Type:</strong></td>
<td>Onboard</td>
</tr>
<tr>
<td><strong>Details:</strong></td>
<td>*Video function provided by onboard VRAM.</td>
</tr>
<tr>
<td><strong>Standard VRAM:</strong></td>
<td>512k*</td>
</tr>
<tr>
<td><strong>Maximum VRAM:</strong></td>
<td>512k*</td>
</tr>
<tr>
<td><strong>Built-in Display:</strong></td>
<td>9.8&quot; Grayscale</td>
</tr>
<tr>
<td><strong>Native Resolution:</strong></td>
<td>640x400</td>
</tr>
<tr>
<td><strong>Details:</strong></td>
<td>9.8&quot; grayscale (4-bit, 16 grays) FSTN passive-matrix LCD.</td>
</tr>
<tr>
<td><strong>Standard Hard Drive:</strong></td>
<td>40, 80, 120 MB</td>
</tr>
<tr>
<td><strong>Int. HD Interface:</strong></td>
<td>SCSI</td>
</tr>
</tbody>
</table>
18 July 2010

**Reading PowerBook 2.5" SCSI Hard Drives**

Back in the 68k era Apple used small form factor 2.5" SCSI hard drives in the PowerBook 100, 500 and Duo lines of notebooks. Use of the SCSI format provided compatibility with desktop Macs of the same era and permitted the use of external **SCSI Disk Mode** - the precursor to today's Target Disk Mode.

The 2.5" SCSI form factor was not widely used throughout the industry, however. Most vendors (and soon Apple themselves) utilized ATA (IDE) drives. ATA remained the industry standard until SATA took over this past decade.

Occasionally somebody with an old SCSI-based PowerBook will contact me about transferring files from their old system. If their PowerBook doesn't boot, or the drive was previously removed from the computer, it becomes necessary to access the disk directly.

The scarcity of the 2.5" form factor means that there aren't many ways to read old PowerBook hard drives without having to install them in another PowerBook. A
Hi,

I need to make a disk image of the hd in a Powerbook 160. The machine has a 40GB 40 pin SCSI hd and I cannot find a 40to50 pin converter. Any idea to make the disk image? Thanks
White SCSI-2 Adapter with Mac HDI-30 to HD-50 Connectors
CODE: BPC0-000HDI30C50SW

List price: $5.94
Price: $5.05
You save: $0.89 (15%)
In stock

Qty: 1 Add to Cart Add to compare list
Hi Peter

Unfortunately I have not found a solution to this problem. I will cancel the $15.00 charge in google checkout.

Thank you
TANGERINE FUSION

SCSI and FireWire Disk Modes

Paulo Rodrigues - 2000.11.29

Our Fair Computer Company has released some quirky yet useful features in its computer systems and OS, and then advertised them very little, if at all. Apple's SCSI Disk Mode, and it's modernized offspring, FireWire Target Disk Mode, are excellent examples.

SCSI Disk Mode

SCSI Disk Mode, introduced way back in October 1991 on the PowerBook 100, allowed you to mount your PowerBook's hard drive on another Mac using a funny $30 cable made called the SCSI Disk Adaptor. (Apple changed its name to HD Target Mode starting with the 5300 and 190, since they used IDE hard drives, but it works the same way. For consistency I'll refer to them both as SCSI Disk Mode.)

While today you can do the same thing with File Sharing and a $15 ethernet "crossover" cable, on most models you're limited to the relatively poor bandwidth of 10 megabit per second ethernet. Also, if your PowerBook has no ethernet port, you'd need to buy either an expensive PC Card ethernet adaptor or a speed-squashing LocalTalk-to-Ethernet bridge.

I'll refer to the Mac acting as an external hard drive as the "target" Mac, like Apple does, and the connecting computer as the "host" Mac.

Setup of SCSI Disk Mode
Future Archives

• Creeley: 593.2 GB; 607,437 files
• STOP Aids: 835.7 GB; 403,115 files
• Mandelbrot: 79 GB; 166,980 files
• Fikes: 100,000 emails
• Creeley: 30,000 emails

• What is MPLP for born digital materials?
• Computer caused the problem. Can computer help to solve the problem?
The Science Behind an Answer

Watson performs so fast that it can rival the greatest human contestants in understanding a Jeopardy! clue and arriving at a single, precise answer. The significance of this accomplishment can be difficult to comprehend.

Watch the video to see how the computing system designed to play Jeopardy! works.

The first person mentioned by name in ‘The Man in the Iron Mask’ is this hero of a previous book by the same author.
Topic Modeling Martha Ballard’s Diary

• Martha Ballard’s exhaustive diary, which records daily entries over the course of 27 years.

• Cameron Blevins, PhD candidate in History, at Stanford used MALLET to analyze the texts.
The power of topic modeling really emerges when we examine thematic trends across the entire diary. As a simple barometer of its effectiveness, I used one of the generated topics that I labeled **COLD WEATHER**, which included words such as **cold, windy, chilly, snowy, and air**. When its entry scores are aggregated into months of the year, it shows exactly what one would expect over the course of a typical year:
MUSE – Memories USing Email

EMAIL ARCHIVE OF ROBERT W. CREELEY
OCTOBER 23, 1994 TO JUNE 24, 2006

28,147 MESSAGES
1,765 outgoing
26,382 incoming

22,826 PEOPLE

19 GROUPS
Covering 556 people
and 4,561 messages

5,479 ATTACHMENTS
1676 images

What next? Click on the icons below to view your sentiment patterns, groups, calendar highlights, and attachments.
New! Try connecting your archives to your web browsing with the Muse lens.
Paste text into the “Archivist Tool” Box

Provide some reference text about the owner of this email archive, or provide a list of URLs.

For example, you can paste in text of the owner’s Wikipedia page.

Muse will generate a set of leads into the archive based on this text.

Robert Creeley (May 21, 1926 – March 30, 2005) was an American poet and author of more than sixty books. He is usually associated with the Black Mountain poets, though his verse aesthetic diverged from that school’s. He was close with Charles Olson, Robert Duncan, Allen Ginsberg, John Wieners and Ed Dorn. He served as the Samuel P. Capen Professor of Poetry and the Humanities at State University of New York at Buffalo. In 1991, he joined colleagues Susan Howe, Charles Bernstein, Raymond Federman, Robert Bertholf, and Dennis Tedlock in founding the Poetics Program at Buffalo. Creeley lived in Waldoboro, Maine, Buffalo, New York, and Providence, Rhode Island where he taught at Brown University. He was a recipient of the Lannan Foundation Lifetime Achievement Award.
Annotated text (Ensure that the browsing lens is on).

Robert Creeley (May 21, 1926 – March 30, 2005) was an American poet and author of more than sixty books. He is usually associated with the Black Mountain poets, though his verse aesthetic diverged from that school’s. He was close with Charles Olson, Robert Duncan, Allen Ginsberg, John Wieners and Ed Dorn. He served as the Samuel P. Capen Professor of Poetry and the Humanities at State University of New York at Buffalo. In 1991, he joined colleagues Susan Howe, Charles Bernstein, Raymond Federman, Robert Bertholf, and Dennis Tedlock in founding the Poetics Program at Buffalo. Creeley lived in Waldoboro, Maine, Buffalo, New York, and Providence, Rhode Island where he taught at Brown University. He was a recipient of the Lannan Foundation Lifetime Achievement Award.
Annotated text (Ensure that the browsing lens is on).

Robert Creeley (May 21, 1926 - March 30, 2005) was an American poet and author of more than sixty books. He is usually associated with the Black Mountain poets, though his verse aesthetic diverged from that school's. He was close with Charles Olson, Robert Duncan, Allen Ginsberg, John Wieners and Ed Dorn. He served as the Samuel P. Capen Professor of Poetry and the Humanities at State University of New York at Buffalo. In 1991, he joined colleagues Susan Howe, Charles Bernstein, Raymond Federman, Robert Bertholf, and Dennis Tedlock in founding the Poetics Program at Buffalo. Creeley lived in Waldoboro, Maine, Buffalo, New York, and Providence, Rhode Island where he taught at Brown University. He was a recipient of the Lannan Foundation Lifetime Achievement Award.
Date: July 12, 2001 7:05am
From: pfox<pfox@ndbooks.com>
To: Robert Creeley <creeley@acsu.buffalo.edu>, Ewwigg@aol.com
Subject: FW: Robert Creeley

Dear Bob,

I’m just back from a little holiday that took us to western PA and Ohio to see relatives and then a few days at the Shakespeare festival in Stratford, Ontario—good productions of the Henriad particularly Henry IV, Part I which leapt to life on the very simple thrust stage.

Meanwhile back at the ranch: I’ve received the following from Yun Kim at the Lannan Foundation. Now, do you want some kind of banner at the top of the backad: “Recipient of the 2001 Lannan Foundation Lifetime Achievement Award” or whatever wording you prefer? I think this would be more affective than just incorporating the information into the backad copy. We should probably change the wording of paragraph one which refers to the Before Columbus Foundation Award to make it all “flow” properly. I assume you got my e-mail of June 22 and checked the new quotes for the backad. I need a confirmation from you on this. I’ll be writing Yun Kim right now about the