



Stanford University LIBRARIES &
ACADEMIC INFORMATION RESOURCES

Accessioning Born-Digital Materials

Northern California Fall Workshop

Society of California Archivists

Peter Chan, Digital Archivist

Nov. 8, 2012

Agenda

- Literature review
- Put accessioning in context
- Stanford work flow
- Hands-on
- Demonstration
- Questions
- Forensic Lab Tour

Oxford and Manchester

Workbook on Digital Private Papers (2007)

<http://www.paradigm.ac.uk/workbook/index.html>

- One section on “Accessioning digital and hybrid personal archives”
- Very comprehensive (except delivery)

Hull, Stanford, Yale, and UVa

AIMS Born-Digital Collections:

An Inter-Institutional Model for Stewardship
(2011)

<http://www2.lib.virginia.edu/aims/whitepaper/>

- High level
- Covering pre-accessioning, accessioning, processing and delivery

OCLC

You've Got to Walk Before You Can Run: First Steps for Managing Born-Digital Content Received on Physical Media (2012)

<http://www.oclc.org/research/news/2012/08-23.html>

- Very basic
- More to follow

Stanford

Born-Digital Archives Program: Forensics
Workflow Documentation: (2012)

<https://sites.google.com/site/workflowdocumentation/home>

- Used in processing the born digital component of the STOP Aids Project Records
- Detail
- Still work-in-progress

Best Practices

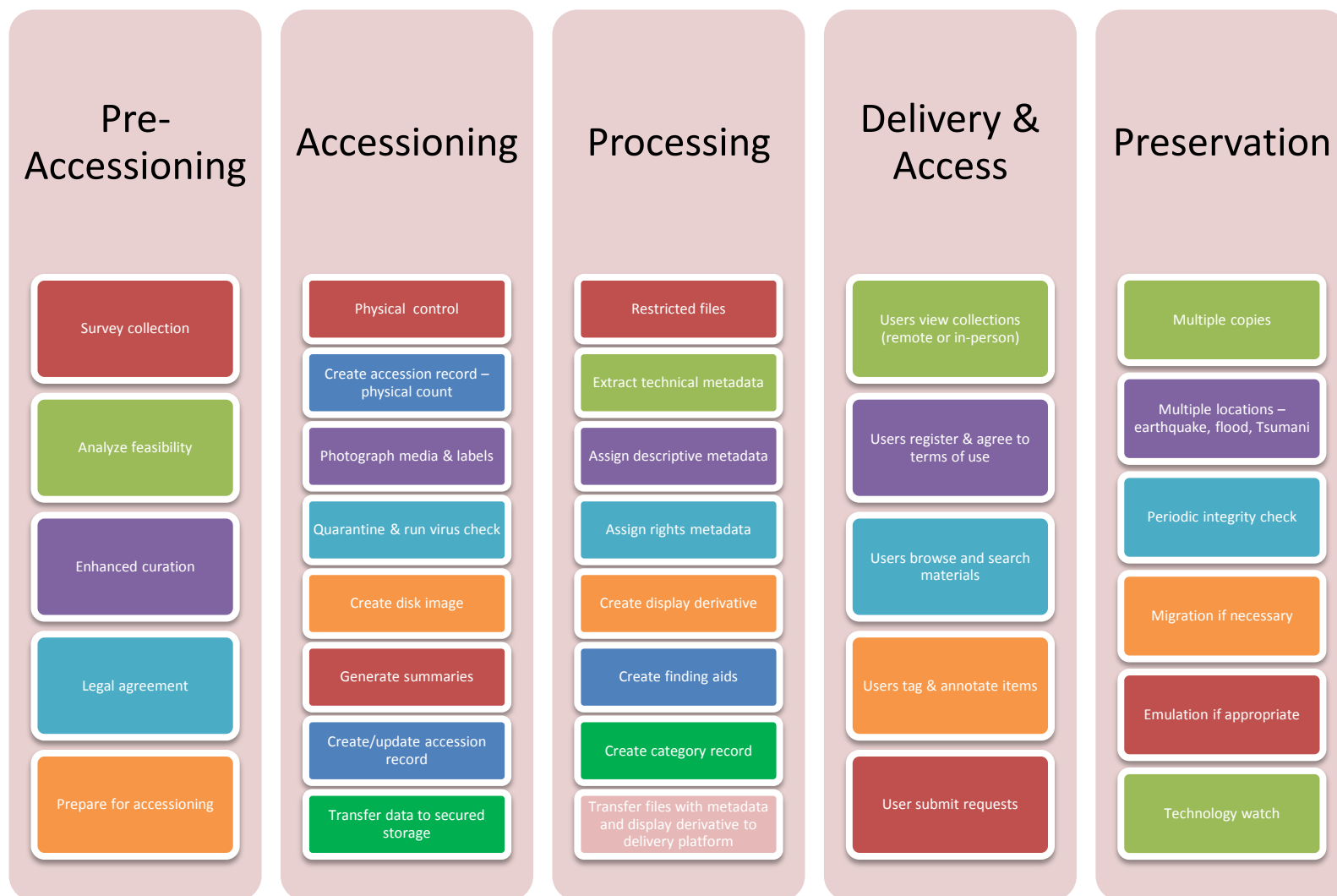
- Harder to agree (compare to processing of paper archives) because the vast difference among institutions in software and hardware platforms
- Not enough people with necessary knowledge to contribute
- Still early stage of development (especially in processing and delivery)

Don't Act Now

- Stanford spent \$10K to recovery files from 4 hard drives received ~10 years ago.
- Only 47 percent of the recordable DVDs tested indicated an estimated life expectancy beyond 15 years. Some had a predicted life expectancy as short as 1.9 years. (note 1)
- Manufacturing of 5.25 inch floppy drive stopped for some years already.

Note 1: <http://www.thexlab.com/faqs/opticalmedialongevity.html>

Born-Digital Workflow – Textual Files



Accessioning

Activities

“Physical” control

Create accession record

Photograph media & labels

Quarantine & run virus check

Create disk image

Generate summaries

Create/update accession record

Transfer data to secured storage

Tools / Persons

Manual / SecureFX™ / Others

Archivist's Toolkit (AT)

Canon EOS T1i with copy stand

Sophos™

FTK™ Imager

AccessData FTK

Archivists' Toolkit (AT)

SecureFX™

Deliverables

Control of media, files held locally

Accession record – physical count

Photographs of labels and media


Virus free files

Disk image, audit log, image log

Summary report

AT accession record – no. of files and size

Files on secure network

 Record 2 of 3

Accessions

Title: Rebecca Solnit papers

Accession #: MSS.2011-119

Basic Information

Accession Notes

User defined fields

Names & Subjects

Acknowledgements, Restrictions & Processing Tasks

Accession Number

MSS

2011-119

Accession Date

2011-8-24

Resources Linked to this accession

Resource Identifier ▲	Resource Title
M1839	Rebecca Solnit papers

Link Resource Remove Link

Resource Type

Papers ▼

Title

Rebecca Solnit papers

Physical Description

Primary Extent

Alternate Extent

Extent Number

31

Linear feet ▼

Container Summary

24 cartons, 3 lap tops

Repository

SPECCOLL

Date Expression

Date

Begin

End

Bulk Dates

Begin

End

Deaccessions

Deaccession Date	Extent	Extent Measurement
------------------	--------	--------------------

Add Deaccession Remove Deaccession

Locations

Location	Note
----------	------

Add Location Remove Location

Accession Notes

no listing available - therefore closed till processed.
no restrictions.

First

Previous

Next

Last

Reports

Close

Save

Created: Aug 24, 2011 by ohanluin | Modified: Mar 30, 2012 by pchan3 | Record Number: 32571

AT Accession Records

- Physical
 - Media count
 - Computer count


[Register](#) or [Login](#)

 Search
[Home](#) [Overview](#) [Download](#) [Support](#) [For Developers](#) [Add Ons](#) [Forum](#) [About Us](#)

Application Bugs

Bug reports can be transmitted using the bug report function in the application. Otherwise they should be sent to the AT project at info@archiviststoolkit.org

Other comments or questions should also be sent to info@archiviststoolkit.org

Copyright 2006-2009

Multiple Extent Plugin

Submitted by [michael_vandermillen](#) on Thu, 08/11/2011 - 08:29

Independent of collection size, archivists count and report on a variety of container and media types as they manage, store, and seek to understand the nature of their collections and meet preservation and access challenges.

In answer to this need, Harvard developed a two-fold approach in the Archivists' Toolkit. First, we re-labeled the AT default extent as "Primary extent." Second, we developed a plug-in* to allow multiple alternate statements of extent and labeled this "Alternate extent." The alternate extent statements have the same data elements as the default AT extent statements, but the alternate extents have a separate drop-down list of extent types. Alternate extents are not required, but multiple alternate extents may be applied to an accession or resource record.

This plugin is an extension of BYU's Multiple Date and Physical Description custom panels plugin.

Attachment	Size
Harvard_Multiple_extent_plugin.pdf	239.88 KB

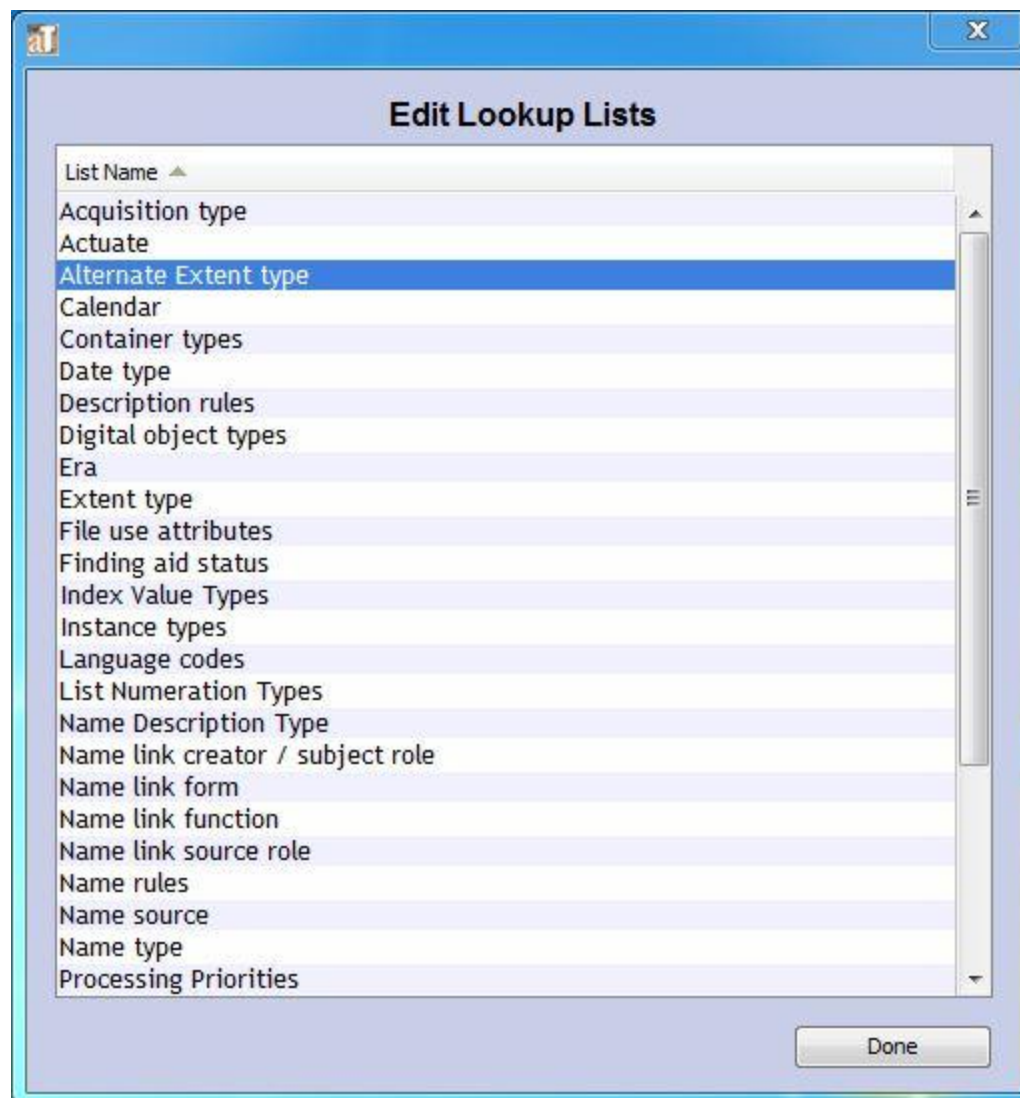
Author: Harvard University Library


Zip File:
 [multipleExtent.zip](#)

Source File:
 [multipleExtentSrc.zip](#)

Description:
 Plugin to allow primary and alternate extent entry for accession and resource records (based on the BYU Multiple Date and Phys. Desc. plugin)

[Login](#) or [register](#) to post comments





Administration **Lookup Lists** Record 3 of 34

List Name

☐ Paired Values ☐ Restrict to NMTOKEM

List Item	Code
album(s)	
architectural drawing(s)	
audiocassette(s)	
audiocassette(s) (DAT)	
audiocassette(s) (microcassette)	
audiotape(s) (reel-to-reel)	
boxes	
computer cartridge(s)	
computer file(s)	
computer media	
computer tape(s)	
computer, desktop	
computer, portable	
cubic feet	
film reel(s) (16mm)	
film reel(s) (8mm)	
floppy disk(s) (3.5 inch)	
floppy disk(s) (5.25 inch)	
floppy disk(s) (8 inch)	
folder(s)	
folio(s)	
gigabyte(s)	
hard drive(s) (external)	

Add Item

Remove Item(s)


Change Item


Merge Items


Import Items


Show Record Count


Items in Red can't be modified
Items in Blue are AT initial values


 First

 Previous

 Next

 Last

 Cancel

 OK

Created: Sep 12, 2011 by dhartwig | Modified: Oct 19, 2012 by dhartwig | Record Number: 67

Accessions

Record 2 of 3

Title: Rebecca Solnit papers

Accession #: MSS.2011-119

Basic Information | Accession Notes | User defined fields | Names & Subjects | Acknowledgements, Restrictions & Processing Tasks

Accession Number MSS 2011-119

Date Expression

Accession Date 2011-8-24

Resources Linked to this accession

Resource Identifier	Resource Title
M1839	Rebecca Solnit papers

Link Resource Remove Link

Resource Type Papers

Title

Rebecca Solnit papers

Physical Description

Primary Extent | Alternate Extent

Extent Display	Container Summary	Extent Number	Extent Type
3 computer, por...	1 Macintosh Pow...	3	computer, portable

Add Description Remove Description

Repository SPECCOLL

First Previous Next Last

Reports Close Save

Created: Aug 24, 2011 by ohanluin | Modified: Mar 30, 2012 by pchan3 | Record Number: 32571

Physical Descriptions

Record 1 of 1

Extent Number 3 computer, portable

Container Summary

1 Macintosh PowerBook 160
1 Macintosh Powerbook G3
1 Apple PowerBook G4

Physical Detail

Dimensions

First Previous Next Last

Cancel OK

Created: Mar 30, 2012 by pchan3 | Modified: Apr 13, 2012 by pchan3 | Record Number:...

Media Count

- Media count by
 - 3, 3.5, 5.25, 8 inch. floppy diskettes
 - Zip disk
 - Open reel, cartridge tape
 - CD, DVD, Optical disk
 - External hard drive
- Computer
 - Desktop / Portable
 - Mac / PC / Others

8-inch, 5.25-inch, and 3.5-inch floppy



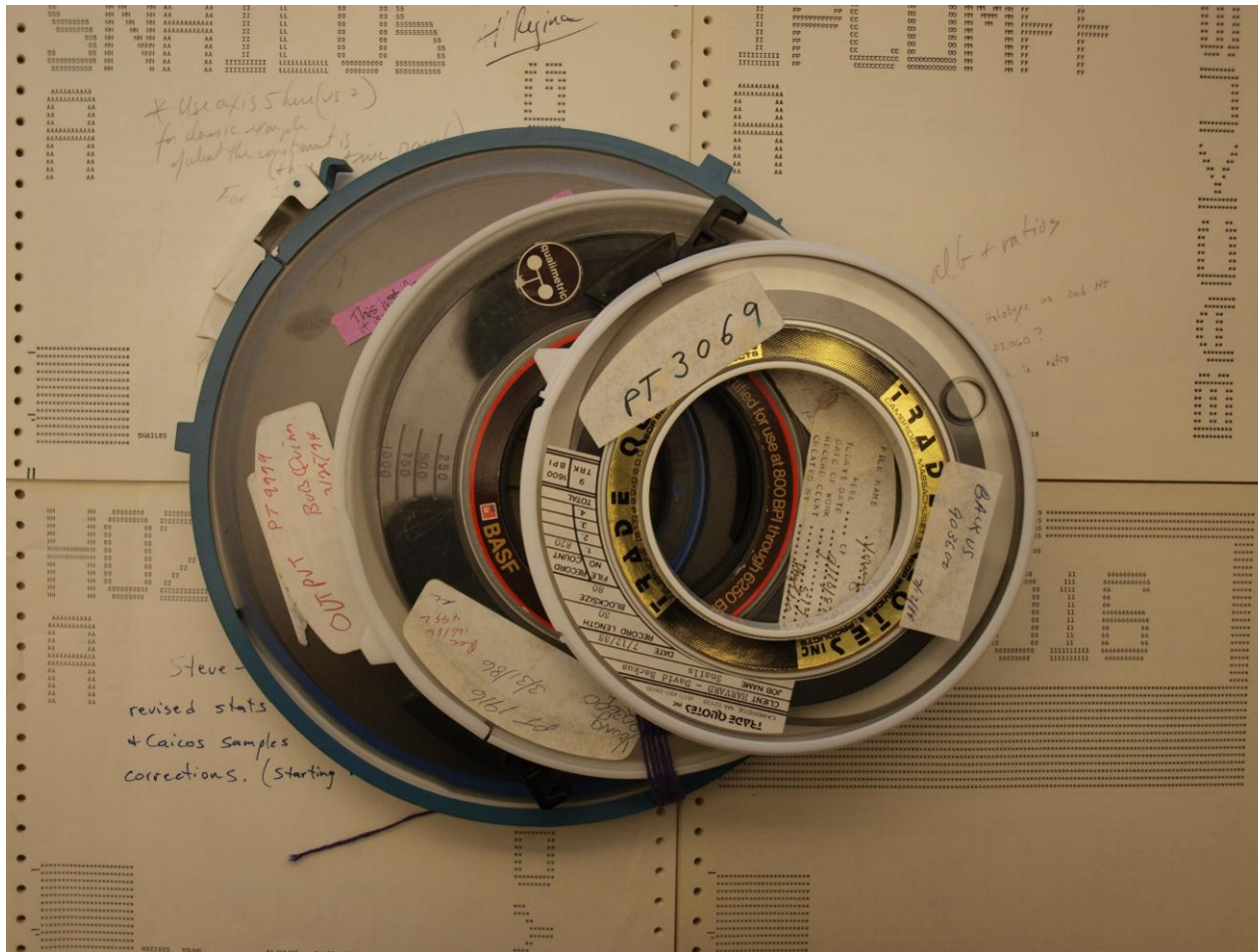
Smith Corona DataDisk 3-inch floppy



100MB Zip Disc for Iomega Zip



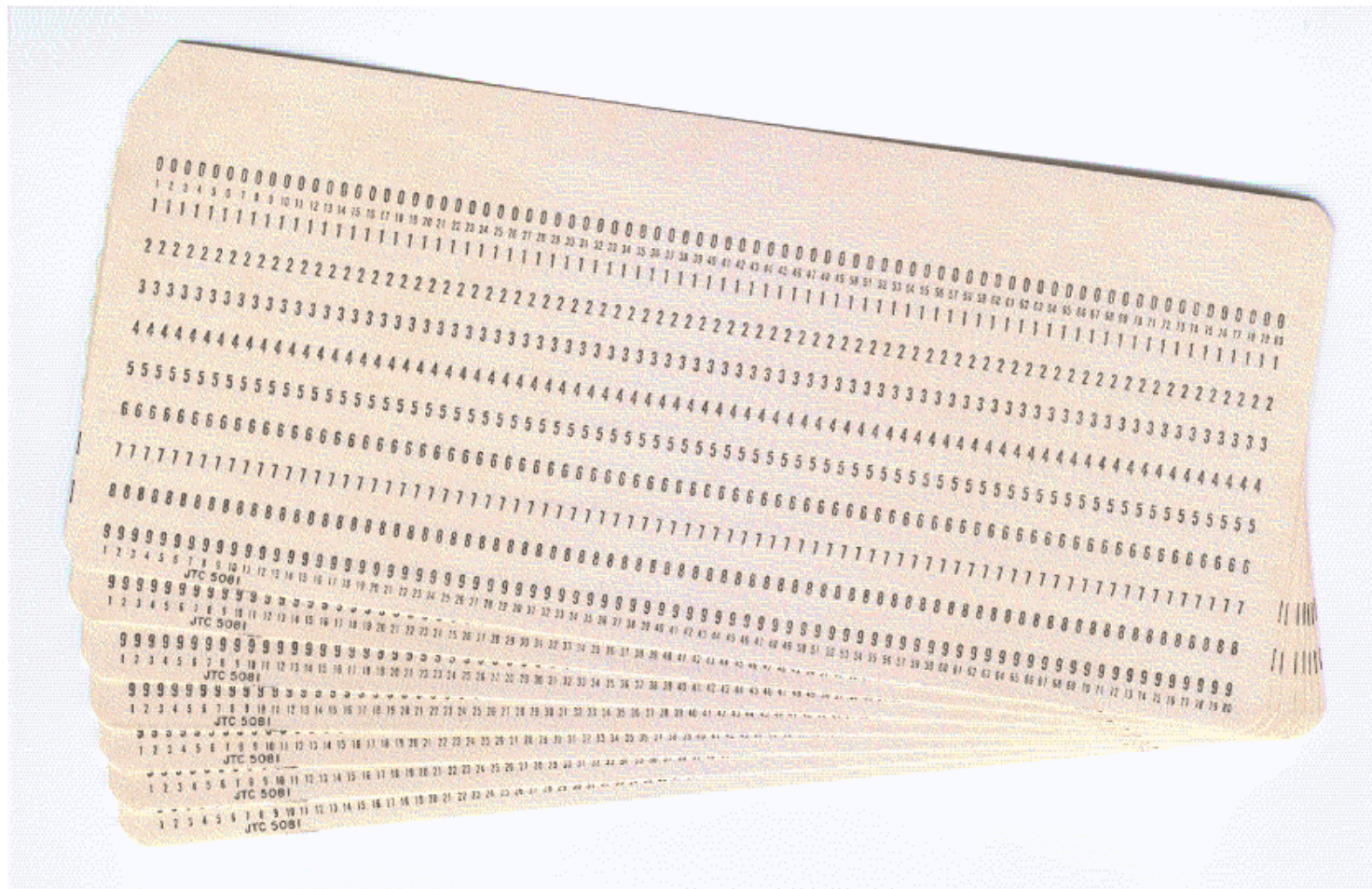
Open Reel Tape



Cartridge Tape



Punch Cards



Apple II



Commodore C64 SX-64



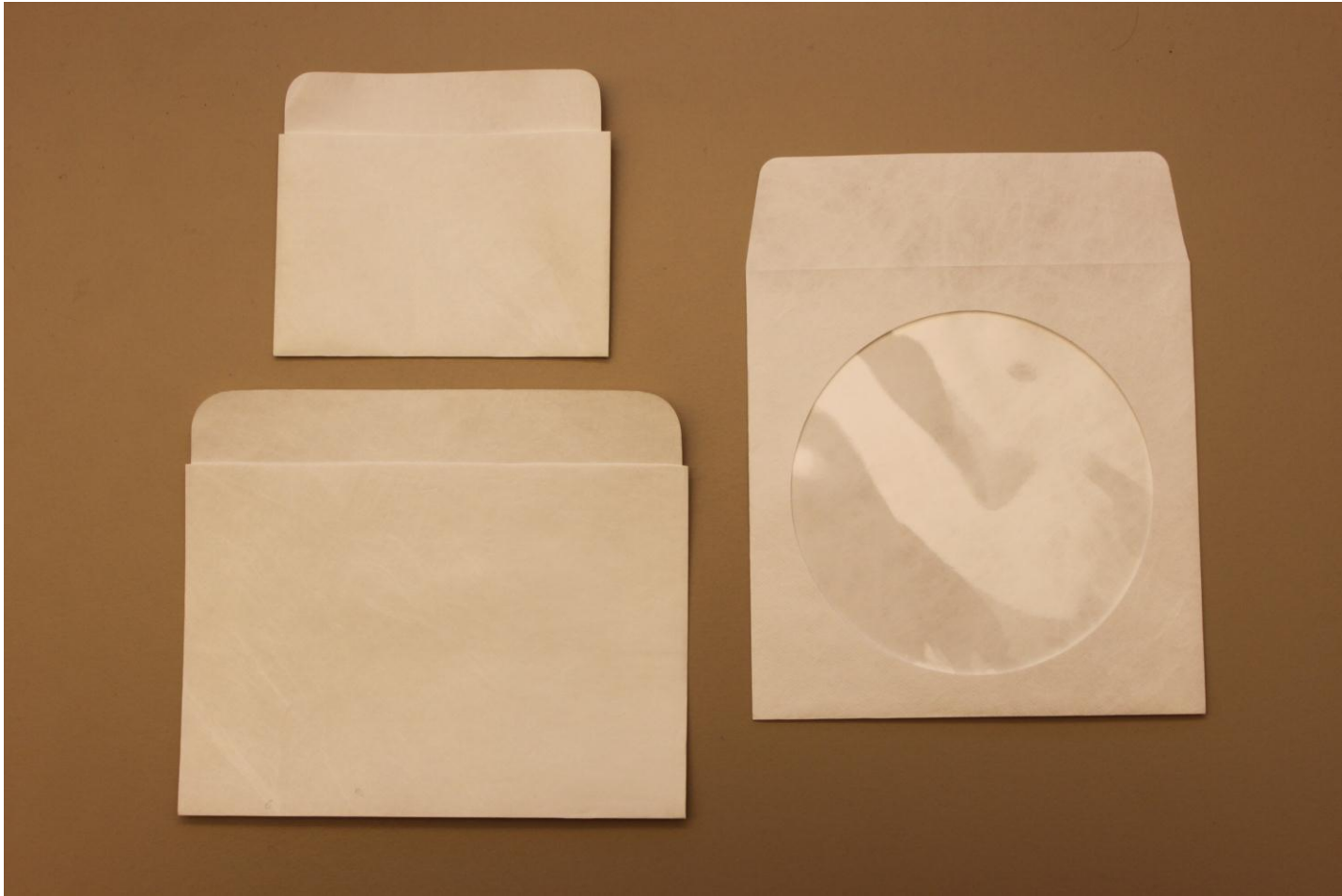
Kaypro 10 portable computer



Media Label /Rehouse

- Label
 - Use “Call No._CMxxx” as label name, 0.5 x 1.875 inch. label
 - Template http://www.avery.com/avery/en_us/Templates-%26-Software/Templates/Labels/Return-Address-Labels/Return-Address-Label-80-per-sheet_Microsoft-Word.htm?N=0&refchannel=c042fd03ab30a110VgnVCM1000002118140aRCRD
- Re-House
 - Follow the same box no. naming convention as other materials (paper, av, etc.).

Media Sleeve



hollingermetalede.com

Floppy Disc Box

Cat #	Description	Size	3 or more	10 or more	20 or more	Qty
CDBB	CD/DVD Box	5 x 4 7/8 x 5 5/8"	\$5.40	\$5.05	\$4.85	<input type="text"/>
CDBT	CD/DVD Box	5 x 4 7/8 x 5 5/8"	\$4.65	\$4.40	\$4.20	<input type="text"/>
FDB442	Floppy Disc Box	3 1/2"	\$3.60	\$3.35	\$3.20	<input type="text"/>
FDB552	Floppy Disc Box	5 1/2"	\$4.05	\$3.85	\$3.70	<input type="text"/>

[Buy](#)

hollingermetalede.com

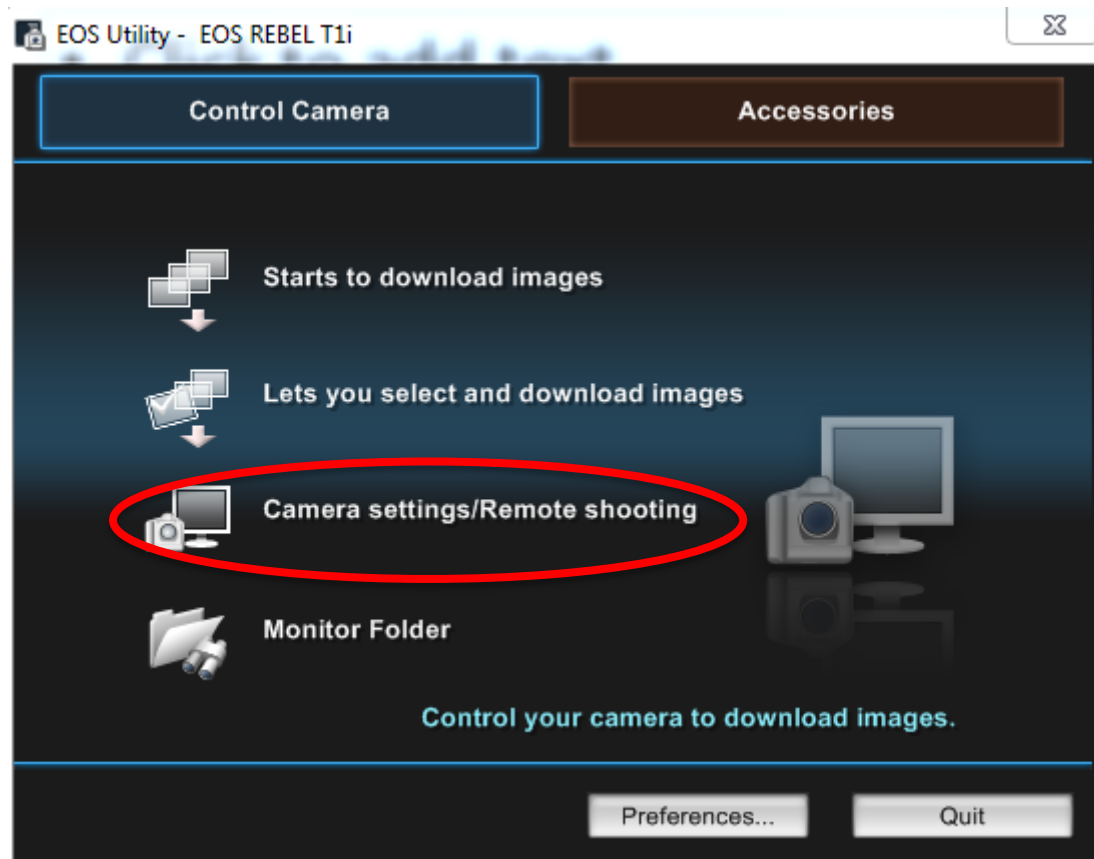


11735HT-A

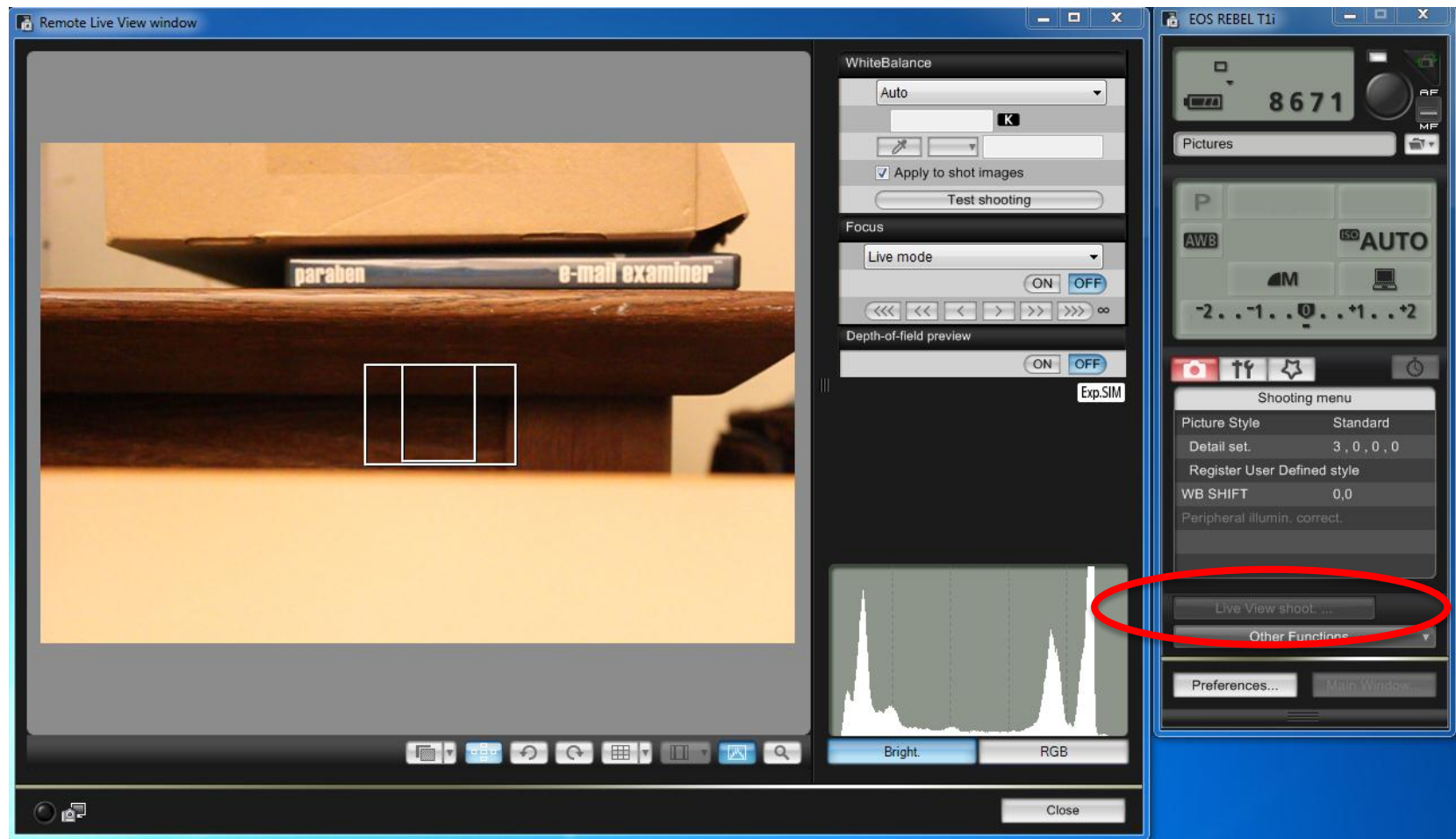
HT-A For 5 x 8 cards, full lid
box

12L X 8 1/8W X
5 1/2H

Canon EOS Utility



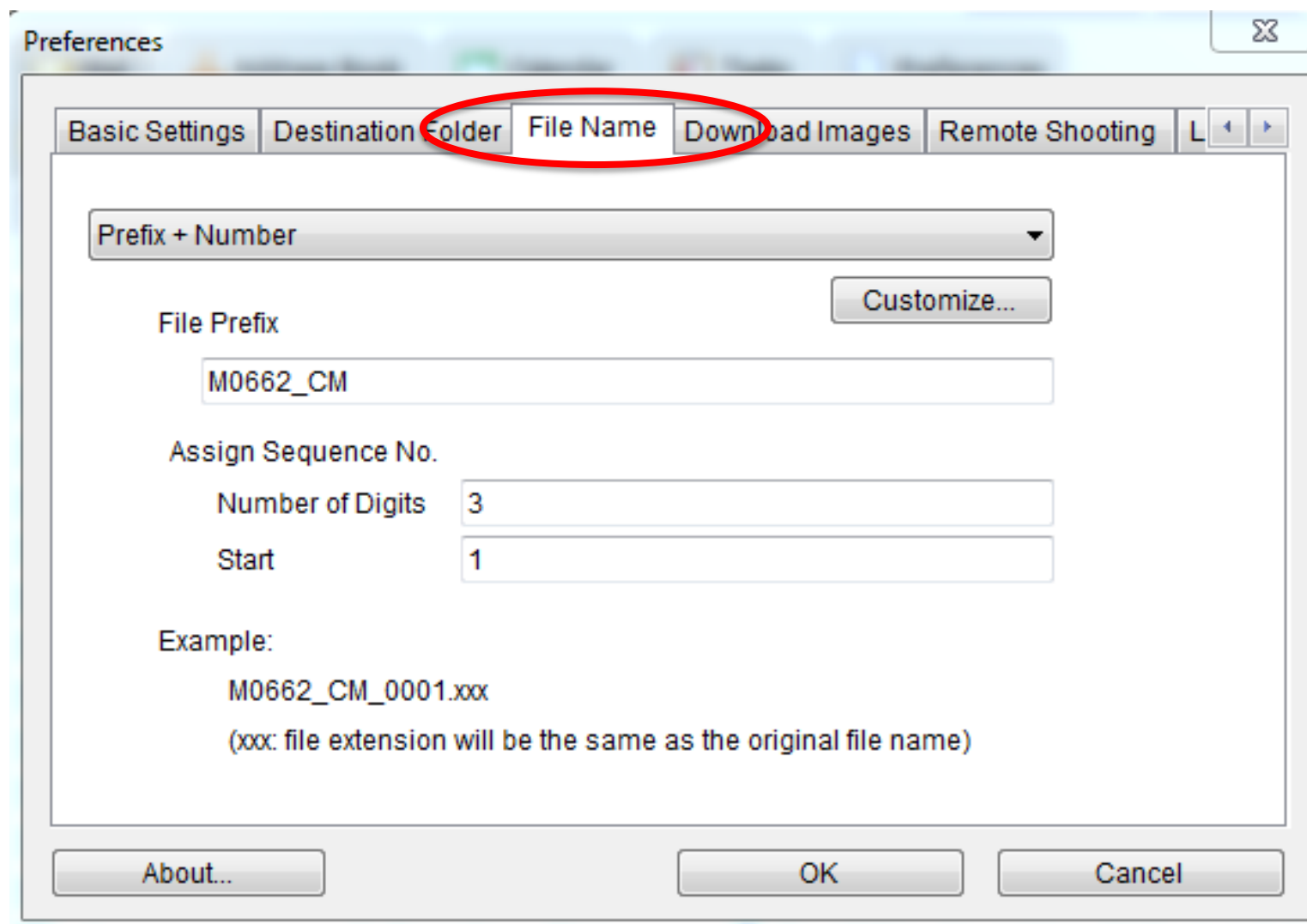
Live View Shooting



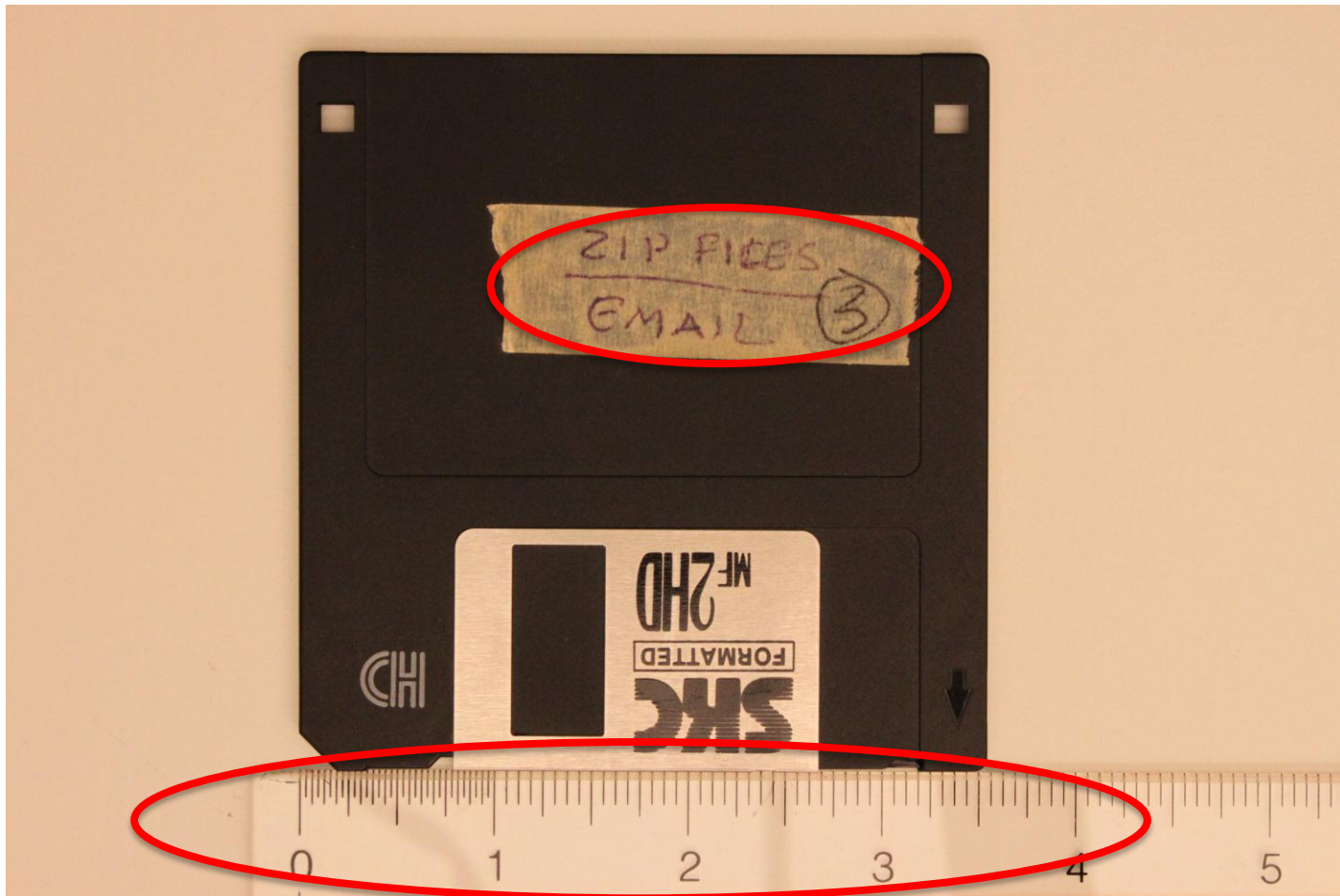
Filename

- Photographing media
 - Use ““Call No._CMxxx “ as filename
 - If more than 1 photo is taken, add _1 for first and _2 for second photo, etc. (e.g. front, back, box, etc.)
 - Use computer to control the camera if you have more than 20 media to photo; otherwise, just use stand alone camera.
 - Store all photos in “Media Photo” folder

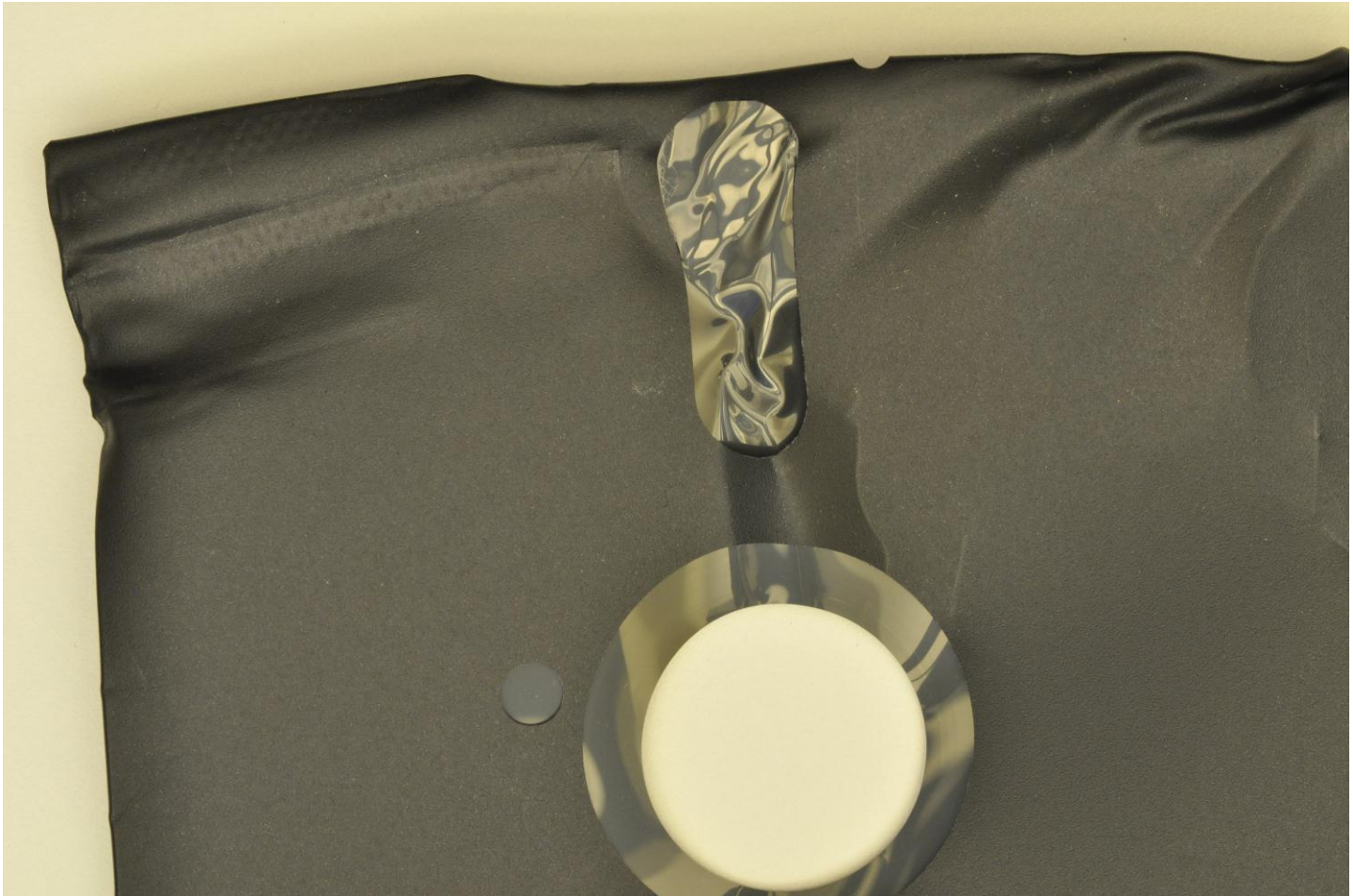
Preference



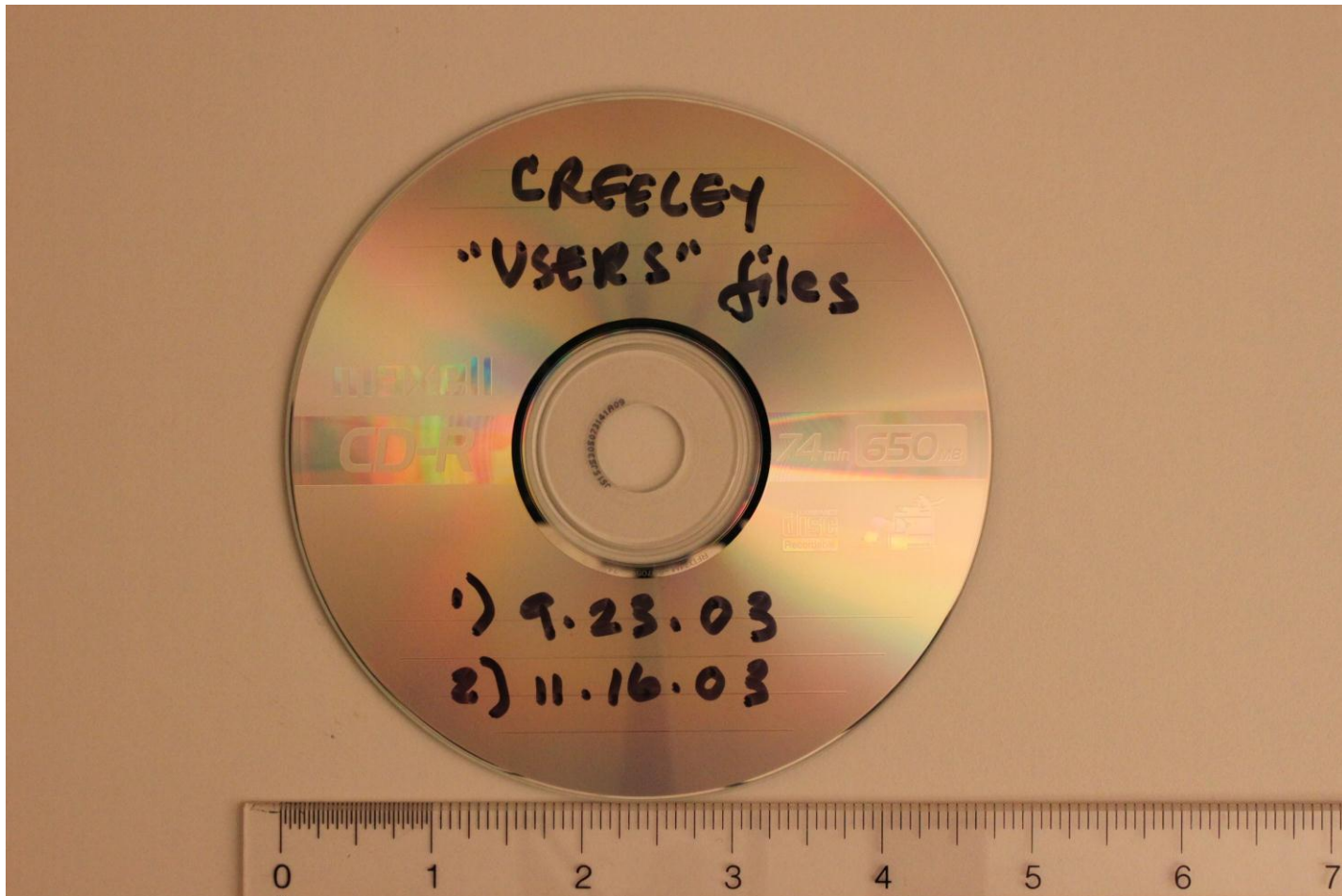
Media Photo



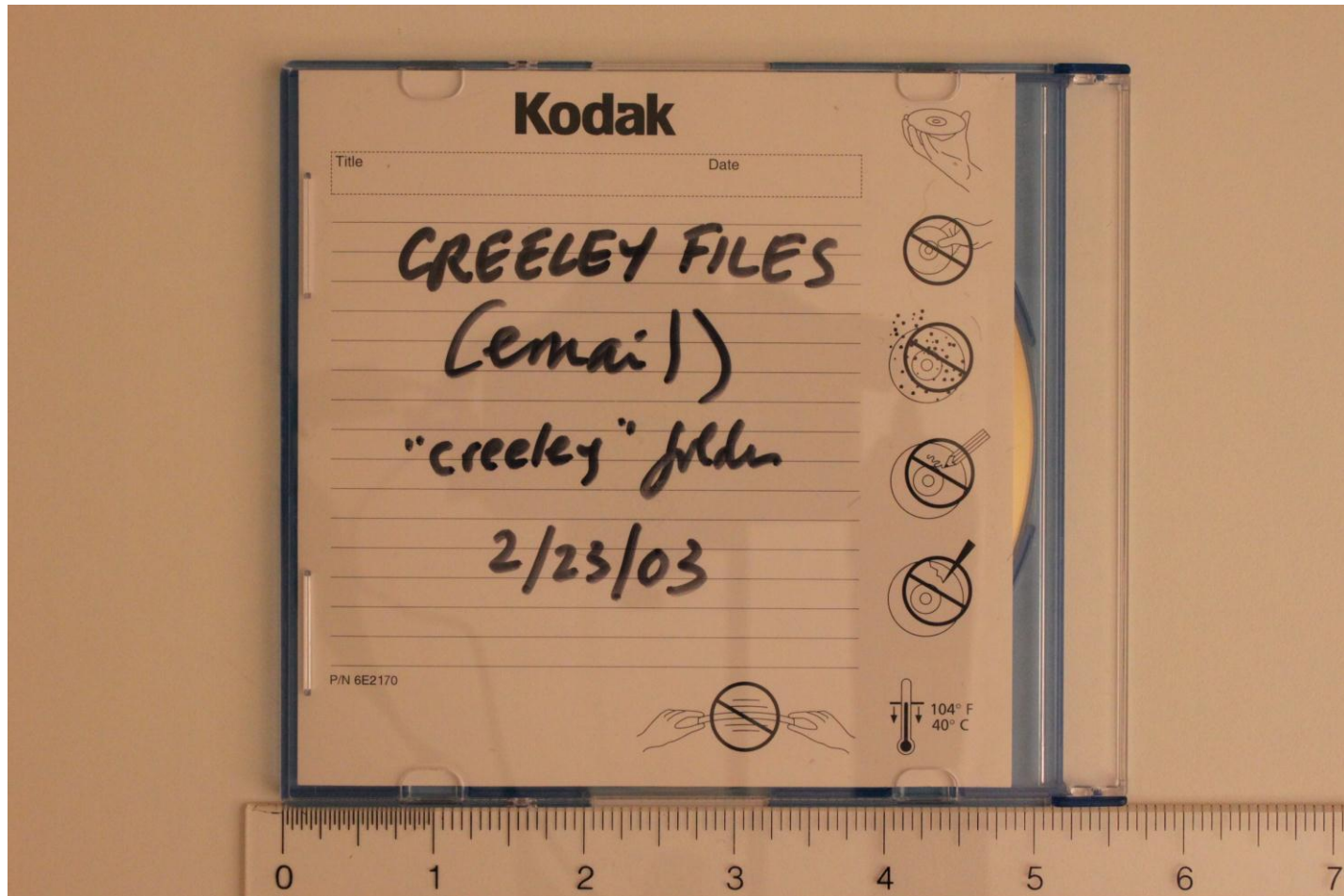
Media Photo



Media Photo



Media Photo



AT Accession Records

- Provide finer information than physical storage media count
 - Size in MB/TB, etc.
 - No. of files
 - Link to
 - Image log spreadsheet
 - Collection summary from FTK

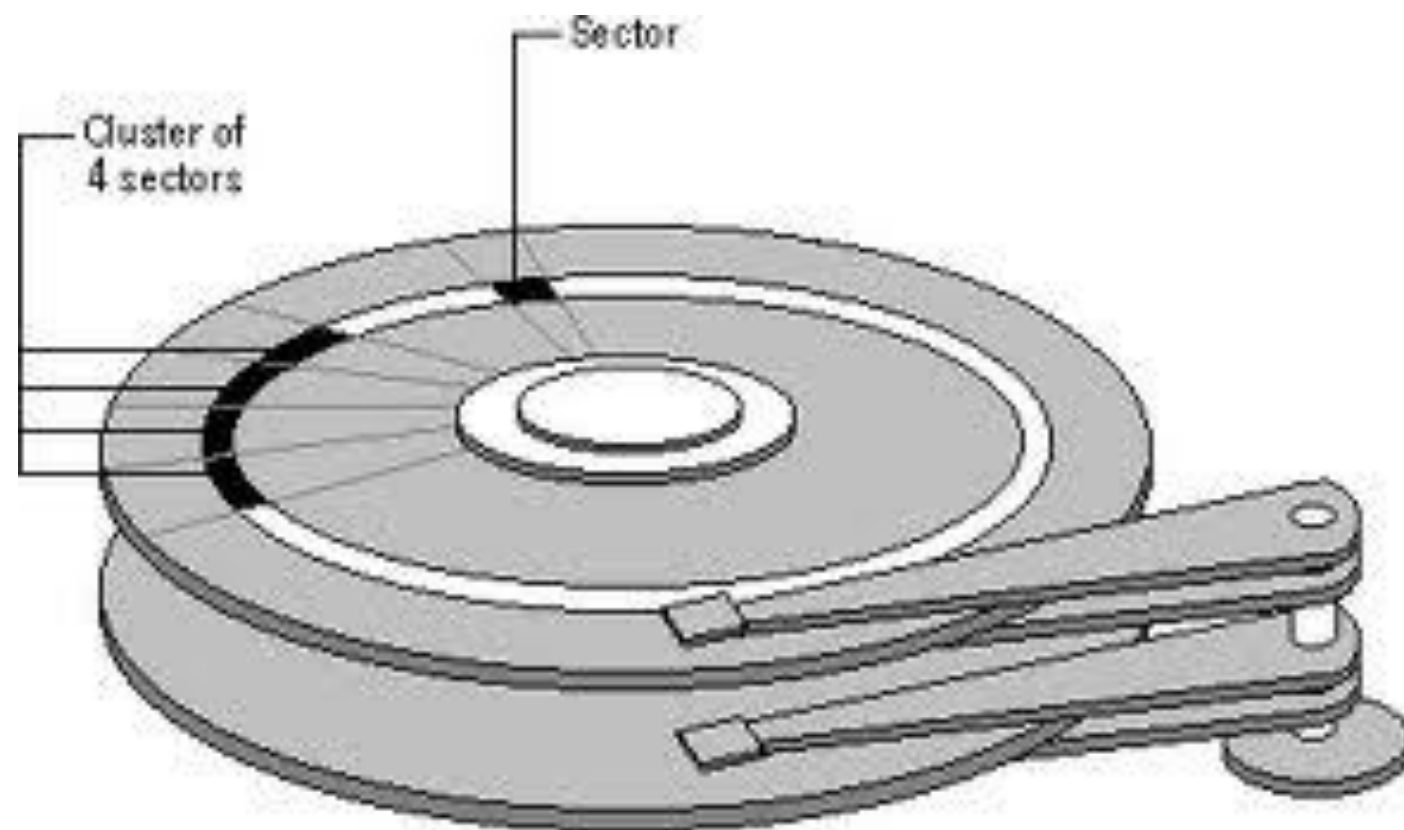
Virus Check

- Quarantine for 30 days (count from the day the media arrive at Stanford)
- Run Sophos
- Remove virus, if any, before creating disk/logical image (unless you are capturing the disk for a researcher in computer virus!)

Forensic vs. Logical Image vs. File Copy

- Do you want to copy deleted files?
- Are you sure you didn't change file dates (creation, modified, last accessed) when copying the files?
- Are you sure you copied all files (files in different partitions)?
- Are you sure you copied all required files (fonts for design files, etc.)
- Are you sure the source and copied files are same?





Disk Partition

- Disk partitioning is the act of dividing a physical hard disk drive into multiple logical storage units.
- The logical units can have different file systems so that we can set the same physical machine as a Window based PC and/or a Linux based PC and/or a Mac PC.

Computer Management

Computer Management

File Action View Help

Computer Management (Local)

- System Tools
 - Task Scheduler
 - Event Viewer
 - Shared Folders
 - Local Users and Groups
 - Performance
 - Device Manager
- Storage
 - Disk Management
 - Services and Applications

Volume	Layout	Type	File System	Status	Capacity	Free Space	% Free	Fault Tolerance	Overhead
(C:)	Simple	Basic	NTFS	Healthy (Boot, Page File, Crash Dump, Primary Partition)	79.20 GB	32.32 GB	41 %	No	0%
(E:)	Simple	Basic	NTFS	Healthy (Primary Partition)	40.04 GB	33.71 GB	84 %	No	0%
(F:)	Simple	Basic	NTFS	Healthy (Primary Partition)	1862.64 GB	399.00 GB	21 %	No	0%
(G:)	Simple	Basic	NTFS	Healthy (Primary Partition)	37.17 GB	27.42 GB	74 %	No	0%
1052253965 (D:)	Simple	Basic	FAT	Healthy (Active, Primary Partition)	96 MB	41 MB	43 %	No	0%
CODEMETER (I:)	Simple	Basic	FAT32	Healthy (Active, Primary Partition)	39 MB	0 MB	0 %	No	0%
System Reserved	Simple	Basic	NTFS	Healthy (System, Active, Primary Partition)	100 MB	69 MB	69 %	No	0%

Disk 0
Basic
37.27 GB
Online

System Reserved
100 MB NTFS
Healthy (System, Active, Primary Partition)

(G:)
37.17 GB NTFS
Healthy (Primary Partition)

Disk 1
Removable
96 MB
Online

1052253965 (D:)
96 MB FAT
Healthy (Active, Primary Partition)

Disk 2
Basic
119.24 GB
Online

(E:)
40.04 GB NTFS
Healthy (Primary Partition)

(C:)
79.20 GB NTFS
Healthy (Boot, Page File, Crash Dump, Primary Partition)

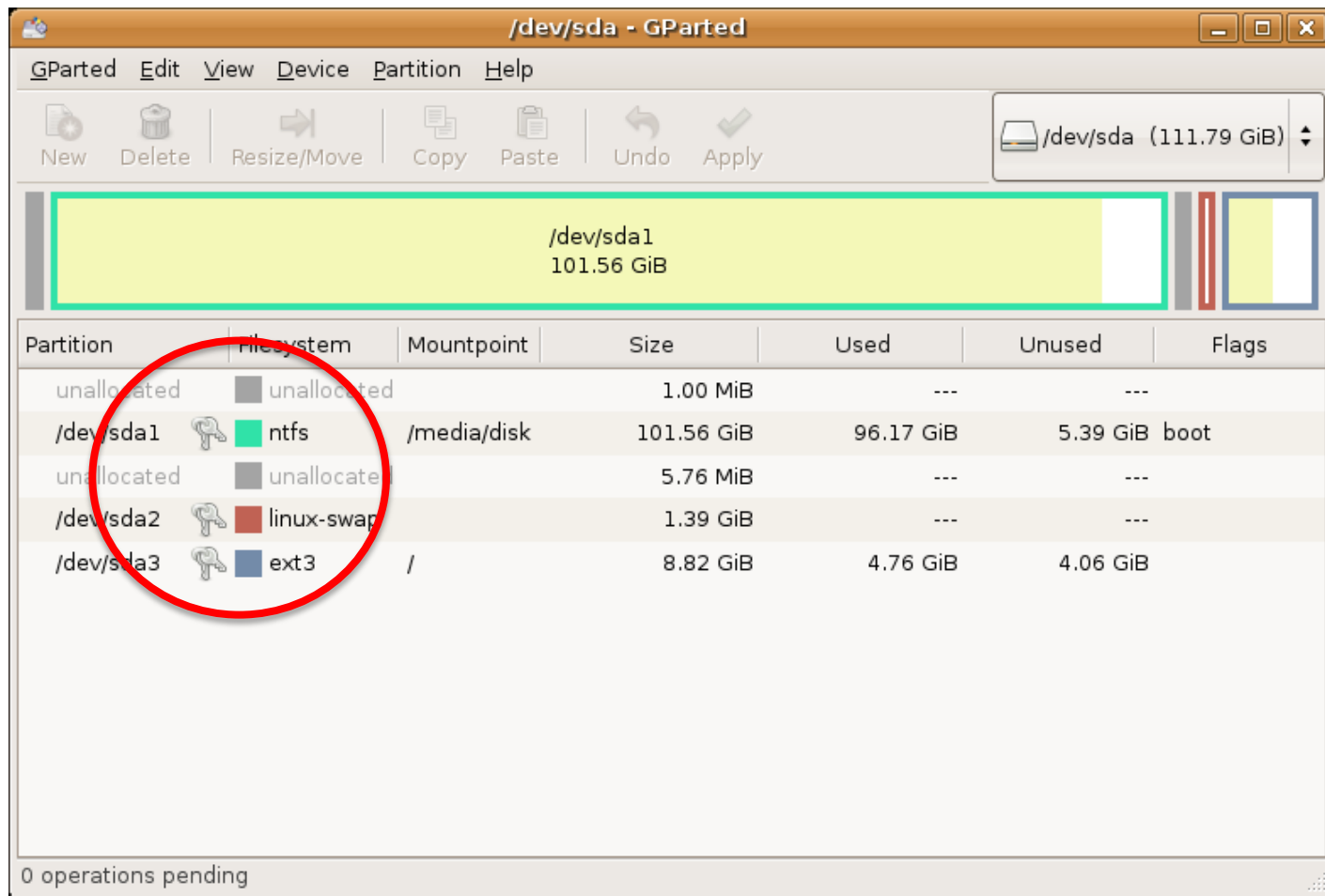
Disk 3
Basic
1862.64 GB
Online

(F:)
1862.64 GB NTFS
Healthy (Primary Partition)

Disk 4
Basic
39 MB
Online

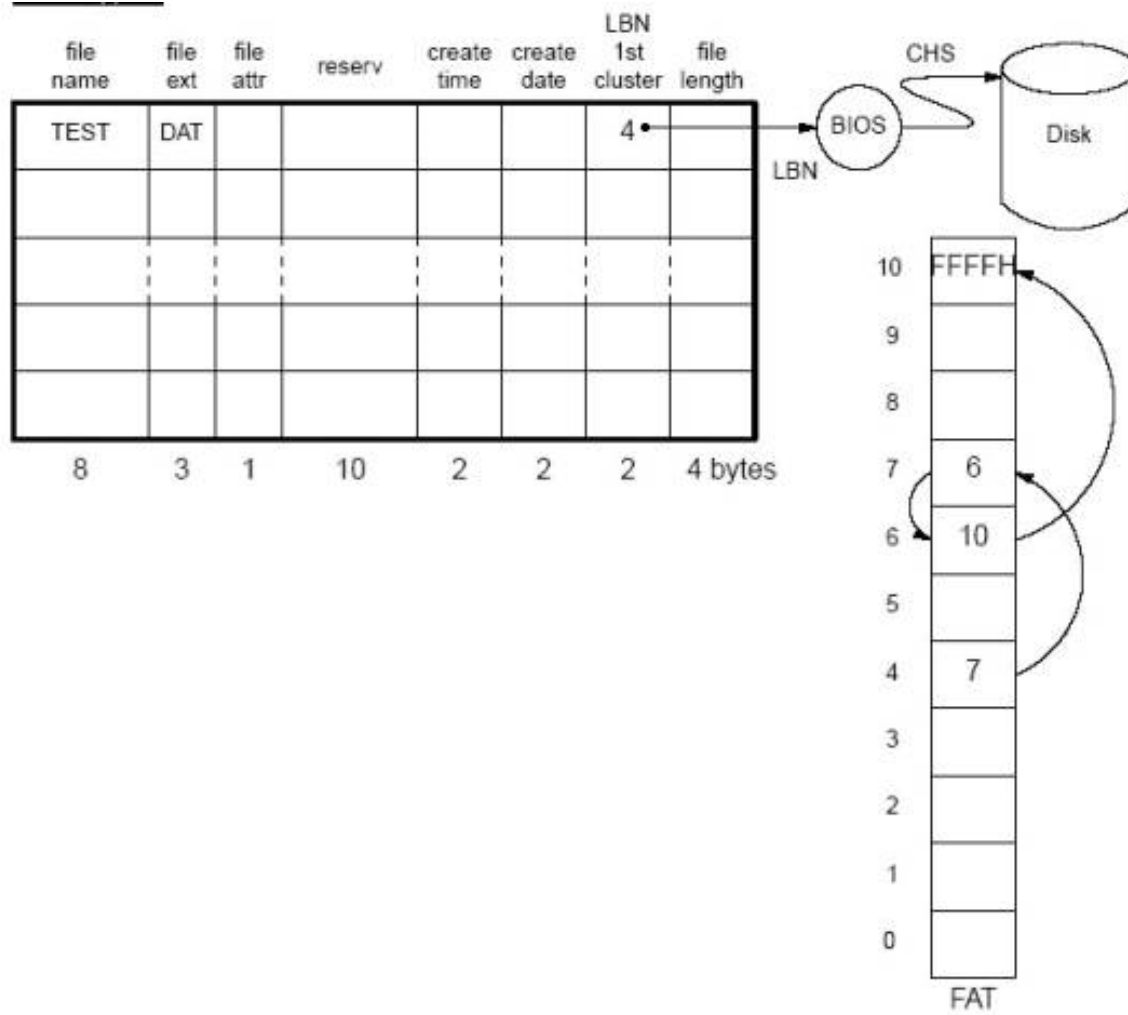
CODEMETER (I:)
39 MB FAT32
Healthy (Active, Primary Partition)

GParted



File System

- Microsoft: FAT12 (File Allocation Table), FAT16, FAT32, exFAT (FAT64), NTFS (New Technology File System)
- Apple: Apple ProDOS, HFS (Hierarchical File System), HFS+
- CPM file system
- Linux: ext (extended filesystem), ext2, ext3, ext4
- Optical Discs: ISO 9660, UDF (Universal Disk Format)



Disk / Forensic image

- A complete bit-by-bit copy of a storage medium or device, such as a hard drive, SSD (solid state drive), tape drive, floppy disk, CD/DVD/BD, or flash memory device.
- The image can be stored in one or more files.
- Deleted files, if any, will be copied in this process.
- All partitions will be copied.

Logical image

- A copy of the files in the directory(folder) / directories(folders) specified in the copy process.
- The full path of each file is recorded and the files are embedded in one or more files in AD1 format.
- Since deleted files and unpartitioned space are not represented in a directory, they are not copied in the process.
- The host operating system has to recognize the target directory (folder) for the operation. (You cannot perform a logical copy of an Apple disk using FTK Imager under Windows perform).

Capture 5.25 Floppy Diskette



Compare 4 5.25 Floppy Drive Solutions

- https://docs.google.com/document/d/1TLY3mn3duadGBLqqb2_XnYoF3jakjmo1hXfzvmWCTzE/edit?hl=en_US

Hard Disk Interface

- SCSI
- IDE
- Firewire 1394
- SATA
- USB

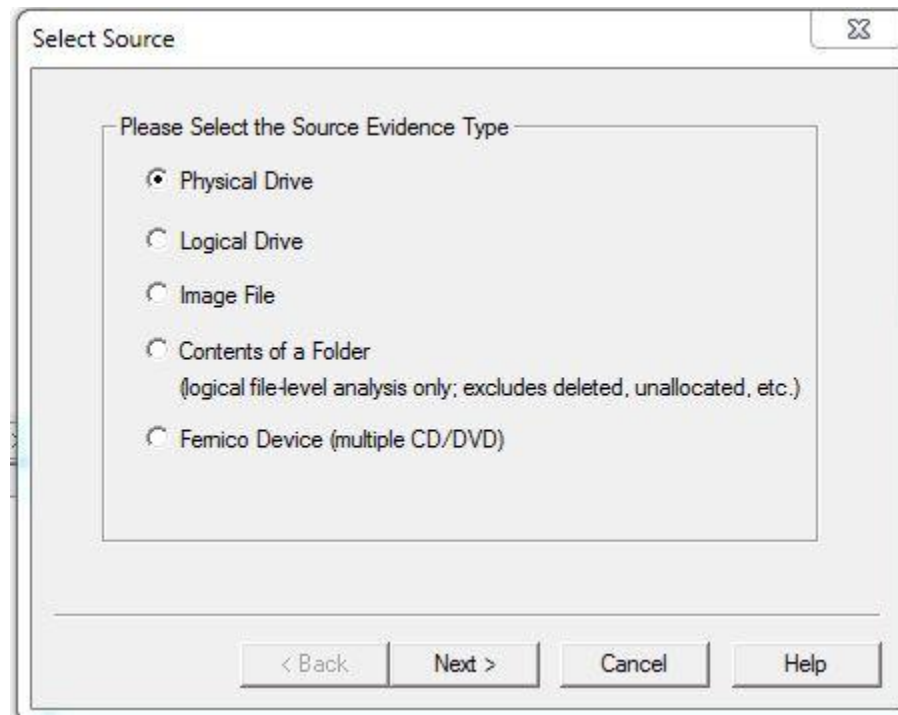
Internal Write Blocker



Image Log Spreadsheet

- Use the following columns
 - Box#
 - CM#
 - Media Details (3.5, 5.25, single/double sided, single/double/high density, etc.)
 - Image Result (successful / unreadable)
 - Note (bad sector, etc.)
 - Manufacturer
 - Earliest modification/creation year of files
 - Folder title (if removed from collection folder)
- Use “Call No. Imaging Log” as file name

FTK Imager



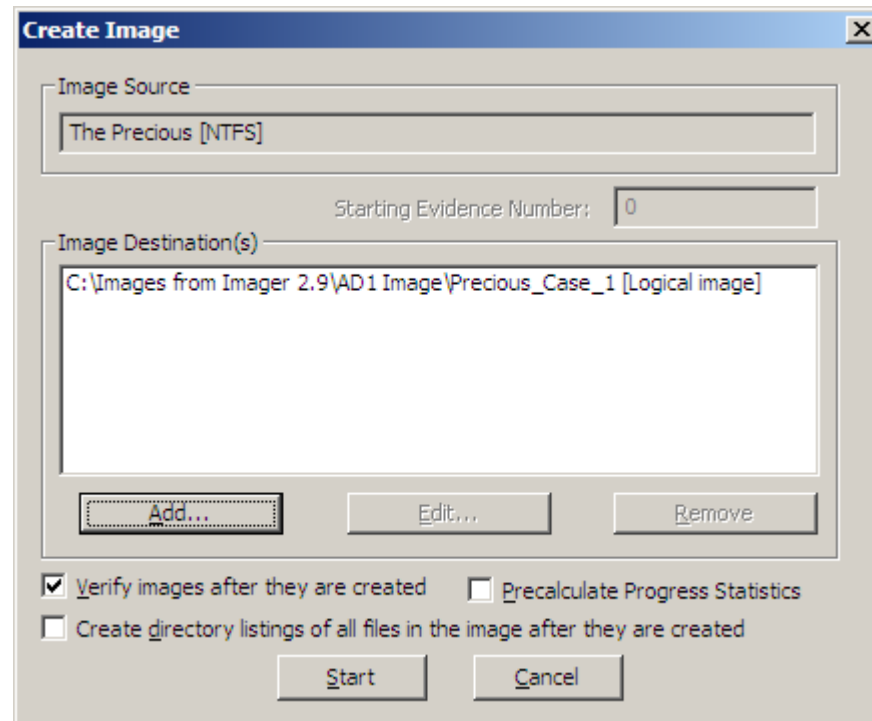
Create Disk Image

- Physical Drive – hard / flash memory drive
- Logical Drive – floppy / CD / DVD drive
- Contents of a Folder – logical image
- Fernico Device – a device which backs up forensic data from network locations or from locally attached hard drives, automatically spanning the content over a series of discs.

Fernico Device



Create Image



Tick “Verify images after they are created”

Tick “Create directory listings of all files in the image after they are created” if you need one.

Image Type

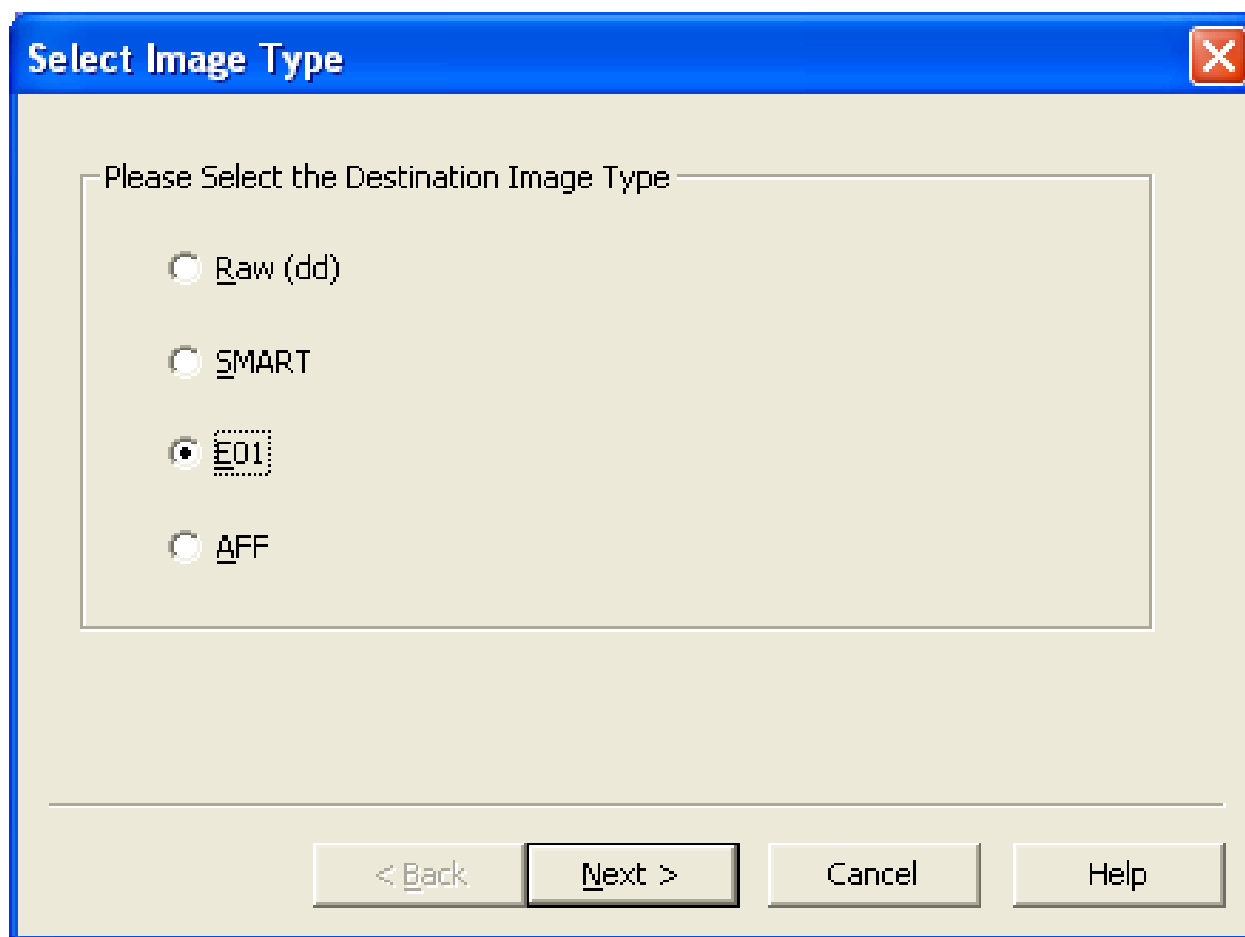
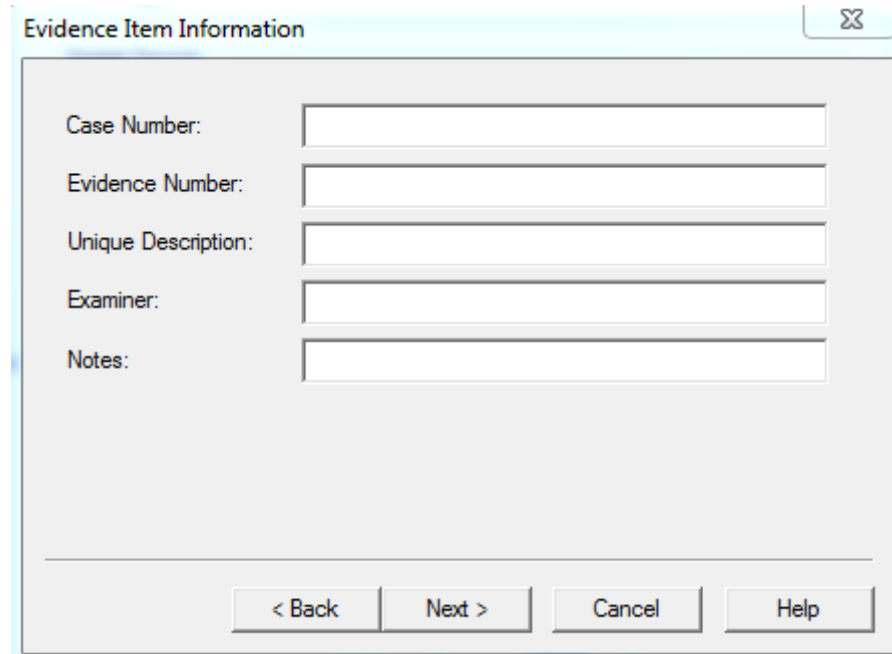


Image Type

- Raw (dd) – commonly used disk image format created by the UNIX command dd
- SMART – proprietary disk image format created by ASR Data.
- E01 – proprietary disk image format created by EnCase.
- AFF – an open source disk image format, allow encryption of disk image
- AD1 - proprietary logical disk image format created by AccessData.

Evidence Item Information



A screenshot of a software dialog box titled "Evidence Item Information". The dialog box has a light blue title bar with a close button (X) in the top right corner. The main area is light gray and contains five input fields, each with a label to its left: "Case Number:", "Evidence Number:", "Unique Description:", "Examiner:", and "Notes:". Each label is followed by a white rectangular text input box. At the bottom of the dialog box, there is a horizontal line separating the input area from the button area. Below the line are four buttons: "< Back", "Next >", "Cancel", and "Help".

Evidence Item Information

Case Number:

Evidence Number:

Unique Description:

Examiner:

Notes:

< Back Next > Cancel Help

Evidence Item Information

- Use “Raw (dd)” as image type
- Use call no. as “Case Number”
- Use CMxxx as “Evidence Number”
- Put your name as “Examiner”
- Use the following in the “Notes” field:
 - 3.5 inch floppy disk; 5.25 inch floppy disk; Zip disk;
 - External hard disk; Internal hard disk;
 - Optical Disk

Select Image Destination

Select Image Destination

Image Destination Folder
C:\Images from Imager 2.9\AD1 Image Browse

Image Filename (Excluding Extension)
Mantooth 1

Image Fragment Size (MB) 1500
For Raw and E01 formats: 0 = do not fragment

Compression (0=None, 1=Fastest, ..., 9=Smallest) 6

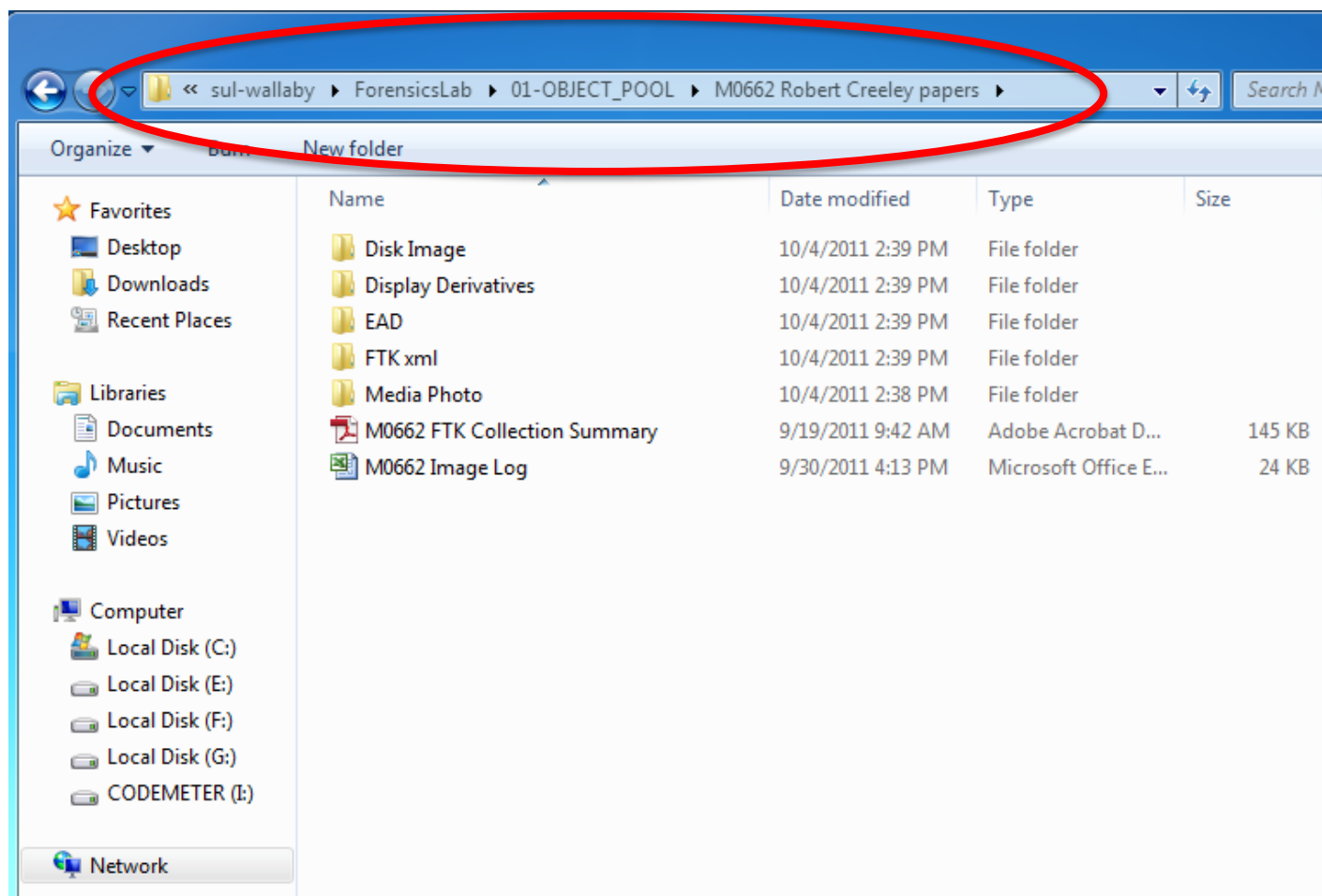
Use AD Encryption ☒


< Back Finish Cancel Help

Select Image Destination

- Store all files under “Disk Image” folder in Desktop
- Use “Case No._CMxxx” as Image Filename for disk image
- Default Image Fragment Size = 1500 MB
- To save images segments that can be burned to a CD, specify 650 MB.
- To save image segments that can be burned to a DVD, specify 4000 MB.
- Selecting 0 (zero) produces the largest file, with no compression.

Storage Locations for Files



 Record 3 of 3

Accessions

Title: Rebecca Solnit papers - captured data from portable computers

Accession #: MSS.2012-056

Basic Information

Accession Notes

User defined fields

Names & Subjects

Acknowledgements, Restrictions & Processing Tasks

Accession Number

MSS

2012-056

Accession Date

2012-4-2

Resources Linked to this accession

Resource Identifier ▲	Resource Title
M1839	Rebecca Solnit papers

Link Resource Remove Link

Resource Type

Title

Rebecca Solnit papers - captured data from portable computers

Physical Description

Primary Extent

Alternate Extent

Extent Number

68.54

gigabyte(s)

Container Summary

megabytes: 72,570
number of files: 732,812
captured from: 3 portable computers

Repository SPECCOLL

Date Expression

Date

Begin

End

Bulk Dates

Begin

End

Deaccessions

Deaccession Date	Extent	Extent Measurement
------------------	--------	--------------------

Add Deaccession Remove Deaccession

Locations

Location	Note
----------	------

Add Location Remove Location

Accession Notes

related ACCN - MSS.2011-119.
Closed until processed.

First

Previous

Next

Last

Reports

Close

Save

Created: Apr 2, 2012 by pchan3 | Modified: Jul 5, 2012 by pchan3 | Record Number: 33677

AT Accession Record

External Documents

Accessions

Record 1 of 9

Title: Stephen Jay Gould papers

Accession #: MSS.2004-148

Basic Information | Accession Notes | User defined fields | Names & Subjects | Acknowledgements, Restrictions & Processing Tasks

method: Gift

contact/office

related collections

Preservation/conservation note

Inventory

Deaccession note

External Documents

First Previous Next Last Reports Close Save

Created: Dec 16, 2008 by amorgan | Modified: Aug 10, 2011 by jejohns1 | Record Number: 22095

Href	Type	Title	Content
\\sul-wallaby\Special Colle...	External Document	Deaccessions List	\\sul-wallaby...
\\sul-wallaby\Special Colle...	External Document	FTK Accession Report	\\sul-wallaby...
\\sul-wallaby\Special Colle...	External Document	Computer Media Imaging Log	\\sul-wallaby...
\\sul-wallaby\Special Colle...	External Document	IM1437 Gould_Accns 2004-148--2007-248_MASTER survey.xls	\\sul-wallaby...

External Documents Location

- \\sul-wallaby\Special Collections\Manuscripts\Collections\M1437 Gould
- M1437 FTK Accession Report
- M1437 Computer Media Imaging Log

Problems in Capturing Floppy Disk

- Sophos under Win 7 will claim the completion of scanning a floppy disk even though it don't recognize the file format.
- FTK Imager under Win 7 will claim the imaging of a floppy disk successful even though it don't recognize the file format.

Floppy Disk Capture

- Virus Check
 - Make sure the system you use can understand the filesystem of the floppy disk by double click the floppy disk

Ensure Accruate Virus Check

- List the directory of the floppy disk
 - Double click on the floppy disk drive

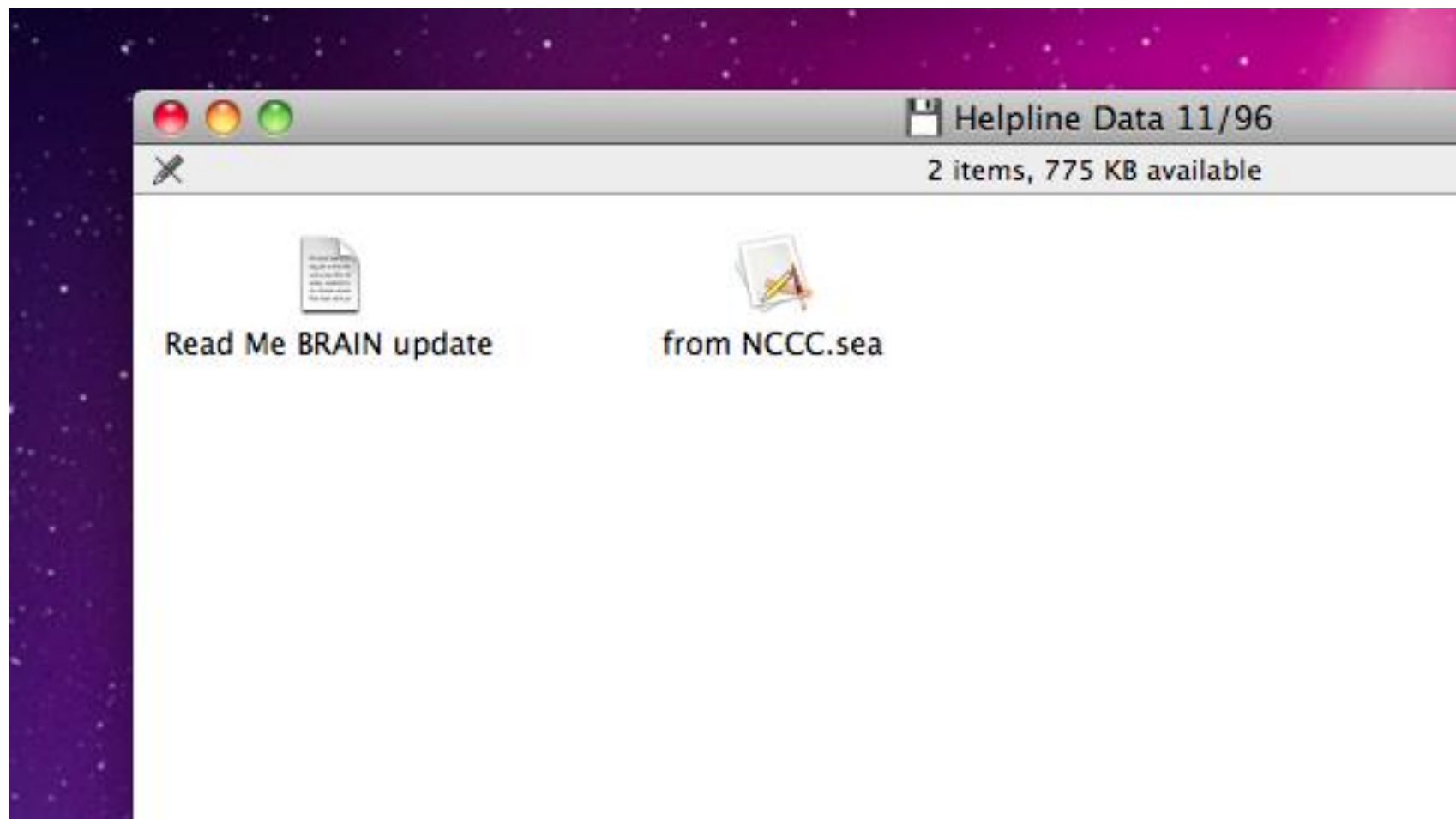
Floppy Disk Capture



FTK



FTK



Virus Scan

- Open Sophos Antivirs
- “+” Custom Scan
- Start Scan

FTK Imager

- click on "Terminal" icon
- `cd ~/Desktop` `diskutil list` (find disk identifier that corresponds to NAME of disk)
- `diskutil unmountDisk /dev/disk1` (match device found from diskutil list)
- `./ftkimager /dev/disk1 floppy1.dd --verify`

Outsourcing

- I have never worked with the following vendors. Don't know the quality of the service.
- Punch Cards (<http://punchcardreader.com/>)
 - 6.5¢ per card + \$3 setup fee + return postage
- 9-track Tape (<http://www.emaglink.com>)
 - \$150 per tape plus the output media and shipping. \$25 for a DVD.

Get Hard drives out from Mac

- PowerBook G4
 - <http://www.ifixit.com/Guide/PowerBook-G4-Aluminum-12-Inch-867-MHz-Hard-Drive-Replacement/208/1>
- Macintosh PowerBook 165c
 - <http://www.ifixit.com/Guide/Installing-Macintosh-PowerBook-165c-Hard-Drive/7022/1>

Take Home

- How to identify a variety of storage media, both obsolete and current: floppy diskettes (3.5 and 5.25 inch), computer tapes, optical disks, and hard drives.
- An introduction to computer storage system interfaces: IDE, SATA, SCSI (HDI, 40 pin, 68 pin, and xx pin connectors).
- An introduction to file systems, including: File Allocation Table (FAT), New Technology File System (NTFS), and Hierarchical File System (HFS)
- An introduction to the difference between logical and forensic images.
- How to create an image log using spreadsheet software.
- How to create basic accession records in Archivists' Toolkit.

Take Home

- How to install FTK Imager for Windows (free/low cost software), and how to create logical and disk images using Access Data FTK Imager software.
- How to image disks using write blocker (acquisition of information on a drive without creating the possibility of accidentally damaging the drive contents). -How to view files using FTK Imager.
- How to use the command line FTK Imager for Mac files.
- Demonstrations will include: how to remove a hard drive from a desktop computer, as well as how to remove a hard drive from an external hard drive enclosure. Peter will lead a discussion about outsourcing (tapes, punch cards, data recovery).