

Digital Forensics: Advanced

Instructors: Christopher (Cal) Lee University of North Carolina at Chapel Hill

Kam Woods University of North Carolina at Chapel Hill

December 13-14, 2021



SOCIETY OF AMERICAN ARCHIVISTS

About These Slides

Authors

Cal Lee, Kam Woods

Description

These are the slides from Cal Lee and Kam Woods's "Advanced Digital Forensics" class. There are a number of hands-on exercises included. The sample data referenced in these slides is available here: <u>https://github.com/BitCurator/bcc-dfa-sample-data/</u>

Learning object type

Lesson plan/materials

Learning objectives

This learning object might be used in a lesson to satisfy the following learning objectives:

- Identify the appropriate tools to: safely acquire born-digital materials from storage media and other modes of transfer; assist in the appraisal of born-digital materials; scan for sensitive information in born-digital materials; and package born-digital materials for preservation and access.
- Practice using tools in the BitCurator Environment.

Digital Archives Specialist (DAS)



Curriculum and Certification Program offered by SAA:

- Foundational Courses—must pass 4
- <u>Tactical and Strategic</u> Courses—*must pass 3*
- <u>Tools and Services</u> Courses—*must pass 1*
- <u>Transformational</u> Courses—*must pass 1*
- Course examinations are administered online.

DAS Core Competencies Addressed

- Understand the nature of records in electronic form, including the functions of various storage media, the nature of system dependence, and the effect on integrity of records over time.
- Formulate strategies and tactics for appraising, describing, managing, organizing, and preserving digital archives.
- Integrate technologies, tools, software, and media within existing functions for appraising, capturing, preserving, and providing access to digital collections.
- Curate, store, and retrieve original masters and access copies of digital archives.

https://www2.archivists.org/prof-education/das-curriculum-structure

Agenda

Day 1

- Welcome and introductions
- Motivation and overview
- Technical fundamentals
- Data acquisition considerations
- Potential elements of your own digital forensics lab
- Bit-level treatment of individual files
- Creating and extracting forensic metadata
- BitCurator reporting features
- Other BitCurator environment tools
- Preview of Day 2

Day 2

- Day 1 Postmortem Questions, Concerns and Insights
- Command-line operations in Linux
- FIDO as an example of a command-line tool
- Regular expressions
- Extracting data from specific types of files: images, office files, email
- Windows artifacts (including the Registry)
- End user access (logistics and technical approaches)
- Incorporating digital forensics into archival workflows
- Challenges, Ethical/legal issues, and donor agreements
- Wrap up and evaluations

Personal Introductions

- Who's teaching you?
- What about you?
 - Who are you (name, institution, job title)?
 - Why are you here (relevance to job, what you hope to get out of the workshop)?
 - What have you done so far to apply digital forensics methods in your institution?

Software You Should Have Installed for the Exercises

- VirtualBox
- VirtualBox Extensions
- BitCurator Virtual Machine
- Exiftool (or on the web via https://exif.tools/)
- Sample data: <u>https://distro.ibiblio.org/bitcurator/samples/saa-dfa-sample-data.zip</u>
- Visit the link above and download this now if you have not already done so!
- Additional tools for Windows users:
 - FTK Imager (Windows only)
 - OSFMount (Windows only)
 - RegRipper (Windows GUI, can also run at the command line in BitCurator environment)

Discussion Scenario

- You've been charged with taking care of data from a prominent community leader who has died unexpectedly
- Her materials include some paper and lots of digital data (on floppies, CDs, and a laptop hard drive)
- What should you do with the floppies?
- CDs?
- Hard drive?

Goals When Acquiring Born-Digital Materials

- Ensure integrity of materials
- Allow users to make sense of materials and understand their context
- Prevent inadvertent disclosure of sensitive data

Fundamental Archival Principles to Apply

Provenance

- Reflect "life history" of records
- Records from a common origin or source should be managed together as an aggregate unit
- Original Order Organize and manage records in ways that reflect their arrangement within the creation/use environment

Chain of Custody

- "Succession of offices or persons who have held materials from the moment they were created"¹
- Ideal recordkeeping system would provide "an unblemished line of responsible custody"²
- 1. Pearce-Moses, Richard. *A Glossary of Archival and Records Terminology*. Chicago, IL: Society of American Archivists, 2005.
- 2. Hilary Jenkinson, *A Manual of Archive Administration: Including the Problems of War Archives and Archive Making* (Oxford: Clarendon Press, 1922), 11.

Digital Forensics Can Help Archivists to Fulfill Their Principles

Provenance

- Identify, extract and save essential information about context of creation
- Original Order Reflect original folder structures, files associations, related applications and user accounts

Chain of Custody

Identifying Sensitive Information

- Documentation of how records were acquired and any transformations to them
- Use well-established hardware and software mechanisms to ensure that data haven't been changed inadvertently
- Identify personally identifying information, regardless of where it appears
- Flag for removal, redaction, closure or restriction

More Product, More (Machine) Processes

- Archivists need to apply many <u>more</u> processes to born-digital records (e.g. integrity checks, metadata extraction, audit trails, characterization)
- The good news is that most of these processes can be performed by <u>software</u>

Digital Forensics in Archives

- In recent years, archivists have been applying various digital forensics methods, for example:
 - use of write blockers
 - generation of disk images
 - applying cryptographic hashes to files
 - capture of Digital Forensics XML (DFXML)
 - scanning bitstreams for personally identifying

Need for Adaptation of Digital Forensics Tools and Tasks for Archivists

- Existing digital forensics tools provide valuable functionality, but they don't always fit well into primary workflows of archives.
- For example, archives are particularly concerned with:
 - structure and persistence of metadata
 - provisions for providing public access to data
 - support for older technologies (e.g. floppy disks, HFS)

From Bitstreams to Heritage:

Putting Digital Forensics into Practice in Collecting Institutions



Christopher A. Lee, Kam Woods, Matthew Kirschenbaum, and Alexandra Chassanoff

https://bitcurator.net/wp-content/uploads/sites/1099/2018/08/bitstreams-to-heritage.pdf

After this class, you should be able to:

- Install and operate the BitCurator environment as a virtual machine in VirtualBox
- Explain and recognize different types of metadata stored in common filesystems
- Identify file types based on magic numbers (file signatures)
- Determine potential hardware options for acquiring data from various types of storage media
- Apply common Linux commands at the command line and compose basic regular expressions
- Evaluate disk image format options based on needs and priorities of your institution and collections
- Generate BitCurator reports and use bulk_extractor to identify potentially sensitive data
- Extract and interpret EXIF metadata
- Capture and analyze Windows Registry artifacts using RegRipper
- Determine essential points in your institution's workflows where it will be beneficial to incorporate forensics tools and methods
- Make and justify decisions of professional ethics that emerge when caring for born-digital records
- Recognize technical strategies for providing access

Caveats and Such

- Advanced doesn't mean "everything we didn't cover in the Fundamental class"
- There's much more about digital forensics that we won't be addressing
- Selective hands-on experience with specific applications
- A license to learn more in the future

Digital Resources - Levels of Representation*

Level	Label	Explanation		
8	Aggregation of objects	Set of objects that form an aggregation that is meaningful		
		encountered as an entity		
7	Object or package	Object composed of multiple files, each of which could also		
		be encountered as individual files		
6	In-application rendering	As rendered and encountered within a specific application		
5	File through filesystem	Files encountered as discrete set of items with associate		
		paths and file names		
4	File as "raw" bitstream	Bitstream encountered as a continuous series of binary		
		values		
3	Sub-file data structure	Discrete "chunk" of data that is part of a larger file		
2	Bitstream through I/O	Series of 1s and 0s as accessed from the storage media		
	equipment	using input/output hardware and software (e.g. controllers,		
		drivers, ports, connectors)		
1	Raw signal stream through	Stream of magnetic flux transitions or other analog electronic		
	I/O equipment	output read from the drive without yet interpreting the signal		
		stream as a set of discrete values (i.e. not treated as a digital		
		bitstream that can be directly read by the host computer)		
0	Bitstream on physical	Physical properties of the storage medium that are		
	medium	interpreted as bitstreams at Level 1		

*Covered in Fundamental class. See also: Lee, Christopher A. "<u>Digital Curation as Communication Mediation</u>." In Handbook of Technical Communication, edited by Alexander Mehler, Laurent Romary, and Dafydd Gibbon, 507-530. Berlin: Mouton De Gruyter, 2012. **18**

Digital Resources - Levels of Representation*

Level	Label	Explanation				
8	Aggregation of objects	Set of objects that form an aggregation that is meaningful				
		encountered as an entity				
7	Object or package	Object composed of multiple files, each of which could also				
		be encountered as individual files				
6	In-application rendering	As rendered and encountered within a specific application				
5	File through filesystem	Files end Levels where digital pciate				
		paths and				
4	File as "raw" bitstream	Bitstream forensics methods and hary				
		to als can provide a lat				
3	Sub-file data structure	Discrete LOOIS Can provide a lot				
2	Bitstream through I/O	Series of of assistance nedia				
	equipment	using inp				
		drivers, ports, connectors)				
1	Raw signal stream through	Stream of magnetic flux transitions or other analog electronic output read from the drive without yet interpreting the signal stream as a set of discrete values (i.e. not treated as a digital bitstream that can be directly read by the host computer)				
	I/O equipment					
0	Bitstream on physical	Physical properties of the storage medium that are				
	medium	interpreted as bitstreams at Level 1				

*Covered in Fundamental class. See also: Lee, Christopher A. "<u>Digital Curation as Communication Mediation</u>." In Handbook of Technical Communication, edited by Alexander Mehler, Laurent Romary, and Dafydd Gibbon, 507-530. Berlin: Mouton De Gruyter, 2012. **19**

BitCurator

- Funded by Andrew W. Mellon Foundation
 - Phase 1: October 1, 2011 September 30, 2013
 - Phase 2 October 1, 2013 September 30, 2014
- Partners: School of Information and Library Science (SILS) at UNC and Maryland Institute for Technology in the Humanities (MITH)

BitCurator Goals

- Develop a system for collecting professionals that incorporates the functionality of open source digital forensics tools
- Address two fundamental needs not usually addressed by the digital forensics industry:
 - Incorporation into the workflow of archives/library ingest and collection management environments
 - Provision of public access to the data

BitCurator Environment*

- Bundles, integrates and extends functionality of open source software
- Can be run as:
 - Self-contained environment (based on Ubuntu Linux) running directly on a computer (download installation ISO)
 - Using "bootstrapping" installation scripts to turn any Ubuntu Linux machine into a BitCurator Environment
 - Self-contained Linux environment in a virtual machine using e.g.
 VirtualBox or VMWare
 - As individual components run directly in your own
 Linux environment or (whenever possible) Windows environment

*To read about and download the environment, see: https://bitcurator.net/

Hands-On Familiarization with VirtualBox and the BitCurator VM*

*For a detailed walk-through of the steps we're following, see the Quickstart Guide:

https://github.com/BitCurator/bitcurator-distro/wiki/Releases#quickstart-guide



BitCurator Resources



January 3, 2019 / Uncategorized

BitCurator 2.0.10 released

A new production release of BitCurator (2.0.10) is now available at the BitCurator release portal.

You can download the VirtualBox VM and installation ISO directly using the following links:

workflows, and other resources can be found on the BitCurator Consortiumhosted Confluence site:

BitCurator Consortium Confluence Site

Get help and ask questions on our Google Group, follow us on Twitter, view our screencasts on YouTube, follow us on Facebook, and browse our code on GitHub. Get the software Documentation and technical specifications Screencasts Google Group People Project overview Publications News

https://bitcurator.net/



Most tasks we will cover in this course are explained in the Quick Start Guide



https://github.com/BitCurator/bitcurator-distro/wiki/Releases#quickstart-guide

From Code to Community: Building and Sustaining BitCurator through Community Engagement



https://bitcurator.net/wp-content/uploads/sites/1099/2018/08/code-to-community.pdf

BitCurator Consortium

- Continuing home for hosting, stewardship and support of BitCurator tools and associated user engagement
- Administrative home: Educopia Institute
- Funding based on membership dues
- Software and documentation are free and open source, but membership provides benefits (e.g. support, training, development priority)

BitCurater CONSORTIUM

About Us + Why Digital Forensics + Using BitCurator + Get Involved +



Membership is open to libraries, archives, museums, and other institutions worldwide that seek a collaborative community within which they may explore and apply forensics approaches and solutions to their digital collections.

Become a member now >

How to Use BitCurator

- Acquire and process digital collections.
- · Maintain the original order of digital materials.
- · Survey the extent and composition of digital collections.
- · Redact personally identifiable information.
- Extract technical and preservation metadata.
- Package digital materials for archival storage.

Learn more about getting started.

How our members are using BitCurator

Member Benefits

- Use of the members-only BCC mailing list and help desk
- Access to the members-only videos and documentation
- Prioritized requests for BitCurator feature development
- Opportunities to serve on the BCC committees
- Voting rights for community governance
- Professional development opportunities
- Discounts for events including the BitCurator User Forum

Members

McMaster University Penn State University Massachusetts Institute of Technology Duke University The University of Maryland, MITH Stanford University Yale University The University of Manchester Library 29 University of

al 1 a 11

Technical Fundamentals

Checksums – Compact Representations of Bitstreams

- A given bitstream, fed into an algorithm, will generate a short string of characters that is <u>extremely</u> unlikely to be generated by a different bistream fed into that same algorithm
- Most common = MD5, SHA-1
- Can determine:
 - □ If bits have changed after a transfer
 - If bits have flipped within a storage environment
 - Whether two different files are identical bitstreams
- A library of hash values can identify "known and notable" (EnCase terminology) files
 - Known files that can be ignored (e.g. software listed in National Software Reference Library)
 - Notable specific bitstreams that you're trying to find

In BitCurator environment: Right Click on File or Directory and Calculate MD5

Pictures					🚨 🖾 I	■ 1 •) 8:26 AM & BitC	urator 🔱
	Computer	• 🙍 Home	Pictures			🔶 🥧 (Q Search
	 Home Desktop Documents Downloads Music Pictures 	bc400px-1280full. bitcurat		Biturater Itor 🔍 Open With 🥑 Open With	Image Viewer Firefox Web Browser Shotwell Photo Viewer		
	Videos			Open With	Other Application		
	💁 File System			Scripts		Disk Image Info	•
	Trash Network			Cut Copy	Encrypt-Decrypt Show File Details	File Analysis FileInfo Find Files	•
	Browse Net			Make Link Rename	Show File Type Show File Type (mime)	Open Browser as Root Safe Mount	
2.45				Move to	View in Hexeditor	Open Scripts Folder	
				Move to Tra	ash		
				Revert to P Compress Send To	revious Version 		
ل_				Properties			
						"bitcurator-grub.png" selected	(43.3 kB)

Calc

6

Calculace	e MD5 (Files and Dire	eccories)			₽ T ↓ 4)))	8:31 AM I BI	ccurator Q
	Computer Home Desktop Couments Downloads Music Pictures Videos File System Trash Network Browse Net	• Mome	Pictures	itor-grub- w.png		-	Q Search
			 Calculate MD5 (Files and Director) Please choose the way you want the MD (1 file(s) selected) Handling Display on screen Save to file (the selected filename 4 	ries) •5 hash to be presented • .md5 extenion) Cancel OK	t		

Calculate MD5 (Files and Directories) 図 🖾 1 (4)) 8:30 AM L BitCurator 役 Computer B Home Pictures 🔶 🤿 🔍 Search ÷. \odot Home Desktop BitCurator Documents or Constant bc400px-1280full. bitcurator-grub.png bitcurator-grub-Downloads png new.png Music Pictures Videos P File System Trash Network Calculate MD5 (Files and Directories) Browse Net... Į The MD5 hash of the selected file: ceb2622125be1231b0fc9babee27942d /home/bcadmin/Pictures/bitcurator-grub.png 240 Cancel OK >_

29

1

°D |

Note on MD5/SHA1 - Potential Collisions

- From a security perspective, MD5 has been "broken" since 2005
- SHA-1 was broken in February 2017
- Someone with malicious intent can create two different bitstreams that result in the same hash (i.e. hash collisions)

Hash Collisions (The Poisoned Message Attack): "The Story of Alice and her Boss"*

Being an intern, Alice does not have any access to secret documents. Not enough for her ...

... tricky Alice decides to fool Caesar. Because Caesar is still relying on the widely used MD5 hash function, she implements the attack from Wang and Yu [WY05] to find MD5 collisions. When she receives her letter of recommendation (on paper), she prepares **two postscript files with the same MD5 hash**:

- · One to display the letter of recommendation, and
- a second one, an order from Caesar to grant Alice some kind of a security clearance.



*Stefan Lucks and Magnus Daum,

http://th.informatik.uni-mannheim.de/people/lucks/HashCollisions/

Wayback link:

https://web.archive.org/web/20160713130211/http://th.informatik.uni-mannheim.de/people/lucks/Has
Julius. Caesar Via Appia 1 Rome, The Roman Empire	Julius. Caesar Via Appia 1 Rome, The Roman Empire
May, 22, 2005 Order: Alice Falbala is given full access to all confidential and secret information about GAUL.	May, 22, 2005 To Whom it May Concern: Alice Falbala fulfilled all the requirements of the Roman Empire intern position. She was excellent at translating roman into her gaul native language, learned very rapidly, and worked with considerable independence and confidence.
Sincerely, Julius Caesar	Her basic work habits such as punctuality, interpersonal deportment, communication skills, and completing assigned and self-determined goals were all excellent. I recommend Alice for challenging positions in which creativity, reliability, and language skills are required.
	I highly recommend hiring her. If you'd like to discuss her attributes in more detail, please don't hesitate to contact me. Sincerely, Julius Caesar

Alternatives to MD5?

SHA (Secure Hash Algorithm)

- Originally developed by the NSA
- Several variants: SHA-0, SHA-1, SHA-2 family
- Early variants (SHA-0, SHA-1) known to be compromised
- Most commonly used is now SHA-256. Can be used to process bitstreams ("messages") up to 3.4 * 10³⁸ bits (very large!)
- Disadvantage: more time, computing power required to produce

Algorith	hm and variant	Output size (bits)	Internal state size (bits)	Block size (bits)	Max message size (bits)	Word size (bits)	Rounds	Operations	Collisions found?	
	SHA-0		540 off4 d					Yes		
	SHA-1	160	160	512	21	32	80	add, and, or, xor, rotate, mod	Theoretical attack (260)[6]	
	SHA-256/224	256/224	256	512	2 ⁶⁴ – 1	32	64			
SHA-2	SHA-512/384	512/384	512	1024	2 ¹²⁸ – 1	64	80	add, and, or, xor, rotate, mod, snift	NO	

Source: https://en.wikipedia.org/wiki/Cryptographic_hash_function

Implications of MD5/SHA1 Being "Broken"

- Rarely a concern when hash is used for integrity checks on known items (e.g. verifying that a file was transferred correctly to a repository or that files in storage are still intact)
- Can be a concern if one is relying on a hash as proof of record authenticity – risks can include cases of internal tampering
- There are more robust hash algorithms to address this (SHA-2 family, including SHA-256) so good practice is to generate one of them along with the MD5
- MD5 is still widely used, because it is fast to calculate and still widely supported

Question:

Can you use a cryptographic hash to determine specifically what any given file **contains**?

If not, what could you use?

File System

- Access controls
- File names & identifiers
- File size (length)
- Where to find files in storage (sectors and clusters)
- MAC times
 - Modified when the content was last changed
 - Accessed time file was last accessed (by person or software)
 - Changed last time metadata changed
 - Created (implemented inconsistently, if at all, across different file systems)





Name	Operating System(s) Using it as Native File System [often other OSs can also recognize it]
FAT12, FAT16	MS-DOS
FAT32 (VFAT)	Windows 95, 98
exFAT	Windows XP SP2 and later (primary use: USB drives, SD cards)
NTFS	Windows NT, 2000, XP, Server 2003, Server 2008, Vista
MFS	Macintosh System 1-3
HFS (Hierarchical File System)	Macintosh System 4-8
HFS+	Macintosh System 8.1 – 9, OS X 10.0 – 10.11
APFS	macOS 10.12
ext, ext2, ext3, ext4 (Extended File System)	Linux
XFS	Linux, typically Enterprise variants (RHEL)
HPFS (High Performance File System)	OS/2
ISOFS (ISO 9660)	Any OS that reads data from a CD
JFS1 (Journaled File System)	AIX (IBM)
ReiserFS	Several Linux distributions
UFS (Unix File System) aka FFS (Fast File System)	Various flavors of Unix

Name

FAT12, FAT16	MS-DOS
FAT32 (VFAT)	Windows 95, 98
exFAT	
NTFS	Hesystems vou're
MFS	
HFS (Hierarchical File System)	Mmoot likely to
HFS+	
APFS	m
ext, ext2, ext3, ext4 (Extended File System)	encounter
XFS	Li
HPFS (High Performance File System)	OS/2
ISOFS (ISO 9660)	Any OS that reads data from a CD
JFS1 (Journaled File System)	AIX (IBM)
ReiserFS	Several Linux distributions
UFS (Unix File System) aka FFS (Fast File System)	Various flavors of Unix

Operating System(s) Using it as Native File System [often other OSs can also recognize it]

NTFS vs. FAT File System Attributes

- Two disk images are in your zip file, and can also be found at: <u>https://digitalcorpora.s3.amazonaws.com/corpora/scenarios</u> <u>/2009-m57-patents/usb/terry-work-usb-2009-12-11.E01</u> <u>https://digitalcorpora.s3.amazonaws.com/corpora/drives/np</u> <u>s-2009-ntfs1/ntfs1-gen1.E01</u>
- Load each disk image into a separate instance of FTK Imager (run them side by side to compare what you see) – if you don't have a Windows computer, look on with a partner
- Look at the properties of some files*
- What differences do you notice?

*Properties are shown in the bottom left corner. If you don't see them, go to the View menu at the top and select "Properties." You may need to drag the top of the properties window up to see all of the values.

Connecting a Device vs. Mounting a Filesystem

HFS+ volume is visible through Windows Disk Management after it's connected

() → (+ Control Panel + Al	I Control Panel Items 🕨 Adm	inistrative Tool	ls				•	++ Searc	h Administr	ative Tool	- 0	x Q				
Organize 🕶 🔲 Open											- 61	0				
Envoriter Name	*	Da	te modified	Туре	S	ize				- 8		-				
Desktop 👘 Cor	mponent Services	7/1	4/2009 12:57 AM	Shortcut		2 KB										
📕 Downloads 🛛 🛃 Cor	nputer Management	7/1	4/2009 12:54 AM	Shortcut		2 KB										
🖳 Recent Places 🛛 🛃 Compu	ter Management					1.00										X
File Act	ion View Help															
📕 callee	2 🖬 🛛 🖬 🖸 🗙 🖆	f 🖻 🔍 📓														
Comp	uter Management (Local Volu	ume l	Layout Type File	System	Status Healthy (EET Sy	rtem Partition)		Capacity	Free Space	% Free	Fault To	lerance	Overhead	Actions		
Documents	Task Scheduler		Simple Basic	í	Healthy (Prima	ary Partition)		14.21 GB	14.21 GB	100 %	No		0%	Disk Manager	nent	<u> </u>
Music	Event Viewer	ystem Vindows (C:)	Simple Basic NT Simple Basic NT	FS I FS I	Healthy (Syste Healthy (Boot.	m, Active, Primary Partition) Page File, Crash Dump, Primary	/ Partition)	200 MB 238.28 GB	168 MB 106.43 GB	84 % 45 %	No No		0% 0%	More Acti	ons	100
Pictures	Local Users and Groups		•													
Subversion Difference Subversion	Performance Device Manager															
A 🤤 Sto	rage															
📲 Computer 🔰 🛼 Ser	Vices and Applications															
Windows (C:)																
SILS Users (I:)																
.																
Network	< [m										
		Disk 0														
	Bas 238	sic 3.47 GB	system 200 MB NTFS			Windows (C:) 238.28 GB NTFS										
Computer N	Onl	line	Healthy (System.	Active Prin	nan Dediti	Handler (David David File Crach	Dump Prin	mary Partitio	on)							
Shortcut																
	Ren	'Disk 1 movable	2//////////////////////////////////////										E			
	14.5 Onl	53 GB line	200 MB Healthy (EFI Syste	em Partitio	14.21 GB Healthy (Pri	mary Partition)	1	L28 MB Unallocated								
				//////					5							
	_0	CD-ROM 0											_			
	DVI	D (D:)														
	No	Media														
										_	_		•			
			Primary partition													

But it's not visible through Windows Explorer, because Windows doesn't know how to mount the file system

Organize 🔻 🛛 System	properties Uninstall or cha	ange a program	» 🗄		
🛧 Favorites 📑	Name	Туре	Total Size	Free Space	
🛜 Libraries	 Hard Disk Drives (1) Windows (C:) 	Local Disk	238 GB	106 GB	
🖳 Computer	Devices with Remov	able Storage	(1)		
🏭 Windows (C:)	BD-ROM Drive (D:)	CD Drive			
🚽 My SILS Home (F	A Network Location (2)	2)			
🖵 SILS Users (I:)	🛒 My SILS Home (H:)	Network Dri	912 GB	164 GB	
🗣 Network	SILS Users (I:)	Network Dri	912 GB	164 GB	

Seeing Attached Devices (Whether or Not They're Mounted) - In Windows

- Control Panel
- Administrative Tools
- Computer Management
- Disk Management



			Manage	Aurninistrative loois				
le Home S	hare	View	Shortcut Tools					~
→ 丶 ♠ «	All C	Control Pan	el Items > Administrat	tive Tools 🛛 🗸 🗸	ō			1
		Name	^	Date mod	fied	Туре	Size	
Quick access		🥙 Comr	onent Services	12/7/2019	1:09 AM	Shortcut	2 KB	
🔜 Desktop	*	Com	outer Management	12/7/2019	1.09 AM	Shortcut	2 KB	
🕹 Downloads	*	Detra	ament and Ontimize Dri	ives 12/7/2019	1:09 AM	Shortcut	2 KB	
Documents	A	deskt		4/9/2021 6	:53 AM	Configuration sett	3 KB	
E Pictures	*	T Disk (Cleanup	12/7/2019	1:09 AM	Shortcut	2 KB	
Music		Event	Viewer	12/7/2019	1:09 AM	Shortcut	2 KB	
Videos		iscsi	Initiator	12/7/2019 1:09 AM Shortcut				
Videos		Local	Security Policy	12/7/2019	1:10 AM	Shortcut	2 KB	
OneDrive		DDBC	Data Sources (32-bit)	12/7/2019	1:10 AM	Shortcut	2 KB	
This PC			Data Sources (64-bit)	12/7/2019	12/7/2019 1:09 AM Shortcut			
E mist c		Perfo	rmance Monitor	12/7/2019	1:09 AM	Shortcut	2 KB	
Network		汩 Print	Management	12/6/2019	1:46 PM	Shortcut	2 KB	
		Recov	very Drive	12/7/2019	1:09 AM	Shortcut	2 KB	
		式 Regist	try Editor	12/7/2019	1:09 AM	Shortcut	2 KB	
		🔊 Resou	irce Monitor	12/7/2019	1:09 AM	Shortcut	2 KB	
		Service	es	12/7/2019	1:09 AM	Shortcut	2 KB	
		🔝 System	m Configuration	12/7/2019	1:09 AM	Shortcut	2 KB	
		👰 Syster	m Information	12/7/2019	1:09 AM	Shortcut	2 KB	
		🕖 Task S	cheduler	12/7/2019	1:09 AM	Shortcut	2 KB	
		🔗 Wind	ows Defender Firewall w	vith Advanc 12/7/2019	1:08 AM	Shortcut	2 KB	
		📷 Wind	ows Memory Diagnostic	c 12/7/2019	1:09 AM	Shortcut	2 KB	



Data Acquisition Considerations

Ports and Connectors

- Adoption comes and goes with changes in the industry
- Those that have had wide industry adoption have usually lasted a decade or more
- Important distinction: hardware protocol standards vs. shape of connectors (related but not always the same)



Different USB connectors. From left to right: male Micro USB B-Type, UC-E6 proprietary (not USB), male Mini USB (5-pin) B-type, female A-type, male A-type, male B-type. Shown with a centimeter ruler. Female A-type connector (4th from left) is "upside down" to show the pins.

https://en.wikipedia.org/wiki/File:Usb_connectors.JPG

Adapter Examples

Micro SD to SD



MicroSATA to SATA



NVMe to USB



SATA to IDE



Ethernet to USB



SATA to Molex Power



A Note Regarding Modern Laptops (and other devices)

Some recent laptops, including all Apple laptops post-2019, the Microsoft Surface series of machines, and certain others, have both RAM and storage soldered to the motherboard; these drives cannot* be removed for imaging.



2 x 128GB Western Digital storage chips, Apple Macbook M1 Air laptop motherboard (2020)



*Except in specialized laboratories.

128GB Toshiba storage chip,

Microsoft Surface Laptop

See <u>https://sumuri.com/how-to-image-an-apple-silicon-mac-with-recon-itr-live/</u> for an example of a commercial tool used to image Apple Silicon Macs without drive removal.

M.2 (also known as "Next Generation Form Factor")

- Current most common way to connect a mass storage device (e.g. solid state storage) to the inside of many desktop and laptop computers. Cable-less, connects via an edge slot.
- Internal M.2 devices mount to the motherboard or a PCI expansion card. External M.2 devices typically housed in USB enclosures



512GB M.2 Mass Storage Device https://en.wikipedia.org/wiki/M.2#/media/File:Intel_512G_ M2_Solid_State_Drive.jpg



M.2 connector on a motherboard https://en.wikipedia.org/wiki/M.2#/media/File:M.2_conne ctor_on_a_computer_motherboard.jpg

Serial Advanced Technology Attachment (SATA)

 The most common way to connect a mass storage device (e.g. hard drive, solid state drive) to the inside of a computer from 2003 through the mid-2010's.



Data Cables: SATA on the left, eSATA on the right https://en.wikipedia.org/wiki/File:SATA2_und_eSATA-Stecker.jpg

 eSATA used for external mass storage devices



SATA Power Cable

Integrated Drive Electronics (IDE) / AT Attach

- Precursor to SATA
- No longer widely used, but you'll still find it on older drives
- More than one drive can share the same cable, with one being the *master* (device 0) and the other being the *slave* (device 1) – these are set with *jumpers*





https://static.daniweb.com/images/attachments/0/maxtor_jumper.jp

40-pin IDE Ribbon Cable https://en.wikipedia.org/wiki/File:Nappe.sv

n

Floppy Disks

- Physical storage is similar to hard drives (magnetic charges in a spinning disk)
- Various types and sizes, e.g. high density, double density, 3.5 inch, 5.25 inch, 8 inch
- 3.5 inch floppies are relatively easy to read using a USB drive, but older ones are more complicated...

Floppy Controller Hardware



FluxEngine²



Disc Ferret³



FC 5025⁴



- 1. https://www.kryoflux.com/
- 2. http://cowlark.com/fluxengine/index.html
- 3. http://discferret.com/wiki/DiscFerret
- 4. http://www.deviceside.com/fc5025.html
- 5. http://disk2fdi.joguin.com/D2FCABLE.htm
- 6. http://www.cbmstuff.com/proddetail.php?prod=SCP

Disk2FDI⁵



SuperCard Pro⁶



Common Floppy Formats (Physical)

Many variations over time, often to increase storage density

Size	Density	Tracks	tpi	bpi	Coercivity	Unformatted capacity per side
2 ¹ / ₂ -inch ^{[16][17]}	Single	16 ^{[16][17]}	48 ^[16]			64 KB ^{[16][17]}
	D 11 [18]	40 ^[18]	67.5 ^[18]	8650 ^[18]	600 Oe	250 KB
	Double	80	135	8717	600-665 Oe	500 KB
3 ¹ /2-inch	High	80	135	17434	720-750 Oe	1000 KB
Extended Triple ^[12]	Extended	80	135	34868	900 Oe	2000 KB
	Triple ^[12]	240 ^[11]	406.5 ^[11]	36700 ^[11]		6500 KB
	Single/Double	40	48	5876	300 Oe	250 KB
	Double	80	62.5			(Apple FileWare)
5 ¹ / ₄ -inch	Quad	77	100		300 Oe	500 KB (Micropolis-compatible)
	Quad	80	96	5922	300 Oe	500 KB
	High	80	96	9646	600 Oe	833 KB
8-inch	Single/Double	77	48		300 Oe	1000 KB

Floppy disk physical characteristics

(capacity and tracks are nominal, per side)

Table source: <u>https://en.wikipedia.org/wiki/List_of_floppy_disk_formats</u>

*Coercivity is how resistant the medium is to being demagnetized (larger number means it requires a larger magnetic field to be demagnetized) ⁶⁴

Actual bits on most IBM PC-type floppies are encoded using "modified frequency modulation" (MFM)

	Cinala	129	26	-	1	250.25 KB ^{[NB 13][19][20][21]}			
	9 ipob	Single	120	20	77	2	500.5 KB ^{[NB 13][19][20][21]}	260	MEN
	8-Inch	Daubla	1024			1	616 KB ^{[NB 13][20][21]}	- 300	
		Double	1024	ö		2	1232 KB ^{[NB 13][19][20][21]}		
				0		1	160 KB ^[NB 13]		
		Dauble		ð	40	2	320 KB ^[NB 13]		
		Double		-		1	180 KB ^[NB 13]	300	
5 ¹ ⁄ ₄ -inch		512	9		2	360 KB ^[NB 13]		MFM	
	Our JINB 151	_		00	1	320 KB ^[NB 13]			
IBM PC		Quad	21	0	00	2	640 KB ^[NB 13]	- 300	
companioles		High		15	80	2	1200 KB ^[NB 13]	360	
				8			320 KB ^[NB 13]	 300	
		Daubla		9		1	360 KB ^[NB 13]		
		Double		8			640 KB ^[NB 13]		
				9	80		720 KB ^[NB 13]		
	37 ₂ -inch (90 mm)		- 512	18	-		1440 KB ^[NB 13]		MEM
	High		01		2	1680 KB ^[NB 13]			
				21	82		1720 KB ^[NB 13]		
	Extended		36	80		2880 KB ^[NB 13]			

Table source: <u>https://en.wikipedia.org/wiki/List_of_floppy_disk_formats</u>

Actual bits on Apple (Macintosh and earlier) may be encoded using "group code recording" (GCR) or "modified frequency modulation" (MFM)

	5 ¹ ⁄ ₄ -inch	Double	256	13	25	1	113.75 KB	200	CCP
		Double	200	16	35	1	140 KB	300	GCR
Apple II		Doublo	510	Variable (8-	90	1	400 KB	394 -	CCD
	3 ¹ / ₂ -inch (90 mm)	Double	512	12)	00	2	800 KB	590	GCR
		High	512	18	80	2	1440 KB	300	MFM
Apple Lisa	$5\frac{1}{4}$ -inch FileWare	Double	512	Variable (15- 22)	46	2	851 KB	218 - 320	GCR
Apple Lisa 2/Macintosh XL		Double	512	Variable (8-	80	1	400 KB	394 -	GCR
3 ² / ₂ -inch (90 mm) Apple Macintosh	1)		12)		2	800 KB	590		
	High	512	18	80	2	1440 KB	300	MFM	

Table source: <u>https://en.wikipedia.org/wiki/List_of_floppy_disk_formats</u>

Issues with Common Floppy Formats

- Why is this important?
 - Floppy disk drives write the actual bitstream using either GCR or MFM (or other, less common encodings)
 - Drive won't read disks encoded in a non-supported format without specialized hardware (e.g. KryoFlux)
 - Other factors:
 - Some devices (Apple in particular) use "CLV" (constant linear velocity), adjusting rotation speed depending on where the head is writing/reading
 - Others vary angular velocity of the disk based on zones ("zoned CAV").
 - Again, need specialized hardware such as Kryoflux to read these disks on a modern drive

Additional Factors: File Systems Used on Floppies

- Common IBM PC compatible floppies:
 - Typically FAT12 with 512 byte clusters (although there are many less common variations)

Attribute	FAT12
Used For	Floppies and very small hard disk volumes
Size of Each FAT Entry	12 bits
Maximum Number of Clusters	4,086
Cluster Size Used	0.5 KB to 4 KB
Maximum Volume Size	16,736,256 bytes

- Common Apple floppies (Macintosh and previous)
 - Apple II: 13 sector disk (5.25", ProDOS 3.2, 113.75K)
 - Apple II: 16 sector disk (5.25", ProDOS 3.3, 140K)
 - Apple Macintosh:
 - Double-density 3.5" Macintosh File System (MFS)
 - Double-density 3.5" Hierarchical File System (HFS)
 - High-density 3.5" Hierarchical File System (HFS)

Potential Problems with ISO 9660 Media (e.g. CD-ROMs)

Factor	Complications
Physical damage	May not be visible to the naked eye
Bad length in volume header	Older CD writing tools sometimes miscalculated sector count, so header metadata doesn't match actual length
Incorrect block size	CDs can misidentify as having 512-byte sectors
File truncation	Filesystem may allow you to navigate to files that subsequently appear damaged or won't render at all (files could be truncated or never fully written to disk)
One sector short when written Track at Once (TAO)	TAO disks often represent a length (in volume header) that is one sector short from the actual length

Not all tools are designed to recognize or address these issues

CD-ROM File System Structure



Note that there are structures, such as the Path Table, that are not generally used, but may contain metadata. For more info on issues in managing optical media, see: https://kamwoods.net/publications/woodsbrownarch09.pdf 70

Creating Disk Images of CDs

- Cdrdao primarily for ripping audio CDs (addressing issues such as the TAO one discussed above)
- Often sufficient to use dd form CD-ROMs
 - In the BitCurator VM, CD-ROM drive should appear as a device called /dev/sr0
 - Command to acquire: bcadmin@ubuntu:~\$ dd if=/dev/sr0 of=FILENAME1.iso

Dealing with Disk Images from CDs

To modify the file system, you can:

- Mount the disk image
- Use mkisofs to create a new ISO file system bcadmin@ubuntu:~\$ mkisofs -r -o FILENAME2.iso /media/sr0
- If you created an EWF image of a CD:
 - BitCurator mounting scripts can't determine whether the underlying disk image data is a raw (dd) image or an ISO
 - To mount the image, first use ewfexport (command-line tool) to pull out the raw (dd) image, then rename it as an ISO
Two Important Considerations for Internal Media that are Used as External Media

- Power internal drive needs different connector (often Molex), not the kind that plugs into the wall
- Cooling when pulled from the computer, you've also separated the drive from the fan, so you should often add an external one to ensure cooling



https://en.wikipedia.org/wiki/File:Molex_female_connector.jpg



https://www.tigerdirect.com/applications/SearchTools/item -details.asp?EdpNo=1648567

Kryoflux installed and running in a mini jukebox



*Adapted from a Mini JukeBox setup designed by the National Library of Australia

Write Blocking – One-Way Street for Data

- Ensures that data can be read from the device, but no bits can be changed
- Doesn't just prevent changes conscious made by user but also changes made by the system
- Options for write blocking (in order of most to least certain to prevent writes to the drive):
 - Dedicated write blockers
 - Writing blocking tabs or settings on the device itself
 - Software-based write blocking

Image source: http://thinng.com/1555-one-way-sign-seat



Dedicated Hardware Write Blockers









A WiebeTech SATA/IDE write-blocker



This end connects to your computer using USB 2.0/3.0, eSATA, or FireWire. The cables are in the box.



This end connects to the drive you want to image using SATA or IDE. There are power cables in the box that can connect to either drive.



This end powers the write blocker itself. There's a power supply in the box.



The WiebeTech Combodock V5 supports both USB 2.0 and USB 3.0. USB 3 is faster, but may not always be compatible with all computers.



The ComboDock has been connected to an IDE drive, a host computer, and powered on. It will allow you to select write-blocked or read/write. "Enter" enables write-blocking.



The drive is now fully powered, and you can use the dock to examine some metadata...



This drive has just been powered on, and it's registering as 14 degrees Celsius (57 Fahrenheit). This is well below room temperature, but it will quickly rise without a fan.



This drive is indicating a raw capacity of 156 thousand megabytes (about 150GB).



The name of the manufacturer is embedded in the metadata. Why might this be important?



The model number is next. This number is *not* unique, but common to many drives.



The serial number, however, *is* unique.



Above is a USB write blocker manufactured by Tableau.



This end has a power connector for the blocker, and a USB 'b' port to connect to the host. The FireWire port is just for updating the write blocker firmware. It has no other use.



This end connects to the USB device. This particular write-blocker will *not* recognize USB floppy drives, only USB flash drives and USB hard disks. And look, a power switch.



Here's what it looks like when everything is plugged in. Plug everything in before turning the write-blocker on.



Once you turn it on, the "Host Detect" and other lights will light up. If "Host Detect" doesn't light up, the write blocker can't see your computer. Something has gone wrong.



Different write blockers may expose different metadata. This write blocker displays the product name, in addition to the manufacturer.

Host-Based Write Blockers

FRED UltraBay

- The "UltraBay" on the FRED provides write blocking for a range of interfaces.
 - USB (top left, next to power switch)
 - SCSI (right of USB) [no longer included in newer FREDs]
 - PATA/IDE (below SCSI)
 - SATA (left of IDE)
- It also includes a MOLEX power connector. There's a cable in the toolbox that converts this to SATA power, if required.
- The ports in the white box on the right (top picture) are NOT write-blocked.

In earlier FREDs:



On newer FREDs:



5.25 Inch Floppy – If light can get through, it's **not** write protected



https://en.wikipedia.org/wiki/File:Floppy_tabs_3x2.jpg

3.5 Inch Floppy – If light can get through, it **is** write protected



Source: <u>http://www.techmint.info/2009/09/security-write-protecting-floppy-disks.html</u> Current:

https://web.archive.org/web/20100125050630/http://www.techmint.info/2009/09/security-write-protecting-flop py-disks.html

Example of Software Write Blocking – Mounted Devices set to Read-Only by Default

			🔍 📜 🖪 💌 🕪) 2:44 PM 🔱
Ó			
	home	Documentation and Help	
0	Imaging Tools		
<u>-</u> ا		Network servers	
	Forensics Tools	😵 🖯 🕕 System Mount Policy	
A		CAUTION! You are about to set the system-wide mount policy to:	
1	Accession Tools	WRITEABLE	
		Click CANCEL to remain in the READ-ONLY state. Currently mounted volumes will not be affected until remounted.	
	Additional Tools	Cancel OK	
9	<u></u>		-
	Shared Folders and Media	BitCurat ∉r	
	sf_bc_share		
	Trash		
-			

Other Potential System Changes to Reduce Risk of Writing to File Systems

- Macintosh: Disk Arbitrator "will block the mounting of file systems to avoid mounting as read-write and violating the integrity of the evidence" <u>https://github.com/aburgh/Disk-Arbitrator</u>
- Windows: See "Digital Forensics: How to configure Windows Investigative Workstations" <u>https://www.sans.org/blog/digital-forensics-how-to-</u> <u>configure-windows-investigative-workstations/</u>

Potential Elements of your own Digital Forensics Lab

FRED Options from Digital Intelligence



https://digitalintelligence.com/products/overview

Mini Jukebox

Configuration #2 for Manuscripts

3 x Double DVD Drive units (MyBorg-006> MyBorg-007)

• 1 x Storage Tower (black) with 4SATA Multilane connector installed ST4SAML-B, 4 bay aluminium unit

• 2 x Plextor PX-800A DVD-RW Super Multi Drive Drives - Drive Letter Mapping W & X

• 1 x 3.5" floppy disk Drive Letter Mapping A or B

• 1 x Western Digital WD5001ABYS CAVIAR RE2/ 500GB Hard Drive Letter Mapping Z

"Prometheus Component Installation Guide Pre-install requirements and methodology." National Library of Australia. October 22, 2008.

ATA

Stanford University Libraries and Academic Information Resources (SULAIR)



The British Library (UK)



School of Information and Library Science at UNC Chapel Hill, North Carolina



USB 3.5" Floppy Disk Drive

Still available new from online retailers, look for a drive that can read both 1.44 MB(HD) and 800 KB (DD) 3.5" diskettes. Most drives support HD diskettes in both PC and Mac format, but only support PC formatted DD diskettes. New units are still available for around S20.

External USB 250MB Zip Drive

These units are available both new and used. We recommend the 250MB model as it is backwards compatible with the 100MB Zip disks. New units retail for around \$200 and used units for around \$50. BitCurater

Device Side Data's FC5025

The FC5025 is a controller card for 5.25" floppy disk drives that can be used as an internal or external—as seen here interface. Device Side Data charges \$55.25 per controller.

5.25" Floppy Disk Drive

These units are no longer available new, but can still be purchased off of eBay for about \$50. We recommend purchasing a number of drives as well as a floppy disk drive cleaning kit.

Wiebetech UltraDock Hardware Write Protector

This unit serves as both an interface with IDE and Serial ATA type hard disk drives and as a write protector. Because it is common for the OS to overwrite metadata on a hard drive, write protection ensures that no interactions of the archivist or researcher affects the integrity of the original media. Wiebetech charges \$250 for the UltraDock Hardware Write Protector.

https://bitcurator.net/2013/08/02/building-a-digital-curation-workstation-with-bitcurator -update/ 108
Outfitting a Born-Digital Archives Program (Ben Goldman, Penn State University)



https://practicaltechnologyforarchives.org/issue2_goldman/

Useful Resource - Mediapedia



https://mediapedia.nla.gov.au/

Creating Exact Copies of Data from Media – Disk Images

- Getting an "image" of a storage medium involves working at a level below the file system
- Can get at file attributes and deleted files not visible through higher-level copy operations

Creating a Disk Image in Guymager

Image: Computer Image: Computer Image: Computer Image: Computer	ition and
File format	
Devices Misc He C Linux dd raw image (file extension .dd or .xxx) Image files Rescan Expert Witness Format, sub-format Guymager (file extension .Exx) Split image files C Advanced forensic image (file extension .aff) Split size 2047 MiB Image files	
Serial nr. Serial Case number I Evidence number 1 Examiner BitCurator User	Progre
VB45b1d326-9557 Description A sample floppy disk image Notes Additional notes go here Destination	
Image directory /home/bcadmin/Desktop/SampleData/ Image filename (without extension) sampleimage Info filename (without extension) sampleimage	Ŀ
Size Size Sector size Image file Info file Calculate MD5 Current speed Re-read source after acquisition for verification (takes twice as long) Source verification ✓ Verify image after acquisition (takes twice as long)	
Cancel Duplicate image Start	

Examples of Disk Image Formats

- RAW and Split RAW (RAW stored across multiple files)
- Advanced Forensics Format (AFF) [no longer recommended]
- EnCase Evidence File (.E01)
- ISO (for CD-ROM)
- IMG (floppy or sometimes CD-ROM)

RAW (dd)

- Copies of the raw media data. Often split into smaller chunks to make them more manageable and so that the resulting images can fit onto limited filesystems and media such as FAT or DVD/CDROM.
- Advantages:
 - Very simple, use simple tools to manipulate the image.
 - Image can be easily split for storage and transport on removable media
 - Output can be piped to other applications for immediate processing
- Disadvantages:
 - Can be very large (no compression). Zipped raw images cannot be operated on directly with regular tools (efficiently perform arbitrary seeks).
 - Often too large to store on FAT formatted media
 - No metadata other than filenames, no hashes.
 - No checksumming on files not robust
 - Missing segments (for example from scratched CD/DVD can sometimes be overwritten with 0's).
 - Overwritten data (unrecoverable no checksums on small blocks in file).

Expert Witness Format (Encase)

- Evidence file consists (in order) of: Acquisition information, Data Block, CRC (cyclic redundancy check), acquisition hash (MD5)
- Can be split for storage, transport
- CRC computed for every 32K block; balance between integrity and speed, also makes it very difficult to tamper with the evidence file (1 in 4 billion chance of collision)
- Cannot be manipulated with simple (open source UNIX) tools; support reverse engineered in libewf
- Previously limited to 2GB size
- Largely proprietary
- Has been reverse engineered by Joachim Metz in libewf (used in open source tools that read EWF) - <u>https://github.com/libyal/libewf</u>



ISO Images (.iso extension) for CD-ROM and DVD

- Similar to raw, but can't contain
 - multiple tracks
 - audio or video tracks
- Don't contain control headers or error correction fields (raw can include these)
- Filesystem usually will be either ISO 9660 (CD-ROM) or UDF (DVDs)

Accessing Data in Disk Images

- Virtualization and emulation
- Mounting the original filesystem
- Accessing (but not mounting) disk images using forensics software
- Two options discussed later for end user access:
 - Remote, dynamic access to disk image contents
 - Cross-drive analysis

Emulation as a Service



https://eaasi.gitlab.io/program_docs/intro-emulation-workshop/06-eaasi/index.html 118

What's the difference between the two options shown in FTK Imager below?

Q	AccessData FTK Imager 3.1.3.2				Q AccessData FTK Imager 3.1.3.2			x
Ē	le <u>V</u> iew <u>M</u> ode <u>H</u> elp g <u>A</u> dd Evidence Item		◎ ★ ★ ★ ★ ● ● ● ● ● ● ● ● ● ● ● ● ● ● ●		Add Evidence Item		1 10 12 12 12 12 12 12 12 12 12 12 12 12 12	
1	Add All All All All All All All All All	File List		×		File List		×
1	Image Mounting	Nos	Size Type	Date Modified	🔓 Image Mou <u>n</u> ting	Name	Size Type Date Mod	lified
	Remove Evidence Item				🚔 Remove Evidence Item			
相	Remove All Evidence Items				Remove All Evidence Items			
G	<u>C</u> reate Disk Image				🖸 Create Disk Image			
5	Export Disk Image				Export Disk Image			
E	Export Logical Image (AD1)				Export Logical Image (AD1)			
4	Add to Custom Content Image (AD <u>1</u>)				Add to Custom Content Image (AD1)			
G	Create Custom Content Image (AD1)				🖾 Create Custom Content Image (AD1)			
	Decrypt AD1 image				Decrypt AD1 image			
1	<u>V</u> erify Drive/Image							
	Capture Memory				Capture Memory			
1	Obtain Protected Files	< [III	•	Detain Protected Files	•	m	•
9	Detect EFS Encryption				Contract EFS Encryption			
	Export Eiles				Export <u>Files</u>			
E	Export File Hash List				Export File Hash List			
E	Export Directory Listing				Export Directory Listing			
	E <u>x</u> it				Exit			

Mounting a Disk Image to Browse the Contents

🔔 📬 En 📧 🕪 2:58 PM 🔱



Mounting a Disk Image to Browse the Contents

Image: Status Image: Sta							
Amme Sigle_cshare buf_exerciseplaning winxp_feports WinXPEDI Besktop Documents Videos Trash Connectto Server Computer Connectto Server State Computer State Computer Secont Sec		X = D wieVB					
ging Tool Geneent Geneent Geneent <th>home</th> <th>() Inst he share</th> <th>huf exercisenlanning winXP</th> <th></th> <th></th> <th></th> <th></th>	home	() Inst he share	huf exercisenlanning winXP				
indication indication <th></th> <th></th> <th></th> <th></th> <th></th> <th></th> <th></th>							
Image: Note winxp_teports WinXPED Image: Desktop winxp_teports Image: Desktop winxp_teports Image: Downloads Image: Downloads Image: Downloads Im	maging Tools	Recent		10 101 1010			
Image: State of the second			winxp_beout winxp_re	ports WinXP.E	01		
naries Tools ↓ Downloads ↓ Downloads ↓ Wideos ↓ Trash ↓ Ocuments ↓ Config.Msi ↓ Config.Msi ↓ Documents and ↓ Config.Msi ↓ Documents and ↓ Config.Msi ↓ Documents and ↓ Documents ↓ Docume		Documents					
Image: Sign Tools Image: Sign Tools Image: Sign Tools <th>Forensics Tools</th> <th>🕹 Downloads</th> <th></th> <th></th> <th></th> <th></th> <th></th>	Forensics Tools	🕹 Downloads					
ssion Tools • Pictures • Videos • Trash • Sf_bc_share • Desktop • Computer • Downloads • Pictures • Videos • Trash • Desktop • Downloads • Pictures • Videos • Downloads • Pictures • Videos • Desktop • Downloads • Pictures • Videos • V		J Music					
 Videos Trash Sf_bc_share Network Computer Connect to Server Connect to Server Setures Videos Trash Sf_bc_share Connect to Server Setures Network Setures Setur	Accession Tool	D Pictures					
Image: Trash Image: Trash <th></th> <th>🗯 Videos</th> <th>WINXP_VOID2</th> <th></th> <th></th> <th></th> <th></th>		🗯 Videos	WINXP_VOID2				
construction sf_bc_share Program Files Program F		🛅 Trash	< > ■1.6 GB Volu	me			
Image: Section of the sec	Addicional lool	🔳 sf_bc_share 🔄	⊘ Recent				
I - 1.6 GB Volume Computer Connect to Server Media Connect to Server Music Pictures Videos Trash AUTOEXEC.BAT SBadClus SBitmap boc.ini I Trash AUTOEXEC.BAT SBadclus SBitmap Jobulan I Trash AUTOEXEC.BAT SBadclus SBitmap Jobulan I Trash AUTOEXEC.BAT SBoot CONFIG.SYS Io.SYS SLogFile Jobulan Jobulan I Trash AUTOEXEC.BAT SBoot CONFIG.SYS Io.SYS SLogFile Jobulan Jobulan Jobulan I Trash AUTOEXEC.BAT SBoot CONFIG.SYS Io.SYS SLogFile Jobulan Jobulan <		⊡ Network	☆ Home	Config.Msi	Documents and	\$Extend	Program Files
Media Computer Connect to Server Image: Documents Image: Documents <	Shared Folders a	📕 1.6 GB Volume 🔺	Desktop		Settings		
Connect to Server Music RECYCLER System Volume WINDOWS \$AttrDef Music Pictures Image: Connect to Server <	media	Computer	Documents				1 10 101
bc_share Pictures Videos Imformation Trash AUTOEXEC.BAT \$BadClus \$Bitmap bc_share Imformation Imformation Imformation <		👤 Connect to Server	A Music	RECYCLER	System Volume	WINDOWS	\$AttrDef
Image: Videos Image: Trash Image: Sf_bc_share	sf_bc_share		Pictures		Information		
Image: Trash AUTOEXEC.BAT \$BadClus \$Bitmap bot.ini Image: Sf_bc_share			Videos			10 101 1010	
Image: sf_bc_share Image: sf			🗊 Trash	AUTOEXEC.BAT	\$BadClus	\$Bitmap	boot.ini
Image: Computer Image: Computer <td< th=""><th></th><th></th><th>sf_bc_share</th><th>1 10 101</th><th></th><th></th><th>1 10 101</th></td<>			sf_bc_share	1 10 101			1 10 101
Trash			🗗 Network	\$Boot	CONFIG.SYS	IO.SYS	\$LogFile
Computer Mail	Trash		💷 1.6 GB Volume 📤	1		1 20	10
SMFTMIRE MSDOS.SYS NTDETECT.COM pagefile.sys			Computer	101 1010		101 1010	101 1010
			Connect to Server	SMFTMILL	MSDOS.SYS	NIDETECT.COM	pagerile.sys
JB Volume	1.6 GB Volume						
\$Secure \$UpCase \$Volume				\$Secure	\$UpCase	\$Volume	

💿 💿 🗬 🖉 🚍 💻 🔛 🔘 🐼 🚱 Right Ctrl

Exporting Selected Files from a Disk Image

J 🖉 💟
ect All Export selections Cancel export
Image Info 4442dcc6-a2a6-8a42-9137-fbe3be4ae9a2
Media information Media type: fixed disk Is physical: yes Bytes per sector: 512 Number of sectors: 4096000 Media size: 1.9 GiB (2097152000 bytes) Digest hash information MD5: e07f26954b23db1a44dfd28ecd717da9
Messages
/home/bcadmin/Desktop/disk-images/terry-work- usb-2009-12-11.E01 >> Generating DFXML file /home/bcadmin/.bcfa/terry-work- usb-2009-12-11.E01_dfxml.xml
>> Success!!! Fiwalk created DFXML file
>> Generating directory tree

Exercise: Multiple Views into Disk Image Files

- Resources we'll be using (also in the zip file you downloaded):
 - 1. ISO file <u>http://www.ils.unc.edu/callee/25.iso</u>
 - 2. IMG file <u>https://distro.ibiblio.org/bitcurator/lab/something.img</u>
 - 3. OSFMount (Windows only)
 - 4. FTK Imager (Windows only)
 - 5. BitCurator Environment
- Step 1 Mount the ISO and IMG files using OSFMount
- Step 2 Find the drives using Windows Explorer and investigate their contents
- Step 3 Open FTK Imager and add both images as evidence items, and explore what we see in the drives
- Step 4 Use the **BitCurator environment** to mount the disk images [Right click on image file, then select: Scripts > Mount Disk Image]
- Step 5 Use the BitCurator environment to select files within the images to export [Use Forensics Tools > BitCurator Disk Image Access]

Bit-Level Treatment of Individual Files

Hex Dump

- A more compact and more humanly readable way of conveying a stream of bits
- Uses hexadecimal notation
 - Each character represents one of 16 possible values (0,1,2,3,4,5,6,7,8,9,A,B,C,D,E,F)
 - Conveniently, a series of two characters represented in hexadecimal can represent exactly one byte (2⁸ = 256 possible values) of data, because 16² = 256
- Hex dumps from computer's memory often used for debugging or reverse engineering software and for data recovery

Hex Dump Tools

- Many free or inexpensive tools available for download, e.g. Cygnus Hex Editor, Hex Workshop, HexAssistant, MiniDumper, Hex Fiend (Mac), GHex (Linux)*
- BitCurator environment has a built-in hex editor (GHex)

Online tool: <u>https://hexed.it/</u>

* See https://en.wikipedia.org/wiki/Comparison_of_hex_editors

In the BitCurator environment:



"bitcurator-grub.png" selected (43.3 kB)

											3:07 PM 🔱
0	× - 🗆 🖡	😸 😑 😐 BitCur	ator-Basic-720px-	2016.png - GHex							
	< >	000000089	50 4E 47 0D	0A 1A 0A 00 0	0 00 0D 49 48	44 52 PNG.	THDR				
	-	0000001000	00 0C D4 00	00 02 18 08 0	6 00 00 00 1F	71 50	qP		ation		
-	O Recent	0000002039	00 00 00 06	62 4B 47 44 0	0 FF 00 FF 00	FF A09b	KGD		P		
	🔂 Home	00000030BD	A7 93 00 00	00 09 70 48 5	9 73 00 00 0B	13 00	.pHYs				
ا 🕑	🖿 Deskt	0000004000	0B 13 01 00	9A 9C 18 00 0 2B 20 73 E3 4	B 87 00 00 00 00	40 45	TIME		DVOF6		
		0000006054	58 74 43 6F	6D 6D 65 6E 7	4 00 00 00 00	00 43TXtCom	mentC		IVEIS		
>_	J. Down	0000007072	65 61 74 65	64 20 77 69 7	4 68 20 47 49	4D 50 reated	I with GIMP				
		0000008064	2E 65 07 00	00 20 00 49 4	4 41 54 78 DA	EC DDd.e	.IDATx				
A	dd Music	0000009079	B8 65 69 59	DF FD EF A9 1	E 69 66 BA 19	44 99y.eiY.	ifD.				
	D Pictur	0000000000000	47 40 50 C0 D4 28 14 34	I7 C4 29 I2 0 RE E6 D2 98 3	8 26 8F F0 04	51.60P	.)8 DE.Q				
	Video:	0000000013	06 51 79 9D	20 A8 20 A0 4	1 14 64 92 19	94 A6Qy.	A.d				
	🗊 Trash	000000D069	81 16 BA 1B	19 9A 9E 9B E	E AA F7 8F 7D	FA A21					
	sf bc	000000E02C	BB CA AA AE	7D CE BD 87 C	F E7 BA EE 6B	9F AA,}	·k				
	D Netwo	000000000000	53 BF BD 9E	F5 AC BD D6 5	E FB DE CF 4E	00 00.5	[^] N				
	En Herme	00000110B7	DC AD C3 7F	BE 69 75 66 7	5 93 1B 78 3C	AD 3A	iufux<.:				
	Comp	00000120E5	28 75 A8 BA	AE BA 76 F7 F	1 F0 9F AF AE	AE 38.(u	.v8				
	🛛 Conne	000001304A	5D 52 7D A0	FA E0 EE E3 E	1 F5 DE EA A2	EA 9AJ]R}			r		
		00000140E9	0D 08 00 00	00 00 00 00 0 20 00 00 00 0	0 B0 6D 76 A6	03 00	mv		1		
		0000015000	D1 3C 73 DB	16 8D 32 FB F		0B 5Bh <s< td=""><td>2 [</td><td></td><td></td><td></td><td></td></s<>	2 [
		0000017034	D7 5C B8 5B	EF AA DE 59 B	D 63 B7 2E 9E	0E 094.\.[.	Y.c				
		0000018000	00 00 00 00	00 00 B0 49 D	6 F9 C3 26 00	00 00	<mark>I</mark> &	1.8 kB)			
	-	0000019000	7A 3B BD BA	47 75 B7 A3 D	4 AD A6 03 AE	90 AB.z;G	iu				
		000001A0FA	48 83 CD B9	BB F5 D6 DD C	7 F3 5B AC 96	03 00.H	l				
		000001C0DB	55 1F 5F 7D	EC 11 75 B7 E	A 94 E9 70 1B	E0 C3.U. }.	. U D				
	6	000001D02D	9A 6A DE 5A	BD B1 7A C3 6	E FD 5D 75 CD	74 38j.Z.	.z.n.]u.t8				
		000001E000	00 00 00 00	00 00 80 55 A	4 A1 06 00 00	00 58	UX				
	Ŧ			2							
		Signed 8 bit:	-119	Signed 32 bit:	1196314761	Hexadecimal:	89				
		Unsigned 8 bit:	137	Unsigned 32 bit:	1196314761	Octal:	211				
		Signed 16 bit:	20617	Signed 64 bit:	1196314761	Binary:	10001001				
		Unsigned 16 bit:	20617	Unsigned 64 bit:	1196314761	Stream Length:	8 - +				
		Float 32 bit:	5.281654e+04	Float 64 bit:	5.292398e-260						
-		Show	v little endian deco	ding	Show unsigned	and float as hex	kadecimal				
		Offset: 0x0									
								-			🔟 🚫 💽 Right Ctrl

Identifying File Types

- Magic numbers and file signatures
- File extensions
- Metadata stored in file system
- MIME types

Magic Numbers and File Signatures

- Distinct string or pattern that is found within files of a given type (most often in the header)
- Most effective searches for magic numbers often involve regular expressions (e.g. grep) in order to indicate multiple variations of a pattern
- Utilities that use this: file (Unix), TrID, DROID, FITS
- Examples:

File Format	Нех	ASCII
DOC	D0 CF 11 E0 A1 B1 1A E1	ÐÏài ±á
JPG	FF D8 FF	ÿØÿ
PDF	25 50 44 46 2D 31 2E	%PDF-1.
ZIP	50 4B 03 04	РК

File Information Tool Set (FITS) https://code.google.com/p/fits



- FITS "identifies, validates, and extracts technical metadata for various file formats. It wraps several third-party open source tools, normalizes and consolidates their output, and reports any errors.
 FITS was created by the Harvard University Library Office for Information Systems for use in its Digital Repository Service (DRS)."
- Tools currently bundled into it:
 - Jhove
 - Exiftool
 - National Library of New Zealand Metadata Extractor
 - DROID
 - FFIdent
 - File Utility (windows)
- FITS may no longer be available as a part of BitCurator due to JRE incompatibilities

Siegfried

http://www.itforarchivists.com/siegfried/

- Signature-based file format identification tool
 - PRONOM file format signatures (National Archives of UK) (default)
 - MIME-info file format signatures (freedesktop.org)
 - FDD file format signatures (Library of Congress)
- Unlike FITS, does not have validation built in, and fewer extraction tools, but much lighter weight.
- Has a lot of customization for output
 - CSV
 - YAML (text)
 - DROID CSV
 - JSON
 - stdout



Drag a file on to Siegfried's anvil!

Brunnhilde

https://github.com/tw4l/brunnhilde

- Reporting companion for Siegfried
 - Requires Siegfried (but running Brunnhilde also runs Siegfried)
 - Command-line and GUI (we'll be using the CLI version later)
- Reports generated
 - HTML (human readable)
 - Siegfried CSV
 - Directory tree
 - Other CSVs extracted from Siegfried logs (e.g., warnings, unidentified files)
- Can run other processes too, but not required
 - Virus scan
 - bulk_extractor
 - Disk image processing

File Extensions

- Changing file extension usually changes default application OS uses to open (i.e. associates with) the file
- The "8.3" (eight characters, followed by three-character extension) limit in the past – based on FAT – resulted in many creative uses of the extension part of the file name (e.g. reports1.994, april-94.rpt)
- Convention is often still to use only three letters
- No authority for standardizing use, so three-letter extensions are often shared by many formats
- Security risks associated with trusting the file extension to be accurate – malicious code masquerading as another type of file (e.g. viruses sent as email attachments)

Manageme	ent-curriculu	m.docx				
		[
Manageme	ent-curriculu	im.docx				
		l –				
Managem	ent-curriculu	um.zip				
		Π				
		₽-				
🗐 WinZip - Manageme	nt-curriculum.zip)				
File Actions Options H	Help					
	12 1	h (77) (78)	भ 🥥	2		Cit.
New Open	Favorites Add	Extract Encry	pt Vie	ew	CheckOut	Wizard
New Open	Favorites Add	Extract Encry	rpt Vie Size	ew Ratio	CheckOut Packed	Wizard
New Open	Favorites Add	Modified [1/1/1980 12:00 AM	rpt Vie Size 590	ew Ratio 59%	CheckOut Packed	Wizard Path _rels\
New Open	Favorites Add Type XML Document XML Document	Modified Extract Encry 1/1/1980 12:00 AM 1/1/1980 12:00 AM	rpt Vie Size 590 1,445	Ratio 59% 74%	CheckOut Packed 243 370	Wizard Path _rels\
New Open Name A .rels [Content_Types].xml app.xml	Favorites Add Type XML Document XML Document XML Document	Modified 1/1/1980 12:00 AM 1/1/1980 12:00 AM 1/1/1980 12:00 AM	rpt Vie Size 590 1,445 1,041	Ratio 59% 74% 50%	CheckOut Packed 243 370 519	Wizard Path _rels\ docProps\
New Open Name Icontent_Types].xml app.xml core.xml	Favorites Add Type XML Document XML Document XML Document XML Document	Modified 1/1/1980 12:00 AM 1/1/1980 12:00 AM 1/1/1980 12:00 AM 1/1/1980 12:00 AM 1/1/1980 12:00 AM	rpt Vie Size 590 1,445 1,041 633	Ratio 59% 74% 50% 48%	CheckOut Packed 243 370 519 331	Wizard Path _rels\ docProps\ docProps\
New Open Name Icontent_Types].xml Icontent_Types].xml Iconce.xml I	Favorites Add Type XML Document XML Document XML Document XML Document XML Document	Modified 1/1/1980 12:00 AM 1/1/1980 12:00 AM 1/1/1980 12:00 AM 1/1/1980 12:00 AM 1/1/1980 12:00 AM 1/1/1980 12:00 AM	rpt Vie 590 1,445 1,041 633 34,242	Ratio 59% 74% 50% 48% 90%	CheckOut Packed 243 370 519 331 3,454	Wizard Path _rels\ docProps\ docProps\ word\
New Open Name Icontent_Types].xml Icontent_Types].xml Icontent.xml Icontent.xml Icontent.xml Icontent.xml.rels	Favorites Add Type XML Document XML Document XML Document XML Document XML Document XML Document XML Document	Modified 1/1/1980 12:00 AM 1/1/1980 12:00 AM 1/1/1980 12:00 AM 1/1/1980 12:00 AM 1/1/1980 12:00 AM 1/1/1980 12:00 AM 1/1/1980 12:00 AM	rpt Vie 590 1,445 1,041 633 34,242 950	Ratio 59% 74% 50% 48% 90% 72%	CheckOut Packed 243 370 519 331 3,454 265	Wizard Path _rels\ docProps\ docProps\ word\ word_rels\
New Open Name Icontent_Types].xml Icontent_Types].xml Icone.xml Ic	Favorites Add Type XML Document XML Document XML Document XML Document XML Document XML Document XML Document XML Document	Modified 1/1/1980 12:00 AM 1/1/1980 12:00 AM	rpt Vie 590 1,445 1,041 633 34,242 950 1,831	Ratio 59% 74% 50% 48% 90% 72% 72%	CheckOut 243 370 519 331 3,454 265 510	Wizard Path _rels\ docProps\ docProps\ word\ word_rels\ word\
New Open Name Content_Types].xml Content_Types].xml Core.xml Cocument.xml Cocument.	Favorites Add Type XML Document XML Document XML Document XML Document XML Document XML Document XML Document XML Document	Extract Encry Modified 1/1/1980 12:00 AM 1/1/1980 12:00 AM 1/1/1980 12:00 AM	rpt Vie 590 1,445 1,041 633 34,242 950 1,831 6,306	Ratio 59% 74% 50% 48% 90% 72% 72% 87%	CheckOut 243 370 519 331 3,454 265 510 845	Wizard Path _rels\ docProps\ docProps\ word\ word_rels\ word\ word\ word\
New Open Name Content_Types].xml Content_Types].xml Core.xml Core.xml Cocument.xml	Favorites Add Type XML Document XML Document XML Document XML Document XML Document XML Document XML Document XML Document XML Document XML Document	Modified 1/1/1980 12:00 AM 1/1/1980 12:00 AM	rpt Vie 590 1,445 1,041 633 34,242 950 1,831 6,306 1,833	Ratio 59% 74% 50% 48% 90% 72% 72% 87% 57%	CheckOut 243 370 519 331 3,454 265 510 845 791	Wizard Path _rels\ docProps\ docProps\ word_rels\ word_rels\ word\ wor
New Open Name Content_Types].xml Content_Types].xml Core.xml Cocument.xml Cocument.	Favorites Add Type XML Document XML Document	Modified 1/1/1980 12:00 AM 1/1/1980 12:00 AM	rpt Vie 590 1,445 1,041 633 34,242 950 1,831 6,306 1,833 15,692	Ratio 59% 74% 50% 48% 90% 72% 87% 87% 87%	CheckOut 243 370 519 331 3,454 265 510 845 791 2,071	Wizard Path _rels\ docProps\ docProps\ word_rels\ word_rels\ word\ wor
New Open Name Content_Types].xml Content_Types].xml Content.xml Cocument.xml Cocume	Favorites Add Type XML Document XML Document	Modified 1/1/1980 12:00 AM 1/1/1980 12:00 AM	rpt Vie 590 1,445 1,041 633 34,242 950 1,831 6,306 1,833 15,692 6,992	Ratio 59% 74% 50% 48% 90% 72% 72% 87% 87% 87% 87% 87%	CheckOut 243 370 519 331 3,454 265 510 845 791 2,071 1,686	Wizard Path _rels\ docProps\ docProps\ word_rels\ word_rels\ word\ wor
New Open Name Content_Types].xml Content_Types].xml Content.xml Co	Favorites Add Type XML Document XML Document	Modified 1/1/1980 12:00 AM 1/1/1980 12:00 AM	rpt Vie 590 1,445 1,041 633 34,242 950 1,831 6,306 1,833 15,692 6,992 260	Ratio 59% 74% 50% 48% 90% 72% 72% 87% 57% 87% 87% 76% 28%	CheckOut 243 370 519 331 3,454 265 510 845 791 2,071 1,686 187	Wizard Path _rels\ docProps\ docProps\ word_rels\ word_rels\ word\ wor
New Open Name Content_Types].xml Content_Types].xml Content_Types].xml Content.xml Content	Favorites Add Type XML Document XML Document	Extract Encry Modified 1/1/1980 12:00 AM 1/1/1980 12:00 AM 1/1/1980 12:00 AM	rpt Vie 590 1,445 1,041 633 34,242 950 1,831 6,306 1,833 15,692 6,992 260	Ratio 59% 74% 50% 48% 90% 72% 87% 87% 87% 87% 87% 28%	CheckOut 243 370 519 331 3,454 265 510 845 791 2,071 1,686 187	Wizard Path _rels\ docProps\ docProps\ word_rels\ word_rels\ word\ wor

15 KB Microsoft Office Wo... 10/4/2008 11:28 AM

15 KB Microsoft Office Wo... 10/4/2008 11:28 AM

15 KB WinZip File

10/4/2008 11:28 AM

1	xml version="1.0" encoding="UTF-8" standalone="yes"?
2 🗢	<w:document <="" th="" xmins:ve="http://schemas.openxmlformats.org/markup-compatibility/2006"></w:document>
з	xmins:o="urn:schemas-microsoft-com:office:office"
4	xmlns:r="http://schemas.openxmlformats.org/officeDocument/2006/relationships"
5	xmins:m="http://schemas.openxmlformats.org/officeDocument/2006/math"
6	xmlns:v="urn:schemas-microsoft-com:vml"
7	xmlns:wp="http://schemas.openxmlformats.org/drawingml/2006/wordprocessingDrawing"
8	xmlns:w10="urn:schemas-microsoft-com:office:word"
9	xmlns:w="http://schemas.openxmlformats.org/wordprocessingml/2006/main"
10	xmlns:wne="http://schemas.microsoft.com/office/word/2006/wordml">
11 🗢	<w:body></w:body>
12 🗢	<w:p w:rsidp="00015E33" w:rsidr="00CD2E84" w:rsidrdefault="00015E33"></w:p>
13 🗢	<w:ppr></w:ppr>
14	<w:jc w:val="center"></w:jc>
15 🗢	<w:rpr></w:rpr>
16	<w:b></w:b>
17	
18	
19 🗢	<₩ï,>
20 🗢	<w:rpr></w:rpr>
21	<w b=""></w>
22	
23	<w:t>INLS 585: Management for Information Professionals</w:t>
24	
25	
26 🗢	<w:p w:rsidp="00015E33" w:rsidr="00015E33" w:rsidrdefault="00015E33"></w:p>
27 🗢	<w:r></w:r>
28 🗢	<w:rpr></w:rpr>
29	<w ii=""></w>
30	
31	<w:t>Texts:</w:t>
32	
33	
34 🗢	<w:p w:rsidp="00015E33" w:rsidr="00015E33" w:rsidrdefault="00015E33"></w:p>
35 🗢	<w:ppr></w:ppr>
36	<w:spacing w:after="0" w:line="240" w:linerule="auto"></w:spacing>
37	
38 🗢	<w;t></w;t>
39	<w:t xml:space="preserve">Robbins, S.P. and D.A. </w:t>
40	sharra 133

MIME types ("Content-type", "internet media type")

- Widely adopted and recognized by applications
- Based on two-level hierarchy (e.g. text/html, application/octet-stream, image/tiff)
- Major advantage is official registration of MIME types through a central authority

MIME types

Name	MIME Type / Internet Media Type	File Extension	More Details
3D Crossword Plugin	application/vnd.hzn-3d-crossword	.x3d	IANA: 3D Crossword Plugin
3GP	video/3gpp	.3gp	Wikipedia: 3GP
3GP2	video/3gpp2	.3g2	Wikipedia: 3G2
3GPP MSEQ File	application/vnd.mseq	.mseq	IANA: 3GPP MSEQ File
3M Post It Notes	application/vnd.3m.post-it-notes	.pwn	IANA: 3M Post It Notes
3rd Generation Partnership Project - Pic Large	application/vnd.3gpp.pic-bw-large	.plb	3GPP
3rd Generation Partnership Project - Pic Small	application/vnd.3gpp.pic-bw-small	.psb	3GPP
3rd Generation Partnership Project - Pic Var	application/vnd.3gpp.pic-bw-var	.pvb	3GPP
3rd Generation Partnership Project - Transaction Capabilities Application Part	application/vnd.3gpp2.tcap	.tcap	3GPP
7-Zip	application/x-7z-compressed	.7z	Wikipedia: 7-Zip

https://www.iana.org/assignments/media-types/media-types.xhtml

National Software Reference Library (NSRL)

The NSRL (<u>https://www.nsrl.nist.gov</u>) includes a library of hashes of files associated with a large number of software tools developed over the past few decades. See the product list at:

https://www.nsrl.nist.gov/RDS/rds_2.41/ProdList.txt

 Various third-party tools can be used to access the NSRL. There's a web interface available at:

https://www.hashsets.com/home/

(Navigate to Hash Set Engines > National Software Reference Library > SEARCH BY NAME / MD5). But it often generates invalid results, so the following instructions are based on running a command-line tool instead.

Using NSRL Hash Sets to Investigate System Files

- Find a directory from your computer that contains system files.
 - For Windows, a good place to look is in Computer > Local Disk (C:) > Program Files. For example, you could select Program Files > 7-Zip.
 - On a Mac, look in /Applications/ and select a specific folder
- Move the contents of the directory to a new folder called system-files on your host computer's desktop.
- Navigate to your shared folders [Desktop > Shared Folders and Media] in the BitCurator environment and copy the folder system-files to the desktop of the BitCurator environment
- Use md5deep to create a set of md5 hashes of the files in the system-files folder, then pipe the output into nsrllookup to generate lists of known and unknown hashes:
 - md5deep –r ~/Desktop/system-files | nsrllookup –s nsrllookup.com -K known-hashes.txt -U unknown-hashes.txt
- What is the above command doing?
- Look at the contents of the two files:
 - type known-hashes.txt
 - type unknown-hashes.txt

For Your Reference: Running NSRL Lookup in Windows

- Visit: <u>https://rjhansen.github.io/nsrllookup/</u> and download the Windows binary (64-bit).
- Open the .zip file and extract the executable to your desktop.
- Visit: <u>https://github.com/jessek/hashdeep/releases</u> and download <u>md5deep-4.4.zip</u>.
- Open the .zip file and extract md5deep64.exe to your desktop.
- Open a command prompt window (in the start box, type "cmd"). Navigate to your desktop (cd Desktop).
- Type: nsrllookup –help
- Same commands as in previous slide but use quotation marks around the file path in the command.

Exercise: Using PRONOM

The PRONOM technical registry contains information about a wide variety of file formats, including versioning information. You can find it at https://www.nationalarchives.gov.uk/PRONOM/Default.aspx. PRONOM has an online search feature that can be used to view the registry.

Click on "Search PRONOM" and navigate to the "File Format" tab. Clicking on the first search button (under "1. File Formats") will allow you to view all of the entries in the registry.

DROID incorporates information from PRONOM. It also uses file magic and file format extensions to provide a "best effort" at identifying file types. If you'd like to know more about DROID, you can find a quick demonstration video at: <u>https://vimeo.com/24718678</u>

Note: We'll see DROID output in the Siegfried exercise later.

Creating and Extracting Forensic Metadata

High-Level View of Metadata Generation and Reporting



See: Woods, Kam, Christopher Lee, and Sunitha Misra. "Automated Analysis and Visualization of Disk Images and File Systems for Preservation." In *Proceedings of Archiving 2013* (Springfield, VA: Society for Imaging Science and Technology, 2013), 239-244.

XML Schema for Digital Forensics XML

• 43 commits	្រ 1 branch	S 9 releases	🛱 1 contributor	<> Code	
branch: master	 dfxml_schema / + 			() Issues	8
Document an XML validati			· · · · · · · · · · · · · · · · · · ·	1 Pull requests	0
ajnelson authored on D	Dec 4, 2014		latest commit 4c8aab566e 🗟	de Pulse	
ref	Allow offline validation with local XSD	cache	2 years ago	to on the	
LICENSE.txt	Add public domain license text		2 years ago	Graphs	
README.md	Document an XML validation step		6 months ago	HTTPS clone URL	
dfxml.xsd	Document an XML validation step		6 months ago	https://github.com/c	
README.md				You can clone with HTTPS of Subversion.	96
				Clone in Deskt	ор
This is the schema	a repository for Digital Forensics X	ML, version 1.1.1.		C Download ZI	P
If you intend to use download two acc downloading the re To report issues, o	e the dfxml.xsd file as a DFXML do ompanying .xsd files under the "re epository as a Git clone, or by dow questions, or feature requests, plea	ocument validator, note th f" directory. The easiest w vnloading the zip archive f ase either:	at you will also need to vay to do this is by from the Github page.		
File a Github	issue, seeing first if it is already file	ed, here.			

 Email the dfxml@nist.gov mailing list. If you wish to join the mailing list, send an email to dfxmlsubscribe@nist.gov (no subject or message body is necessary), and a moderator will grant access.

https://github.com/dfxml-working-group/dfxml_schema
Operationalizing Original Order - Filesystem Metadata Output from fiwalk*

```
<fileobject>
      <filename>Documents and Settings/All Users/Documents/
                 My Pictures/Sample Pictures/Blue hills.jpg
      </filename>
      <filesize>28521</filesize>
      <alloc>1</alloc>
      <used>1</used>
      <inode>6245</inode>
      <uid>0</uid>
      <gid>0</gid>
      <mtime>1208174400</mtime>
      <ctime>1257729636</ctime>
      <atime>1257729636</atime>
      <crtime>1257729636</crtime>
      <seq>2</seq>
      libmagic>JPEG image data, JFIF standard 1.02</libmagic>
      <byte runs>
       <run file offset='0' fs offset='0' img offset='363200512'
         len='0'/>
      </byte runs>
      <hashdigest type='MD5'>
          6fb2a38dc107eacb41cf1656e899cf70
      </hashdigest>
      <hashdigest type='SHA1'>
          4eee44b18576e84de7b163142b537d2fe6231845
      </hashdigest>
</fileobject>
```

*Developed by Simson Garfinkel

PREMIS (Preservation) Metadata Generated from Running BitCurator Tools – Recorded as PREMIS Events



Provenance – DFXML Output from fiwalk

```
BitCurator-0.2.0 [Running]
                                                                                                                                                                                                            🖂 📼 ҵ 🜒 8:08 PM 👤 BitCurator 🔱
Mozilla Firefox
                                                                                                                                                                                                    🗍 file:///home/b...mpleimage.xml 📑
    0
                              Intersection of the section of th
                                                                                                                                                                                                                  V C
                                                                                                                                                                                                                              8 - Google
                                                                                                                                                                                                                                                                                          Q 🏫
                    This XML file does not appear to have any style information associated with it. The document tree is shown below.
                     <dfxml version="1.0">
                       -<metadata>
                               <dc:type>Disk Image</dc:type>
                           </metadata>
                       -<creator version="1.0">
                               <program>fiwalk</program>
     - - -
                                <version>4.0.2</version>
                           -<build environment>
                                    <compiler>GCC 4.6</compiler>
                                    library name="afflib" version="3.7.1"/>
                                    library name="libewf" version="20130303"/>
                                </build environment>
    P
                           -<execution environment>
                                -<command line>
                                        fiwalk -f -X /home/bcadmin/Desktop/SampleData/sampleimage.xml /home/bcadmin/Desktop/SampleData
                                        /sampleimage.E01
                                    </command line>
                                    <start time>2013-03-12T00:08:28Z</start time>
                                </execution environment>
                           </creator>
   6
                        -<source>
                               <image filename>/home/bcadmin/Desktop/SampleData/sampleimage.E01</image filename>
                           </source>
     •
                           <!-- fs start: 0 -->
                        -<volume offset="0">
                               <partition offset>0</partition offset>
                                <block size>2048</block size>
                                <ftype>2048</ftype>
                                <ftype str>iso9660</ftype str>
                                <body><block count>36839</block count>
                                                                                                                                                                                                                                     일 💿 🖉 🗗 🛄 🚺 🔇 🖳 Left 🕷
```

147

Identifying "Features" of Interest in Disk Images or Directories

Bulk Extractor

0 0	BitCurator-0.3.0 [Running]	
k Extractor Viewer		E I (1) 2:00 AM BitCurator
File Edit View Tools He	8 Run bulk_extractor	
X Highlight: Reports Feature Filter Feature File None	Required Parameters Scan: Image File Cop/SampleData/sampleimage.E01 Cotput Feature Directory Ceneral Options Use Banner File Use Alert List File Use Stop List File Use Find Regex Text File Comparison	Scanners bulk wordlist accts aes base16 base64 elf email exif gps
	Use Find Regex Text Tuning Parameters Use Context Window Size 16 Use Page Size	 ✓ gzip ✓ hiber ✓ json ✓ kml ✓ net ✓ ad5
Referenced Featu Referenced Featu	Use Margin Size1048576Use Min Word Size6Use Max Word Size14Use Block Size512Use Number of Threads1	 ✓ par ✓ vcard ✓ windirs ✓ winpe ✓ winprefetch ✓ zip
	Scanner Controls Use Plugin Directory Use Scan Option Name Restore Defaults Start bulk_extractor Cancel	

Bulk Extractor* – Identifying Potentially Sensitive Information

× 😐 🖦 🕇	Pequired Parameters	Scappers
	Scan: Image File Raw Device Directory of Files	
Reports Fe		□ wordlist
Fe		S accts
	General Options	🧭 aes
	🗌 Use Banner File	🗹 base16
	Use Alert List File	👿 base64
1		🧭 elf
		🞯 email 📲
	Use Find Regex Text File	🗹 exif
	Use Find Regex Text	Market Find
	Use Random Sampling	gps
	Tuning Parameters	i gzip
	Use Context Window Size	
		Son Son
Re		M net
		Pdf
	U USE BLOCK SIZE	👿 rar
	Use Number of Threads	🞯 vcard
	Use Maximum Recursion Depth 7	🞯 windirs
	Use Wait Time 60	🞯 winpe
	Parallelizing	🜌 winprefetc
	Use start processing at offset	🜌 zip
e.//foron	sicewiki xyz/wiki/indox pho2title	-Bulk oxtractor

*Developed by Simson Garfinkel

O O BitCurator-0.3.0 [Running]	
Bulk Extractor Viewer	🔜 💌 🗊 🛊 🕪) 2:02 AM L BitCurator 🔱
File Edit View Tools Help	
🤨 🗴 🔒 👒 🖷 🗰 📕	
Highlight: 🛛 🗹 Match case	
Reports Feature Filter 🗌 Match case Navigation	
× bulk_extractor Scan	•
F Image File sampleimage.E01 Feature Directory bulk-extractor-output	
Progress Done	
Dulk_extractor scan completed. See Status, below, for details.	
'bulk_extractor' Report is Ready	
bulk_extractor has completed.	
//home/bcadmin/Desk //home/bcadmin/Desk	and is ready for viewing.
Status	
R Elapsed time: 0.4985 sec. R Overall performance: 2.958 MBvtes/sec.	
Total email features found: 0	
Close	
	Taut O Llaw / A A S
	🐸 🕑 🖉 🖃 🛄 🚺 💆 Left 📽 🏼 🏸

151

Histogram of Email Addresses (Specific Instances in Context on Right)

000		BitCu	urator-0.2.0 [Running]	2
Bulk Ex	tractor Viewer		🚨 🖂 📼 🃭 🕕 8:44 PM 👤 BitCurator	ψ
۲	x 🔒 🎕 🖷 🛛	E &		
	🗶 Highlight:	✓ Mat	ch case	
	Reports	Feature Filter 🔲 Match case	Navigation	
	▼ beoutput		🧝 🕱 sampleimage.E01, 42273785, privacy@Motorola.com	
	domain.txt domain histogram	Histogram File email histog.	Image File sampleimage.E01	
	email.txt	n=12 privacy@motorola.cor	Feature File email.txt	
>_	email_histogram.t>	n=3 0mj5nj@0itgx.ib.dj	Feature privacy@Motorola.com	
	ether.txt	n=3 john@humaniz.com	readic physicy@motorota.com	
	ether_histogram.ti	n=3 newton@planetb.fr	Image	1.
	packets pcap	n=3 sales@integrationnew n=1 5kda c@kgahw.sl	42272000 by MotorolaMotorola uses Secure Sockets Laver (SSL) encrypti	
	rfc822.txt	n=1 dqf@40mt.ro	42272064 on technology, the highest level of security on the Internet. Th	
	tcp.txt	n=1 fodfv@nwa4.ck	422/2128 e SSL protocol provides server authentication, data integrity, a 42272192 nd privacy on the Web. This security measure helps ensure that n	
H	tcp_histogram.txt	n=1 jqnmq@17.pn	42272256 o impostors, eavesdroppers, or vandals get your personal informa	
	url.txt	n=1 kjph@sj.gr	42272320 tion. SSL not only encrypts your personal and financial informat	
	urt_nistogram.txt	n=1 pdcnfb@tft.ao	42272448 ifies the identity of the server and that the original message a	
	windirs.txt	n=1 qyf@j65.de	42272512 rrives safely at its destination. However, no data transmission	\cap
0000	winpe.txt	n=1 tw+4vsa@x1.ms +	42272576 Over the internet can be guaranteed to be 100% secure. As a res 42272640 ult. while we strive to protect your personal information. Motor	
202			42272704 ola cannot ensure or warrant the security of any information you	
		Referenced Feature File e	42272768 transmit to us or from our Web site, and therefore you use our	
10		Referenced Feature pri	42272896 our best effort to ensure its security on our systems000200	
		34807246 privacy@Motor	42272960 0007AE000038B6.7A8,As a global company Motorola has internationa	Н
		34808676 privacy@Motor	422/3024 i sites and users all over the world, when you give Motorola per 42273088 sonal information, that information may be sent electronically t	
24		42271602 privacy@Motor	42273152 o servers outside of the country where you originally entered th	
3.5		42274743 privacy@Motor	42273216 e information. In addition, that information may be used, stored	
		42347307 privacy@Motor	42273344 ormation. Whenever Motorola handles personal information, regard	
		42350448 privacy@Motor	42273408 less of where this occurs, it takes steps to ensure that your in	0
		74735841 privacy@Motor	422/34/2 formation is treated securely and in accordance with the relevant 42273536 t Terms of Use and this Privacy Policy. How can I correct or ch	
		74738019 privacy@Motor 74738989 privacy@Motor	42273600 ange my personal information? .If you would like to review, corr	
		Princip Grinder	42273664 ect or change any personal information you have provided, or rem	
			42273792 @Motorola.com. If you have established a "user profile" on a Mot	
1			42273856 orola website, you may change the information you provided at an	-
8985	4 (iii) +	۰ <u>ا</u>	🖲 Text 🔿 Hex 🛛 🖶 🎽 🗭	
5			일 🕟 🖉 🖃 🛄 🚺 Left 🕊	10

152

BitCurator Reporting Tool

Run All	Fiwalk XML	Annotated Features	Reports		
Run fiwo	alk, annotate the	e bulk_extractor output, o	and generate Off	fice / PDF reports.	
If you ha	ven't run bulk_ex nd run it first.	tractor yet, use the button	to the right to	Launch BEVie	we
Image Fi	le				
/media	/sf_bc_share/	buf_exerciseplanning/	8-jpeg-search/8	3-jpeg-search.dd	
Bulk Ext	ractor Feature	Directory			
/media	/sf_bc_share/	buf_exerciseplanning/	8-jpeg-search/j	peg_search_beou	Jt
Output I	Directory (fiwal	k output, annotated feat	ures, and report	s will appear in her	e)
/media	/sf_bc_share/l	ouf_exerciseplanning/	8-jpeg-search/j	peg_search_repo	rts
Config F	ile (Optional)				
/Path/	To/File				
Comman	d Line Output				
>> Ima search, >> Ann /media search, >> Succ o /me search, >> Gen /media search, >>> Gen	ge file selecte /8-jpeg-search otate: bulk_ex /sf_bc_share// /jpeg_search_l cess. fiwalk credit dia/sf_bc_share/ /jpeg_search_l nerating PRE nerating bulk_ ating annotate	d: /media/sr_bc_share .dd .tractor feature directo buf_exerciseplanning/ beout eated the following file re/buf_exerciseplannir reports/fiwalk-output. S event for fiwalk in: buf_exerciseplanning/ reports/reports _extractor PREMIS eve ed features	/bur_exercisep ory selected: 8-jpeg- 9: 9: 9: 9: 9: 8-jpeg- nt	Nanning/8-jpeg-	
>> Crea					
>> Crea			Close Ca	ancel Run	

🔜 🔃 En 📧 🗤) 3:22 PM 🔱

BitCurat<u></u>

and Help

Various Specialized BitCurator Reports



Detail: Specialized BitCurator Reports

File	Content
bc_format_bargraph.pdf	histogram of file formats found on the volume
bulk_extractor_report.pdf	high-level overview of feature locations on disk
fiwalk_deleted_files.pdf	shows paths to any deleted materials found in a given partition
fiwalk-output.xml.xlsx	Excel converted DFXML output (file system metadata)
fiwalk_report.pdf	high-level overview of file system characteristics
format_table.pdf	long-form file format names for formats shown in bar graph
premis.xml	PREMIS preservation metadata

Nautilus Scripts

- Scripts that can be run using Nautilus (GNOME file manager)
- Most provide more convenient access (right click and menu selection) to functions performed by applications that could also be run directly

Right-click on file or directory and create MD5



Quick access to a hex view



Other functionality to meet user needs

Function	Tool(s)
Identify duplicate files	FSLint
Characterize files	FITS, FIDO
Scan for viruses	ClamTK
Examine, copy and extract information from old Mac disks	HFS Utilities (including HFS Explorer)
Capture AV file metadata	MediaInfo, FFProbe
Extract text from older binary (.doc) Word files	antiword
Read contents of Mircosoft Outlook PST files	readpst
Examine embedded header information in images	pyExifToolGUI
Generate images of problematic disks or particular disk types (I addition to Guymager	dd, dcfldd, ddrescue, cdrdao (for audio CDs)
Extract and analyze data from Windows Registry files	regripper
Identify files that are partially similar but not identical	sdhash, ssdeep
Package files for storage and/or transfer	Baglt (Java) library, Bagger
File preview (left-click on file then hit space bar)	gnome-sushi

Other functionality to meet user needs

Function	Tool(s)
Play and examine metadata from AV media files	VLC media player
Damaged/lost partition recovery	TestDisk
Damaged/lost file recovery	PhotoRec
Identify the filesystem on a disk	disktype
Index and search for keywords in documents	recoll
Find blacklist data by using hashes calculated from hash blocks	hashdb
Generate hashes of files and blocks	GTK Hash, md5deep, md5sum
Compare hashes of files to hashes in the National Software Reference Library (NSRL) of known system files	nsrllookup
View and edit bytestreams (hex editor)	Bless Hex Editor, GHex

Command Line Operations

- Opens up many more possibilities, such as:
 - stringing tools together
 - performing batch operations
 - changing parameters from their default values
 - using tools that are only available through the command line (no GUI)

Some Considerations

- Role of pipes feed output from one process into another process
- Switches settings that can be applied to a command (e.g. -a, -r)
- Argument a specific piece of data that is processed by a program (e.g. a variable or fixed value)
- Regular expressions used to find patterns (more on this later)
- Text created in Windows and Unix, even though they're both ASCII, will encode new lines differently, so you may need to translate using a tool such as dos2unix or unix2dos.

Some Important Commands and Tasks

- **mkdir** make a directory
- cd change the directory that you're in ["cd .." goes to the parent of the current directory]
- **Is** list contents of a directory
- md5sum generate cryptographic hashes
- cat output content of a text file (can be concatenation of contents of two files)
- **file** determine file types based on magic numbers
- strings matches patterns in the text (ASCII) parts of a file (file can be binary)
- diff compare two files
- hexdump very basic (non-GUI) hex viewer

General Unix/Linux CLI Tips

- man manual page that explains how to run a command or some other technical information (e.g. ascii page)
- **control-z** quit currently running program
- clear clear the screen (hide text from previous commands)
- Up arrow cycles through previous commands, so you can rerun (or adapt) them
- Tab hit this key after you've started typing a string that the operating system can fill in for you (e.g. a long file name). Hitting tab multiple times will cycle through available options.

64

- The saa-dfa-sample-data.zip file you downloaded earlier contains a folder of sample files named file_ident_ex
- If you haven't done this already, add shared folder to BitCurator VM, pointing to the desktop of the host

 General System Display Storage Audio Network Serial Ports USB Shared Folders 	Shared Folders Folders List Name Path Machine Folder Path: C:\Users\callee\Desktop Folder Name: Desktop Read-only Auto-mount OK Cancel
--	--

Move the file_ident_ex folder to the BitCurator VM desktop

Open a terminal in the BitCurator environment (using the Terminal icon in the dock)



Command	Reason/Explanation
pwd	Show the directory you're currently in
ls	List the contents of the current directory
cd Desktop	Change the current directory to Desktop
ls	List the contents of the current directory
unzip files.zip	Decompress and unpack content of files.zip
ls	List the contents of the current directory
cd files	Change the current directory to files
ls	List the contents of the current directory
md5sum [file name of first file] > firsthash	Create a hash of a file and output it to a text file
less firsthash	Display the content of the output to the screen
Control-z	Stop the "less" program
md5sum [file name of second file] > secondhash	Create a hash of a second file and output it to a text file
cat firsthash secondhash > bothhashes	Combine the context of the two output files
more bothhashes	Display the content of the output to the screen
most bothhashes	Display the content of the output to the screen (follow instructions for adding it), then run this command again

Gives you the right administrative permissions



Uses Advanced Packaging Tool to get the program₁₆₇

Command	Reason/Explanation
rm firsthash	Delete (remove) firsthash file
rm secondhash	Delete (remove) secondhash
ls	List the contents of the current directory
hexdump [file name] -C less	Show hex dump of a given file [-C switch shows the standard view of hex on left and ASCII on right]
Use up and down arrows	Navigate within the hex view of the file's content
:q	Quit the "less" program

Exercise: Siegfried and Brunnhilde

- Over the next few slides, we will run Siegfried and Brunnhilde in over the same set of files in several different ways
- Goals
 - Generate characterization and related technical metadata
 - Illustrate how the data can be configured for different uses
 - Identify decision points when data are unclear
- Source files to analyze: file_ident_ex directory in the the zip file you downloaded and extracted earlier

Installing Siegfried and Brunnhilde

Start up the BitCurator VM (if it's not already running) Open a Terminal Window

Image: Second min@ubuntu: ~ bcadmin@ubuntu: -\$ Image: Second min@ubuntu: -\$ Image: Sec	Documentation and Help Network Servers
	BitCurat ∉ r

Installing Siegfried and Brunnhilde

Enter the following commands in Terminal:

wget -qO -

https://bintray.com/user/downloadSubjectPublicKey?username=bintr ay | sudo apt-key add –

- echo "deb <u>http://dl.bintray.com/siegfried/debian wheezy main" | sudo</u> tee -a /etc/apt/sources.list
- sudo apt-get update && sudo apt-get install Siegfried
- sudo pip install brunnhilde

 Note: there is a text file in the sample files that includes these commands if you want to copy/paste them

Running Siegfried and Brunnhilde I

- Ensure you have extracted the saa-dfa-sample-data.zip file
- Create a shared folder (we'll step you through this)
- Start up the BitCurator VM
- Create "sieg-out" and "brunn-out" folders on the desktop
- Drag "file_ident_ex" folder from the shared saa-dfa-sample-data directory to the Desktop



Running Siegfried and Brunnhilde II

- Open a Terminal window
- At the prompt, enter the following command:
 - sf ~/Desktop/file_ident_ex/ > ~/Desktop/sieg_out/sieg_out.yaml
- Open the sieg_out directory and look around
 - What did the command do?
 - What does the file tell you?
 - How would you characterize the data presented in the file?
 - Does anything strike you as odd? Particularly useful?

Running Siegfried and Brunnhilde III

- In the same Terminal window, enter the following commands:
 - sf -droid ~/Desktop/file_ident_ex/ > ~/Desktop/sieg_out/sieg_out-droid.csv
 - sf -json ~/Desktop/file_ident_ex/ > ~/Desktop/sieg_out/sieg_out-json.json
- Open the sieg_out directory and look around
 - What did the commands do?
 - How do these files differ from the one created on the previous slide?
 - Between the three output files, which do you think is most useful presentation of the data? (Hint: there may be more than one answer)
 - Do any of these files strike you as particularly useful? Particularly worthless?

Running Siegfried and Brunnhilde IV

- In the same Terminal window, enter the following command:
 - brunnhilde.py -w ~/Desktop/file_ident_ex/ ~/Desktop/brunn_out/DAS-FileIdent
- Open the brunn_out directory and look around
 - Did Brunnhilde perform any tasks over and above Siegfried?
 - How do the Brunnhilde output files differ from those generated by Siegfried?
 - Inspect csv_reports. How would you characterize what you see here?
 - Are the files here that Brunnhilde and Siegfried found problematic?
 What conclusions might you draw from them?
 - Is there information that Brunnhilde highlighted that you missed in Siegfried's output?

😸 🗇 💷 DAS-FileIdent								
< > 🔂 Home	Desktop	brunn_out	DAS-FileIdent	csv_reports			Q	↓ [Ⅲ Ⅲ
⊘ Recent	Nam	e				Size	Туре	Modified
☆ Home		csv_reports				8 items	Folder	22:03
Desktop		logs			1 item	Folder	21:59	
Documents	())	DAS-FileIdent.html			12.1 kB	Text	21:59	
🕹 Downloads		siegfried.csv			4.1 kB	Text	21:59	
d Music	H	siegfried.sqlite			7.2 kB	Unknown	21:59	
D Pictures		tree.txt				789 bytes	Text	21:59
Videos								

Regular Expressions

- What is a regular expression, or regex?
 - Simply a pattern for matching bits of text.
- What are regex's useful for?
 - Three things:
 - 1. Matching: Does this text contain a pattern?
 - 2. Replacement: Replace some part of the text with other text
 - *3. Extraction:* Yanking out a bit of the text to use somewhere else.

Regular Expressions

- Regular expressions may contain ordinary letters, numbers, and *a few special symbols* that allow you to match a wide range of patterns with a small amount of syntax.
- A regex needs to be interpreted by a program (such as a Perl or Python script) or by an application (such as the Forensic Toolkit)
- Regular expression syntax may look different in different languages and programs.

Regular Expressions – Special Characters

Character	Meaning
•	Match anything
+	Match one or more occurrences
*	Match zero or more occurrences
٨	Match only at the start of the text
\$	Match only at the end of text
\w	Match an alphanumeric word
\d	Match a number
\s	Match any whitespace
\S	Match anything except whitespace

Regular Expressions - Examples

Expression	Explanation
r+	Match the letter 'r' one or more times
t*	Match the letter 't' zero or more times
\d	Match any single digit (shorthand for the expression [0-9]
\dd	Match any pair of digits (alternatively, [0-9][0-9])

Regular Expressions – More Complex Examples

Expression	Explanation
\d+\.\d+	Any two-digit decimal number (e.g. 5.0, 15.2345)
\d+(\.\d+)*	Any integer or decimal number (e.g. 5, 5.0, 15.2345)
(\d\d\d)-(\d\d)-(\d\d)	Match any date strings of the form YYYY-MM-DD

What's "wrong" with the second and third regex patterns shown here? What else might they match?
Regular Expressions in FTK

- Tools for building regular expressions one part at a time
- Various default regular expressions that you can use or adapt

For Fun on Your Own – Regex Golf!

R	egex Golf			
С	lassic Teukon Holiday			
War	mup – Type a regex in the box.			
				0
N	latch all of these	a	nd none of these	
	afoot	×	Atlas	
1	catfoot	×	Aymoro	
1	dogfoot	×	Iberic	
1	fanfoot	×	Mahran	
1	foody	×	Ormazd	
-	foolery	×	Silipan	
-	foolish	×	altared	
1	fooster	×	chandoo	
1	footage	×	crenel	
1	foothot	×	crooked	
1	footle	×	fardo	

Try it at https://alf.nu/RegexGolf

Extracting Data From Specific Types of Files

Exchangeable Image File Format (EXIF)

Possible tags:

https://exiftool.org/TagNames/EXIF.html

Camera manufacturer	Canon
Camera model	Canon EOS 1200D
Author	Praveen. P
Exposure time	1/60 sec (0.0166666666666667)
F-number	f/11
ISO speed rating	200
Date and time of data generation	22:29, 22 November 2018
Lens focal length	41 mm
Show extended details	

Example of EXIF Metadata from a JPEG File (Generated Using exiftool*)

ExifTool
ExifTool Version Number : 9.38
System
File Name : IMG 20130823 151811.jpg
Directory : C:/Users/callee/Documents/images/digital-forensics-lab
File Size : 1785 kB
File Modification Date/Time : 2013:08:23 16:36:44-04:00
File Access Date/Time : 2013:10:14 17:13:02-04:00
File Creation Date/Time : 2013:08:23 16:36:44-04:00
File Permissions : rw-rw-
File
File Type : JPEG
MIME Type : image/ipeg
Exif Byte Order : Big-endian (Motorola, MM)
Image Width : 2592
Image Height : 1944
Encoding Process : Baseline DCT. Huffman coding
Bits Per Sample : 8
Color Components : 3
Y Cb Cr Sub Sampling : YCbCr4:2:0 (2 2)
GPS
GPS Img Direction : 83
GPS Img Direction Ref : Magnetic North
GPS Latitude Ref : North
GPS Latitude : 35 deg 55' 2.24"
GPS Longitude Ref : West
GPS Longitude : 79 deg 2' 57.55"
GPS Altitude Ref : Above Sea Level
GPS Altitude : 0 m
GPS Time Stamp : 19:18:06
GPS Processing Method : NETWORK
GPS Date Stamp : 2013:08:23
IFD0
Orientation : Unknown (0)
Camera Model Name : Galaxy Nexus
Modify Date : 2013:08:23 15:18:11
Y Cb Cr Positioning : Centered
Y Resolution : 72
Resolution Unit : inches
X Resolution : 72
Make : Samsung
ExitIFD
Create Date : 2013:08:23 15:18:11
Date/Time Original : 2013:08:23 15:18:11
Exit Version : 0220
Hash Energy : 0
Image Unique ID : UAEL01

Scene Type : Directly photographed Exposure Index : undef **Components Configuration** : Y, Cb, Cr, -F Number : 2.8 Compressed Bits Per Pixel : 0 Sensing Method : One-chip color area Exposure Program : Aperture-priority AE Aperture Value : 2.6 **Brightness Value** : 0 Subject Distance Range : Unknown Shutter Speed Value : 1/15 Subject Distance :0 m Saturation : Normal Color Space : sRGB Contrast : Normal Metering Mode : Multi-spot Flashpix Version ÷ **Exposure Compensation** :0 : 1944 Exif Image Height Max Aperture Value : 2.6 Sharpness : Normal : 2592 Exif Image Width Focal Length : 3.4 mm **Digital Zoom Ratio** :1 Light Source : Fluorescent Scene Capture Type : Standard Flash : Off, Did not fire Custom Rendered : Custom White Balance : Auto Exposure Mode : Auto ---- IFD1 ----Compression : JPEG (old-style) Image Width : 160 : 120 Image Height Thumbnail Offset : 1239 Thumbnail Length :7164 ---- Composite ----: 2.8 Aperture **GPS** Altitude : 0 m Above Sea Level **GPS** Date/Time : 2013:08:23 19:18:06Z **GPS** Latitude : 35 deg 55' 2.24" N **GPS** Longitude : 79 deg 2' 57.55" W **GPS** Position : 35 deg 55' 2.24" N, 79 deg 2' 57.55" W Image Size : 2592x1944 Shutter Speed · 1/17 : (Binary data 7164 bytes, use -b option to extract) Thumbnail Image : 3.4 mm Focal Length Light Value : 6.7

*https://exiftool.org/TagNames/EXIF.html (Also available through the BitCurator environment)

Exiftool Exercise in BitCurator

- Start up the BitCurator VM
- Download one or more pictures to your desktop that you'd like to examine
- Options for viewing EXIF:
 - 1. PyEXIFToolGUI:
 - Navigate to Desktop > Forensics Tools > PyEXIFToolGUI
 - Open the tool and select File > Load Images
 - Let's also add some GPS coordinates: Select Edit Data, enter the values, then select Save to Selected Image(s) [Make sure that the image is selected]
 - 2. File info menu:
 - Navigate to the image file
 - Right click on it and select Scripts > File Info > Meta Information [Pick EXIF Data]
 - 3. exiftool at the command line:
 - Open a command prompt window
 - Navigate to where you stored the image (e.g. cd Desktop)
 - Type: exiftool [Filename]
 - Note: You can scroll up and down by using Shift + Page Up/Page Down, or you can invoke the command as *exiftool [Filename]* | *less* (type "q" to quit)

Optional Exiftool Exercise (Windows or Mac)

- Download one or more pictures to your desktop that you'd like to examine
- Download and unzip the latest Windows executable (or Mac package) from: <u>https://exiftool.org/</u>
- Save exiftool(-k).exe to your desktop
- Change file name to: exiftool(-k -a -u -g1 -w txt).exe [NOTE: This is changing the parameters for running the software – same as if you were to add these switches at the command line. This trick might not work on a Mac, but you can always issue the commands directly.]
- -k = pause the program before terminating
- -a = allow extraction of duplicate tags
- -u = extract unknown tags
- -g1 = organize output by tag group
- -w = write output text file
- Drag and drop pictures onto the exiftool icon and examine the results
- Change file name to: exiftool(-X -k -a -u -g1 -w xml).exe
- Drag and drop pictures onto the exiftool icon and examine the results
- For more about exiftool, see: <u>https://github.com/exiftool/exiftool</u>

Stripping Metadata From Images

Social Media site/system	Summary	Disp	ectly?	Displays 4Cs?	er	Save A	s ed?	Den	ownloa nbedde	ad ed?
500px - www.500px.com Tested in late 2015	Some embedded metadata fields are shown, all correctly, but not the rights-relevant 4C fields. Metadata preserved in SaveAs file. Compared to 2013: SaveAs preserves metadata now = improvement	Exif	IPTC	IPTC	Exif	IPTC IIM	IPTC XMP	Exif	IPTC IM	IPTC XMP
BEHANCE - www.behance.net Tested in late 2015	All rights-relevant fields and more are shown, all corectly. Embedded metadata is preserved in the SaveAs and the downloaded image file. Compared to 2013: not tested then	Exif	IPTC	IPTC	Exif	IPTC	IPTC XMP	Exif	IPTC	IPTC XMP
Dropbox - www.dropbox.com Tested in late 2015	No embedded metadata shown. Embedded metadata only preserved in the downloaded image file but not in the SaveAs. Compared to 2013: also SaveAs files preserved metadata then = decline	Exif	IPTC	IPTC	Exif	IPTC IIM	IPTC XMP	Exif	IPTC	IPTC XMP
EyeEm - <u>www.eyeem.com</u> Tested in late 2015	No embedded metadata shown. SaveAs file was downscaled and all metadata was stripped off. Compared to 2013: not tested then	Exif	IPTC	IPTC	Exif	IPTC	IPTC XMP	Exif	IPTC	IPTC XMP
Facebook - www.facebook.com Tested in late 2015	No embedded metadata shown. SaveAs file preserved Copyright Notice and Creator in IIM, anything else is stripped off. Surprise: 2 IIM fields contain data generated by Facebook. Compared to 2013: at least 2 fields in IIM survive now = slight improvement	Exif	IPTC	IPTC	Exif	IPTC IIM	IPTC XMP	IPTC	IPTC IIM	IPTC XMP
Flickr FREE account- www.flickr.com Tested in late 2015	Some embedded metadata fields are shown, all correctly, but not all rights-relevant 4Cs. Embedded metadata is stripped off SaveAs files but preserved in downloaded files. Compared to 2013: plus = any downloded file preserves metadata now; minus = even high resolution SaveAs file does not preserve it now.	Exif	IPTC	IPTC	Exif	IPTC	IPTC XMP	Exif	IPTC IIM	IPTC XMP
Google Photo - photos.google.com Tested in late 2015	Some embedded metadata fields are shown, all correctly, but not all rights-relevant 4Cs. SaveAs works only for downscaled files - only Exif metadata is preserved. Downloaded files preserved all metadata. Compared to 2013/Google+ photos: SaveAs file gets IIIM and XMP metadata stripped off now = decline	Exif	IPTC	IPTC	Exif	IPTC IIM	IPTC XMP	Exif	IPTC IIM	IPTC XMP
Img.ly - <u>www.img.ly</u> Tested in late 2015	No embedded metadata shown. Embedded metadata is preserved in the high resolution/original size SaveAs image file but stripped off in a downscaled file. Compared to 2013: the loss of metadata in downscaled images was not tested in 2013.	Exif	IPTC	IPTC	Exif			Exif	IPTC	IPTC XMP
Instagram - <u>instagram.com</u> Tested in late 2015	Tested using the Instagram iOS app v 6.4.1: No embedded metadata fields are shown. No retrieval of image files possible. Compared to 2013: then SaveAs was possible - with stripped off metadata.	Exif	IPTC	IPTC	Exif	IPTC IIM	IPTC XMP	Exif	IPTC IIM	IPTC XMP
Joomeo - www.joomeo.com Tested in late 2015	Some embedded metadata fields are shown, all correctly, but not the rights-relevant 4Cs. Embedded metadata preserved in the downloaded image files. Compared to 2013: more embedded metadata were shown then, including $4Cs = slight decline$	Exif	IPTC	IPTC	Exif	IPTC IIM	IPTC XMP	Exif	IPTC IIM	IPTC XMP
LINKED IN 2015 - www.linkedin.com Tested in late 2015	No embedded metadata shown. Only embedded Exif fields are preserved in SaveAs files. Compared to 2013: not tested then.	Exif	IPTC	IPTC	Exif	IPTC IIM	IPTC XMP	Exif	IPTC IM	IPTC XMP
Pictify - <u>www.pictify.com</u> Tested in late 2015	No embedded metadata shown. No retrieval of image files possible. Compared to 2013: then SaveAs was possible - with stripped off metadata.	Exif	IPTC	IPTC	Exif	IPTC IIM	IPTC XMP	Exif	IPTC IM	IPTC XMP
Pinterest - www.pinterest.com	No embedded metadata shown. Embedded metadata preserved in high resolution/original size images, but IIM and XMP metadata is									

https://www.embeddedmetadata.org/social-media-test-results.php

Office Documents

- Are the "new" office formats (ODF and OOXML) better or worse for forensics?
- What kinds of information can you get out of them?
- What sorts of approaches might you take to view and/or extract the information?

Office Documents – PPTX File Example

 Your zip file contains a document named "The NDSA Levels of Digital Preservation.pptx"

(also at

http://library.harvard.edu/sites/default/files/The%20NDSA%20Levels%20of%20Digital%20Prese rvation_3.pptx

- Change the file extension to .zip
- Open it with 7-Zip or WinZip
- Extract all the files
- Examine the contents of the resulting directory
 - □ Can you find a thumbnail of the first slide?
 - □ Where are the slides stored?
 - □ Where are embedded images stored?
 - □ Can you determine who created the file?

Jonathan Larson Fast Save Example

👙 File Edit View Insert Format Font Tools Window 🕧 🕲	
***1/16/96	
	FIND THE ONE SONG BEFORE YOU ENTER THE LIGHT
FIND THE ONE SONG BEFORE THE VIRUS TAKES HOLD THE GLORY LIKE A SUNSET ONE SONG TO REDEEM THIS EMPTY LIFE	THE GLORY LIKE A SUNSET ONE SONG TO REDEEM THIS EMPTY LIFE
THE FLIES AND THEN- NO NEED TO ENDURE ANYHORE THE DIES (A knock on the door) THE DOOR	TIME FLIES AND THEN- NO NEED TO ENDURE ANYMORE TIME DIES (A knock on the door)
07. LIGHT HY CARDLE ROGER WHAT'D YOU FORGET?	THE DOOR 08. LIGHT MY CANDLE
(RE opens the door. HIHI stands, with a candle.) Page 13 Normal+	ROGER WHAT'D YOU FORGET?
	(HE opens the door. MIMI stands, with a candle.)

00028b60	09	09	09	Za	Za	Za	31	2f	31	36	2f	39	36	4f	55	52	1***1/16/960UR1
00028b70	20	57	45	44	44	49	4e	47	4f	4e	20	54	48	45	20	53	I WEDDINGON THE SI
00028b80	4f	46	41	53	4f	46	41	54	48	45	20	56	49	52	55	53	IOFASOFATHE VIRUS!
00028690	20	54	41	4b	45	53	20	48	4f	4c	44	4d	45	45	54	20	I TAKES HOLDMEET I
00028ba0	59	4f	55	20	41	54	20	54	48	45	20	53	48	4f	57	49	IYOU AT THE SHOWI!
00028660	27	4c	4c	20	54	52	59	20	41	4e	44	20	43	4f	4e	56	I'LL TRY AND CONVI
00028bc0	49	4e	43	45	20	52	4f	47	45	52	20	54	4f	20	47	4f	IINCE ROGER TO GOI
00028bd0	43	4c	4f	53	45	20	4f	4e	43	41	4e	20	49	20	48	45	ICLOSE ONCAN I HE!
00028be0	4c	50	4d	69	73	73	20	50	6f	72	74	65	72	27	73	46	ILPMiss Porter'sFI
00028bf0	4f	52	47	45	54	20	49	54	50	41	55	4c	20	Za	20	Za	IORGET ITPAUL****

https://www.nypl.org/blog/2011/04/22/no-day-today-look-jonathan-larsons-word-files

Hidden Data Exercise – Using a Hex Editor

Your zip file contains a document named ar99.doc (also at: http://web.archive.org/web/20000816164723/http://microsoft.com/msft/ar99/downloads/ar99.doc)

- Open the file in HxD (or upload to <u>https://hexed.it</u>)
- Go to offset 73B90 (use Search > Goto or just Control+G)
- What do you see there?
- What does it tell you about the document?

HxD - [C:\U	sers\ca	allee	Doc	umer	nts\te	each	ing\i	nls69	90-20	014-f	fall\e	xerci	ise-fi	iles\a	r99.	doc]		23
🔝 File Edit	Searc	h V	iew	Ana	lysis	Ext	ras	Win	dow	?								5
🗋 👌 🗸 🗐	-	U.	++	<mark>1</mark> 6		-	AN	SI		•	he	ex		-				
📓 ar99.doc																		
Offset (h) 00	01	02	03	04	05	06	07	08	09	OA	OB	oc	OD	OE	OF		
00073B80	00	1A	00	07	00	1A	00	07	00	1A	00	07	00	07	00	07		
00073890	00	FF	FF	14	00	00	00	OC	00	4B	00	65	00	72	00	72	.ÿÿK.e.r.r	
00073BA0	00	79	00	20	00	4C	00	65	00	69	00	6D	00	65	00	72	.yL.e.i.m.e.r	:
00073BB0	00	20	00	4A	00	61	00	79	00	27	00	73	00	20	00	47	J.a.y.'.sG	;
00073BC0	00	33	00	3A	00	44	00	65	00	73	00	6B	00	74	00	6F	.3.:.D.e.s.k.t.o	1
00073BD0	00	70	00	20	00	46	00	6F	00	60	00	64	00	65	00	72	.pF.o.1.d.e.r	
00073BE0	00	3A	00	61	00	Got	0							X	<u> </u>	6F	.:.a.r.9.9d.c	,
00073BF0	00	63	00	0C	00		-	-	-	-	-	_	-	-	_	20	.cK.e.r.r.y.	
00073C00	00	4C	00	65	00	Of	fset:									4A	.L.e.i.m.e.r.+.J	ŗ
00073C10	00	61	00	79	00									73B9	0	ЗA	.a.y.'.sG.3.:	
00073C20	00	54	00	65	00										- 1	72	.T.e.m.p.o.r.a.r	
00073C30	00	79	00	20	00		he	x	C	dec		0	oct			3A	.yI.t.e.m.s.:	
00073C40	00	57	00	6F	00										-	72	.W.o.r.dW.o.r	
00073C50	00	6B	00	20	00	(Offse	t rela	ative	to					- 1	41	.kF.i.l.eA	
00073C60	00	20	00	33	00		be	gin								79	3K.e.r.r.y	,
00073C70	00	20	00	4C	00	0) cu	rrent	offe	et						2B	L.e.i.m.e.r.+	
00073C80	00	4A	00	61	00		, cu	1 /I	. ons							33	.J.a.y.'.sG.3	1
00073C90	00	3A	00	54	od	C) en	d (ba	ackw	ards)	1					61	.:.T.e.m.p.o.r.a	
00073CA0	00	72	00	79	00											73	.r.yI.t.e.m.s	
00073CB0	00	3A	00	57	00					OK			Car	ncel		6F	.:.W.o.r.dW.o	
00073000	00	72	00	6B	od				-		_					20	.r.kF.i.l.e.	
00073CD0	00	41	00	20	00	33	00	UC	00	TD	00	03	00	12	00	72	.A3K.e.r.r	2
00073CE0	00	79	00	20	00	4C	00	65	00	69	00	6D	00	65	00	72	.yL.e.i.m.e.r	
00073CF0	00	20	00	4A	00	61	00	79	00	27	00	73	00	20	00	47	J.a.y.'.sG	;
00073D00	00	33	00	3A	00	44	00	65	00	73	00	6B	00	74	00	6F	.3.:.D.e.s.k.t.c	
00073D10	00	70	00	20	00	46	00	6F	00	6C	00	64	00	65	00	72	.pF.o.1.d.e.r	
00072020	00	27	00	61	00	77	00	20	00	20	00	75	00	61	00	65		6 1
ffset: 73B90																	Overwrite	

Hidden Data Exercise – Inspection in MS Word

- Do the following:
 - Open it in Word what is it?
 - If prompted to do so at the top, select "Enable Editing"
 - Select: File > Options > Trust Center > Trust Center Settings...
 - Then Privacy Options >
 Document Inspector >
 Inspect



Hidden Data Exercise – Inspection in MS Word

- Are you prompted with this?
- Why do you think this is?
- If you see this, click OK, then save the document
- Run Document
 Inspector again
- What does it tell you?



Email

- What's in an email header?
- Which parts of the header would be of most interest to you as someone responsible for managing and preserving a collection that includes email?
- Which parts of the header would be of most interest to future researchers?

Windows Artifacts



Desktop Operating System Market Share



https://gs.statcounter.com/os-market-share/desktop/worldwide

Let's make sure you can see all the files on your computer.







	Search Quick access
Folder Ontions	7
General View Search	Downloads This PC
Folder views You can apply this view (such as Details or Icons) to all folders of this type. Apply to Folders Reset Folders	Pictures This PC
Advanced settings:	Videos This PC
 Display file icon on thumbnails Display file size information in folder tips Display the following the	
 Hidden files and folders ○ Don't show hidden files, folders, or drives ● Show hidden files, folders, and drives ✓ Hide empty drives 	we'll show the most recent ones here.
 Hide extensions for known file types Hide folder merge conflicts Hide protected operating system files (Recommended) Launch folder windows in a separate process Bestore previous folder windows at logon 	
	older Options > General View Search Folder views You can apply this view (such as Details or Icons) to all folders of this type. Apply to Folders Reset Folders Reset Folders Advanced settings: Pisplay file icon on thumbnails Display file icon on thumbnails Display file size information in folder tips Display file size information in folder tips Display the full path in the title bat. Hidden files and folders O Don't show hidden files, folders, or drives Show hidden files, folders, and drives Hide empty drives Hide empty drives Hide folder merge conflicts Hide folder merge conflicts Hide protected operating system files (Recommended) Launch folder windows in a separate process Restree previous folder windows at longon

Windows Registry

- Information about:
 - Applications installed
 - Application settings
 - Hardware installed
 - Hardware settings
 - User interface and system preferences
 - User accounts
 - Locations of files and recent activities, e.g.
 Most Recently Used (MRU)
 - Lots of online activities, e.g. usernames and passwords, browsing and search query history

Five Main Registry Files

File	Description
NTUSER.DAT	One for each user account, includes information such as Most Recently Used (MRU) file lists, desktop settings, default application behaviors
SAM (Security Accounts Manager)	User account information (including passwords) and security settings
SECURITY	User and group security policies, e.g. which accounts can load device drivers, get remote access to the machine
SOFTWARE	Information about all install programs, including settings and directory paths
SYSTEM	Windows systems settings, such as drive letter mappings, storage volume information, system boot profile, last known good configuration, system name, Windows setup information, hardware profile information

Where are They Located?

_

▶ Computer ► Windows (C:) ► Windows ► System32 ► config ►

 Include in library ▼ Share with ▼ 	New folder		
Name	Date modified	Туре	Size
📙 Journal	7/13/2009 10:34 PM	File folder	
📙 RegBack	10/21/2013 12:39	File folder	
📙 systemprofile	11/20/2010 9:41 PM	File folder	
📕 TxR	2/21/2011 2:10 PM	File folder	
BCD-Template	6/28/2013 6:36 AM	File	28 KB
	10/22/2013 3:50 PM	File	43,008 KB
COMPONENTS.LOG	11/21/2010 1:33 AM	Text Document	1 KB
COMPONENTS.LOG1	10/22/2013 3:50 PM	LOG1 File	256 KB
COMPONENTS.LOG2	7/13/2009 10:34 PM	LOG2 File	0 KB
DEFAULT	10/22/2013 3:40 PM	File	512 KB
DEFAULT.LOG	11/21/2010 1:33 AM	Text Document	1 KB
DEFAULT.LOG1	10/22/2013 3:40 PM	LOG1 File	256 KB
DEFAULT.LOG2	7/13/2009 10:34 PM	LOG2 File	0 KB
netlogon.ftl	10/22/2013 3:17 PM	FTL File	3 KB
SAM	10/22/2013 7:24 AM	File	256 KB
SAM.LOG	11/21/2010 1:33 AM	Text Document	1 KB
SAM.LOG1	10/22/2013 7:23 AM	LOG1 File	21 KB
SAM.LOG2	7/13/2009 10:34 PM	LOG2 File	0 KB
SECURITY	10/22/2013 3:18 PM	File	256 KB
SECURITY.LOG	11/21/2010 1:33 AM	Text Document	1 KB
SECURITY.LOG1	10/22/2013 3:18 PM	LOG1 File	25 KB
SECURITY.LOG2	7/13/2009 10:34 PM	LOG2 File	0 KB
SOFTWARE	10/22/2013 5:13 PM	File	85,504 KB
SOFTWARE.LOG	11/21/2010 1:33 AM	Text Document	1 KB
SOFTWARE.LOG1	10/22/2013 5:13 PM	LOG1 File	256 KB
SOFTWARE.LOG2	7/13/2009 10:34 PM	LOG2 File	0 KB
CVETENA	10/22/2013 5:14 PM	File	19,456 KB
SYSTEM SYSTEM.LOG	11/21/2010 1:33 AM	Text Document	1 KB
SYSTEM.LOG SYSTEM.LOG1	11/21/2010 1:33 AM 10/22/2013 5:14 PM	Text Document LOG1 File	1 KB 256 KB

Computer ► Windows (C:) ► Users ► callee ►

clude in library 🔹 Share with 💌	New folder		
Name	Date modified	Туре	Size
🌗 .VirtualBox	10/21/2013 11:37	File folder	
AppData	3/19/2012 9:39 AM	File folder	
Application Data	7/15/2013 9:55 AM	File folder	
길 Backup	7/15/2013 12:04 PM	File folder	
📙 Contacts	9/24/2013 7:16 AM	File folder	
🔊 Cookies	7/15/2013 9:55 AM	File folder	
📜 Desktop	10/22/2013 9:26 AM	File folder	
鷆 Downloads	10/22/2013 8:36 AM	File folder	
퉬 Dropbox	7/15/2013 12:16 PM	File folder	
🙀 Favorites	9/24/2013 7:16 AM	File folder	
📴 GodMode	2/1/2010 6:40 PM	File folder	
📝 Links	9/24/2013 7:16 AM	File folder	
Local Settings	7/15/2013 9:55 AM	File folder	
My Documents	10/16/2013 12:19	File folder	
My Documents	7/15/2013 9:55 AM	File folder	
🚺 My Music	9/24/2013 7:16 AM	File folder	
E My Pictures	9/24/2013 7:16 AM	File folder	
📓 My Videos	9/24/2013 7:16 AM	File folder	
NetHood	7/15/2013 9:55 AM	File folder	
📕 Oracle	7/15/2013 11:47 AM	File folder	
PrintHood	7/15/2013 9:55 AM	File folder	
Recent	7/15/2013 9:55 AM	File folder	
퉬 Roaming	6/28/2013 4:40 AM	File folder	
Baved Games	9/24/2013 7:16 AM	File folder	
📝 Searches	9/24/2013 7:16 AM	File folder	
SendTo	7/15/2013 9:55 AM	File folder	
📧 Start Menu	7/15/2013 9:55 AM	File folder	
Templates	7/15/2013 9:55 AM	File folder	
VirtualBox VMs	10/17/2013 5:53 PM	File folder	
gitconfig	9/29/2013 5:13 PM	GITCONFIG File	0 KB
NTUSER.DAT	10/22/2013 7:26 PM	DAT File	5,888 KB
ntuser.dat.LOG1	10/22/2013 7:26 PM	LOG1 File	256 KB
ntuser.dat.LOG2	7/15/2013 9:55 AM	LOG2 File	0 KB

Registry Hives

Structure:

Hive

- Key
 - Subkey
 - Value



Example:

HKEY_CURRENT_USER\Software\Microsoft\Windows\CurrentVersion\Explorer\ RecentDocs

What do you think this is?

Registry Hives

Name	Description
HKEY_CLASSES_ROOT	Just points to HKEY_LOCAL_MACHINE\Software\Classes
HKEY_CURRENT_USER	User setting information, which is generated dynamically from HKEY_USERS when a user logs into Windows
HKEY_LOCAL_MACHINE	Hardware and software settings that are specific to this computer but shared across users (generated at startup from SYSTEM.DAT)
HKEY_USERS	Information about each of the user accounts on the computer, e.g. desktop settings, default software behaviors - generated at startup from NTUSER.DAT files, and when user logs out of applications or out of Windows, data are written back to the ntUSER.DAT files
HKEY_CURRENT_CONFIG	Just points to HKEY_LOCAL_MACHINE\Config

Question: Where would you find these registry hives on a disk image? (Hint: This is a trick question)

Registry I	Editor	-				J		
File Edit	View Favorites Help							
	A - B RecentDocs	-s *	Name	Туре	Data			
			(Default)	REG_SZ	(value not set)	1		
	, a m	-	1220	REG_BINARY	6e 00 74 00 66 00 73 00 31 00 2d 00 67 00 65 00 6e 00 30 00 2e 00 61 00 66 00 66 00 00 00 74 00 32 00 00 00 00 00 00 00 00 00 00 6e 74 66 7			
•		E F	•	DEG DELLOV				
Computer\H	Computer\HKEY_CURRENT_USER\Software\Microsoft\Windows\CurrentVersion\Explorer\RecentDocs\.aff							
	croryy		And the second second					
Registry Editor								
File Edit	View Favorites Help							
	✓ Status Bar	-	Name	Туре	Data			
	Split		(Default)	REG_SZ	(value not set)			
	-1		RU 0	REG_BINARY	6e 00 74 00 66 00 73 00 31 00 2d 00 67 00 65 00 6e 00 30 00 2e 00 61 00 66 00 66 00 00 00 74 00 32 00 00 00 00 00 00 00 00 00 00 6e 74 66 7			
•	Display Binary Data	•	•	BEG BRUDDY				
Displays a va	Refresh	F5				-		



Registry Hive Value Types

Туре	Description
REG_BINARY	Raw binary data displayed as hexadecimal*
REG_DWORD	32-bit unsigned integer (4 bytes)
REG_EXPAND_SZ	Variable-length string, usually in UTF-16 (Unicode)
REG_FULL_RESOURCE_DESCRI PTOR	Series of nested arrays used by a hardware device, binary data displayed as hexadecimal*
REG_LINK	Symbolic link to another registry key (Unicode)
REG_MULTI_SZ	Ordered list of strings (multi-string value), usually in UTF- 16
REG_NONE	No specific type – displayed as hexadecimal*
REG_QWORD	64-bit integer (8 bytes)
REG_RESOURCE_LIST	Series of nested arrays used by a hardware device, binary data displayed as hexadecimal*
REG_RESOURCE_REQUIREMEN TS_LIST	Series of nested arrays used by a hardware device, binary data displayed as hexadecimal*
REG_SZ	Fixed-length text string, usually in UTF-16

* Can be opened and viewed in a hex editor

Security ID (SID)

- One assigned to each user account
- Associated with various resources, including files, folders and Recycling Bins



Always an "S", indicating that this is an SID.

Revision level (version of the SID specification being used).

Authority that issued the SID. Value is usually "5", indicating NT Authority.

Domain identifier – value can be up to 500.


S-1-5-21-1180590209-877416012-3186324384-1002

Relative Identifier (RID), designating a specific user in the SAM file. Those below 1000 are default accounts (e.g. 500 = Administrator), and those 1000 or above are created for specific groups or users.

Examining an NTUSER.DAT File

- The files in registry (a folder within the zip file you downloaded earlier) were extracted from a full-drive (including the operating system) disk image
- The following is an example of how these files can be extracted using FTK Imager

- Navigate to: Partition 1 > [root] > Documents and Settings > Charlie > NTUSER.DAT
- Right click on NTUSER.DAT and select Export Files.



Then export the other four registry files from Windows\System32\config

🔍 AccessData FTK Imager 3.1.2.0	
<u>F</u> ile <u>V</u> iew <u>M</u> ode <u>H</u> elp	
🏩 🏩 😜 🚔 🔂 🖬 🖬 🚛 🚑 📟 🚥	🖸 🥄 🗋 🖻 🐱 😹 😵 🖕
Evidence Tree ×	File List ×
🖻 🖓 📩 system 32	Name Size Type Date Modified ^
	SAM 59 AM
	SAMLOG EileStack
	SecEvent.Evt Add to Custom Content Image (AD1) PM
	SECURITY 230 Regular The 12/10/2009 12/30/59 AM
	SECURITY.LOG 1 Regular File 12/9/2009 11:29:01 AM
	software 12.800 Regular File 12/10/2009 12:36:59 AM
	software.LOG 1 Regular File 12/10/2009 12:36:58 AM
	software.LOG.FileSlack 3 File Slack
E CatRoot	
	SysEvent.Evt 256 Regular File 12/10/2009 12:36:55 AM
E-Config +	system 25,564 Regular File 12/10/2009 12:36:59 AM
< +	4
Custom Content Sources ×	000000 72 65 67 66 11 11 00 00-11 11 00 00 74 17 49 E2 regft 🗚
Evidence:File System Path File Options	000010 30 79 CA 01 01 00 00 00-05 00 00 00 00 00 00 00 0yÊ
5	000020 01 00 00 00 20 00 00 00-00 90 34 00 01 00 00 00 ···· ····4··
	000050 00 00 00 00 00 00 00 00 00 00 00
	000000 00 00 00 00 00 00 00 00 00 00 00
New Edit Remove All Create Image	
Properties Hex Value Int Custom Conte	Cursor pos = 224; dus = 2057026; log sec = 16456208; phy sec = 16456271
Exports files from the image to a local folder	NUM //

220

Perform these same tasks in the BitCurator environment



RegRipper Instructions - BitCurator

- Navigate to Forensics Tools, and click on the RegRipper icon
- NOTE: IGNORE examples that it presents, because they use commands and syntax for Windows, not Linux
- Issue each of the following commands:

bcadmin@ubuntu: /usr/share/regripper bcadmin@ubuntu:~\$ cd /usr/share/regripper cadmin@ubuntu:/usr/share/regripper\$ perl rip.pl Rip v.2.8_20130801 - CLI RegRipper tool Rip [-r Reg hive file] [-f plugin file] [-p plugin module] [-l] [-h] Parse Windows Registry files, using either a single module, or a plugins file. -r Reg hive file...Registry hive file to parseGuess the hive file (experimental) [profile]....use the plugin file (default: plugins\plugins) p plugin module...use only this modulelist all pluginsOutput list in CSV format (use with -l) -s system name....Server name (TLN support) -u username.....User name (TLN support) -h.....his information) Ex: C:\>rip -r c:\case\system -f system C:\>rip -r c:\case\ntuser.dat -p userassist C:\>rip -l -c All output goes to STDOUT; use redirection (ie, > or >>) to output to a file. opyright 2013 Quantum Analytics Research, LLC cadmin@ubuntu:/usr/share/regripper\$

perl rip.pl -r ~/Desktop/sample-data/registry/NTUSER.DAT > ~/Desktop/ntuser-report -f ntuser

perl rip.pl -r ~/Desktop/sample-data/registry/SAM > ~/Desktop/sam-report -f sam

perl rip.pl -r ~/Desktop/sample-data/registry/SECURITY > ~/Desktop/security-report -f security

perl rip.pl -r ~/Desktop/sample-data/registry/SOFTWARE > ~/Desktop/software-report -f software

perl rip.pl -r ~/Desktop/sample-data/registry/SYSTEM > ~/Desktop/system-report -f system

*Enter each command in its entirety before hitting enter (line breaks above are simply to fit the text onto the slide, not ones that you should type yourself). Remember that the up arrow and tab can save you time when typing commands.

RegRipper Instructions – Windows I

📙 🛃 🧧 🗧 regripper-exercise				_ 0	×
File Home Share View					~ 🕐
\leftrightarrow \rightarrow \checkmark \uparrow \square \rightarrow This PC \rightarrow Desk	ctop > regripper-exercise		ٽ ~	Search regripper-exercise	Q
Home Folder mjf33 ^	Name	Date modified	Туре	Size	
💻 This PC	registry	5/18/2015 7:47 PM	File folder		
Desktop					
Documents					
🖶 Downloads					
👌 Music					
E Pictures					
Videos					
🛀 Windows (C:)					
🛖 Vol1 (F:)					
🛖 mjf33 (G:)					
👳 Duke Archives (H:)					
🛖 Vol1 (l:)					
👳 lib_archives (\\tts-ozziesmith.					
KBFS (K:)					
RBMSCL (S:)					
素 rbmscl (\\cx4-fe-nas01.oit.du					
素 rbmscl (\\oit-nas-fe01.oit.duk					
👳 rbmscl (\\oit-nas-fe01.oit.duk					
→ rbmscl (\\oit-nas-nb01.oit.du ¥					

- Create a folder called regripper-exercise on your desktop
- Find the registry directory in the folder you extracted from the saa-das-sample-data zip file earlier
- Copy this folder to the regripper-exercise folder on your Desktop

RegRipper Instructions – Windows II

- Navigate to saa-dfa-sample-data\RegRipper3.0
- Run rr.exe (it may simply appear as **rr**) by double clicking it.
- The next set of steps will be run 5 times once for each of the files in regripper-exercise\registry
- Next to the Hive File window, select Browse
 - Navigate to regripper-exercise\registry and select the first Hive File
 - E.g., NTUSER.DAT
- Next to Report File, select Browse
 - Create a new file in regripper-exercise that corresponds to the Hive File above
 - E.g., NTUSER_report.txt
- In the Profile dropdown, select the appropriate profile

- E.g., ntuser profile selection is not required in RegRipper 3.0

- Select Rip It.
- Repeat the above steps for SAM, SECURITY, SOFTWARE, and SYSTEM

RegRipper Output Questions

Examine ntuser- report.txt	Are you able to identify files that the user recently opened? If so, what were they? Can you determine what the most recently opened files of specific types (e.g. txt) were?
Examine sam-report.txt	How many accounts were there on the computer that is represented in the disk image? What is the Relative Identifier (RID) for the user account you're examining? What other interesting information can you gain from the SAM report about this user account and how might you use that information?
Examine security- report.txt	What is the Machine SID for the computer represented in the disk image? Why would you want to know this? How does it relate to the RID that you identified above?
Examine software- report.txt	Identify three different applications that were installed on the computer and the file paths where the applications were stored.
Examine system- report.txt	Find the devclass output. What does this output tell you? How might this information be useful?

RegRipper Output Discussion – ntuser-report

- Are you able to identify the files that the user recently opened? If so, what were they?
 - How did you go about finding this information?
 - What line number(s) points to this information?
- Can you determine what the most recently open files of specific types (e.g. txt) were?
 - How did you go about finding these?
 - What line numbers have this information?
- Look at lines 1109-1117 what type of information are you looking at?
- Is there any other information you find particularly compelling in this report?
- What might you do with this information?

RegRipper Output Discussion – sam-report

- How many accounts were there on the this computer?
 - How did you go about finding this information?
 - What line number(s) points to this information?
- What was the Relative Identifier (RID) for the user account you're examining?
 - How did you go about finding this?
- How many logins did Pat make on this machine?
- Is there any other information you find particularly compelling in this report?
- What might you do with this information?

RegRipper Output Discussion – security-report

- What is the Machine SID for the computer represented here?
 - How did you go about finding this information?
 - What line number(s) points to this information?
- Why would you want to know this information
- How does this relate to the RID in the previous report?

```
😑 security-report.txt 🔣
    auditpol v.20121128
 2
    (Security) Get audit policy from the Security hive file
 3
 4
    auditpol
 5
    Policy\PolAdtEv
 6
    LastWrite Time Sun Nov 8 15:34:54 2009 (UTC)
 7
 8
    Length of data: 44 bytes.
    9
 10
    11
    0x00000020: 00 00 00 00 00 00 00 00 09 00 00 00
 12
    **Auditing is NOT enabled.
13
14
    lsasecrets v.20100219
15
    (Security) TEST - Get update times for LSA Secrets
16
17
18
    Domain secret - $MACHINE.ACC
    Error: Can't call method "get value" on an undefined value at
19
    C:\Users\mjf33\Desktop\das stuff\das-forensics-flash-drive-files-excluding-vbox-bitcurator-slides-20160513\
20
```

RegRipper Output Discussion – software-report

- Identify three different applications that were installed on this computer
 - How did you go about finding this information?
 - What line number(s) points to this information?
- Why would you want to know this information?
- How might it aid description?

```
😑 software-report.txt 🔀
     Launching appinitdlls v.20130425
  2
     appinitdlls v.20130425
  3
    (Software) Gets contents of AppInit DLLs value
  4
  5 AppInit DLLs
  6 Microsoft\Windows NT\CurrentVersion\Windows
  7 LastWrite Time Fri Nov 20 18:55:34 2009 (UTC)
  8
      AppInit DLLs : {blank}
  9
      LoadAppInit DLLs : 1
    *LoadAppInit DLLs value globally enables/disables AppInit DLLS.
 10
     0 = disabled (default)
 11
 12
     Wow6432Node\Microsoft\Windows NT\CurrentVersion\Windows not found.
 13
     Analysis Tip: The AppInit DLLs value should be blank; any DLL listed
 14
 15
      is launched with each user-mode process.
 16
 17
      apppaths v.20120524
 18
      (Software) Gets content of App Paths subkeys
 19
 20
     App Paths
     Microsoft\Windows\CurrentVersion\App Paths
 21
```

RegRipper Output Discussion – system-report

- · Find the devclass output
- What does this output tell you?
- How might this information be useful?

```
🔚 system-report.txt 🔣
     ControlSet001\Control\Session Manager\AppCertDlls not found.
  2
  3
     appcompatcache v.20130425
  4
     (System) Parse files from System hive Shim Cache
  5
  6
     Signature: Oxdeadbeef
  7
     WinXP, 32-bit
  8
     C:\Program Files\AVG\AVG9\avgsrmax.exe
  9
     ModTime: Mon Dec 7 23:47:51 2009 Z
 10
     UpdTime: Tue Dec 8 17:52:13 2009 Z
 11
     Size : 361752 bytes
 12
 13
     C:\Program Files\RealVNC\VNC4\winvnc4.exe
 14
     ModTime: Thu Oct 16 01:13:58 2008 Z
     UpdTime: Tue Dec 8 01:45:23 2009 Z
 15
 16
     Size : 439632 bytes
 17
 18
     C:\RAM\mddbak.exe
 19
     ModTime: Sat Nov 14 01:07:38 2009 Z
     UpdTime: Sun Dec 6 16:10:51 2009 Z
 20
 21
     Size : 95104 bytes
 22
     C:\Program Files\Java\jre6\bin\jqsnotify.exe
 23
 24
     ModTime: Sun Oct 11 12:17:34 2009 Z
```

Viewing and Copying Registry Information if You're Running the Original Environment

- What if you're logged in to the original computer? How might you get information out of the registry?
- What if you wanted to replicate that registry information on another computer?
- Hint: There are tools built into Windows for this.

Restore Points

- Snapshots of Registry hives and some other essential system (including .EXE, .INI, .LNK) files. They're created:
 - when there are major system changes, e.g. installing software
 - at regularly scheduled intervals
 - □ if the user manually creates one
- Let's look at some restore points: Start Button > All Programs > Accessories > System Tools > System Restore [or just "System Restore" in the Start box]

Examining the Recycle Bin

- 1. In the start menu box, type cmd
- Type: cd c:\\$recycle.bin (What is this doing?)
- Type dir /a
 (What is this doing?)
- 4. Type dir *.* /s(What is this doing?)
- Put one or more files into the Recycle Bin (by moving there or by deleting)
- 6. Repeats steps 2-4. What do you see now?

A Brief Discussion of Mac Forensics

- No Registry, so where is all the good stuff stored?
- See:

https://forensicswiki.xyz/wiki/index.php?title=Ma c_OS_X_10.9 - Artifacts_Location but note that this is information a snapshot in time; artifact locations tend to change between versions of macOS.

Archival Importance and Role of SID

- If the volume is NTFS, you can find the SID associated with a specific file
- If you also have registry files from the original computer (particularly SAM.DAT), you can get information associated with that SID, such as the name of the user/group, last time he/she logged in, and various other account details

setuplog.txt

• See disk image example below: Partition 1 > [root] > WINDOWS > setuplog.txt

AccessData FTK Ima	ager 3.1.3.2			and a second sec	
File View Mode	Help				
Evidence Tree	×	File List	1		
Charlie-2009-12-08	8.aff	Name	Size Type	Date Modified	
- Partition 1 [97:	SUMB]	SET3.tmp	1,267 Regular File	4/14/2008 12:0	
	E[((1)5]	SET4.tmp	1,064 Regular File	4/14/2008 12:0	
😟 🗀 s/	AVG	SET8.tmp	17 Regular File	4/14/2008 12:0	
- te se	BadClus	setupact.log	209 Regular File	11/9/2009 1:20:	
🕀 🧰 SE	Extend	setupapi.log FileSlack	4 File Slack	11/50/2009 51	
- [# SS	Secure	setuperr.log	0 Regular File	11/8/2009 5:05:	
	en Incumente and Settinge	setuplog.txt	690 Regular File	11/9/2009 1:32:	
dr	rvitmp	Soap Bubbles.bmp	65 Regular File	4/14/2008 12:0	
🖭 🛅 Pr	rogram Files	spupdsvc.log	9 Regular File	11/10/2009 12:	
😟 🧰 Py	ython26	spupdsvc.log.FileSlack	4 File Slack		
E C R	ECYCLER	Sti_Trace.log	0 Regular File	11/8/2009 5:10:	
E Sy	visition volume information	system.ini	1 Regular File	11/8/2009 5/07	
	Shf migs	tabletoc.log	21 Regular File	11/25/2009 11:	
E -	SNtUninstallKB898461\$	TASKMANIEVE	4 File Slack	4/14/2009 12-0	
i 🔁 🔁	\$NtUninstallKB923561\$	tsoc log	182 Regular File	4/34/2000 12/0 11/05/2000 11	
🖻 🗎	SNtUninstallKB946648\$	tsoc.log.FileSlack	3 File Slack		
e 🤤	\$NtUninstallKB950762\$	li twain.dll	93 Regular File	4/14/2008 12:0	
	SNtUninstallKB950974S	twain_32.dll	50 Regular File	4/14/2008 12:0	
	SNELDrinstall/CB3510063	twunk_16.exe	49 Regular File	4/14/2008 12:0	
• E	SNtUninstallKB951748\$	twunk_32.exe	25 Regular File	4/14/2008 12:0	
•	m +	updspapi.log	39 Regular File	11/25/2009 11:	
Properties	×	updspapi.log.FileSlack	2 File Slack		
		vb.ini	1 Regular File	11/9/2009 1.18	
o Z↓		vbaddin.ini	10 Regular File	1/9/2009 118 ////////////////////////////////	
Read Data	True	WgaNotify log	5 Regular File	1/1///2009.12-0	
Write Data	True	WgaNotify.log.FileSlack	4 File Slack		
Append Data	True	wiadebug.log	1 Regular File	11/8/2009 5:10:	
Delete	True	wiaservc.log	1 Regular File	11/8/2009 5:10:	
Read Permissions	True	🗑 win.ini	1 Regular File	11/9/2009 1:22:	
Change Permissions	s False	WindowsShell.Manifest	1 Regular File	11/9/2009 1:21:	
Take Ownership	False	WindowsUpdate.log	993 Regular File	12/9/2009 12:1	
I NTES Access Cont	trol Entry	and pressess starter and			
ACE Type	Allow Access	Time, File, Line, T	ag,Message		
SID	5-1-5-32-544	11/08/2009 17:05	:52.812,d:\xpsp\	base\ntsetup\syssetup\syssetup.c,6539,BEGIN_SECTION,Installing Windows NT	
Name	Administrators	11/08/2009 17:05	:54.843,d:\xpsp\	base\ntsetup\syssetup\wizard.c,1568,,SETUP: Calculating registery size	
Accore Mark	0016016	11/08/2009 17:05	:54.906,d:\xpsp\	pase\ntsetup\syssetup\wizard.c,1599,SETUP: Calculated time for Win9x migration = 120 seconds	
Frequite File	True	11/08/2009 17:05	:54.921, d: \xpsp\	pase\ntsetup\syssetup.c.65/0,BEGIN_SECTION,Initialization	
Execute File	True	11/08/2009 17:05	:55.125,d:\xpsp\	Dase intsetup syssetup syssetup.c. 0050, BEGIN SECTION, Common Initialization	
kead Data	True	11/08/2009 17:05	.55.234 d.\xpsp\	Deschitsetup/syssetup/syssetup.or.if///pedim_scitton/initializing action log basehitsetup/syssetup.log.clill.mode Satur bas startad	
vvrite Data	True	11/00/2005 1/:05		Aussintsetup (5755etup (109.07155)/001 mode betup has statted.	
Append Data	True	11/08/2009 17:05	:55.343.d:\xnen\	base\ntsetup\syssetup\syssetup.c.1782.END_SECTION.Initializing_action_log	
Delete	True	11/08/2009 17:05	:55.781.d:\xnsn\	base htsetup) syssetup c, 1867, BEGIN SECTION, Creating setup background window	
Read Permissions	True	11/08/2009 17:05	:57.843.d:\xpsp\	base htsetup) syssetup system c.1878 EDD SECTION. Creating setup background window	
Change Permissions	s True	11/08/2009 17:05	:57.843.d:\xnsn\	base\ntsetup\syssetup\syssetup.c.1929.BEGIN SECTION.Initializing SMS support	
Take Ownership	True	11/08/2009 17:05	:57.968,d:\xpsp\	base\ntsetup\syssetup\syssetup.c,1938,,Setup: (non-critical error): Failed load of ismif32.dll.	
I NTFS Access Cont	trol Entry	11/08/2009 17:05	:57.968, d:\xpsp\	base\ntsetup\syssetup\syssetup.c,1940,END SECTION,Initializing SMS support	
	· · · · · ·	11/08/2009 17:05	:57.968, d:\xpsp\	base\ntsetup\syssetup\syssetup.c,1971,BEGIN SECTION,Shutting down power management	
		11/00/2000 17.05	. EO 021 2. \	hasa atastua aussatua a 1074 END CECETON Chutting darm sauar management	
					•
chanle-2009-12-08.aff/	/Partition 1 [9750MB]/NONAM	IE [NTFS]/[root]/WINDOWS/setu	plog.txt		

- What do you see in this file?
- What information could be useful for digital curation?
- When/how might you use it?

End User Access Scenarios*

- Virtualization and emulation
- Mounting the original filesystem
- Accessing (but not mounting) disk images using forensics software
- Remote, dynamic access to disk image contents
- Cross-drive analysis

*Note: The first three were discussed earlier

BitCurator Access

- Two-year project (October 1, 2014 September 30, 2016) at School of Information and Library Science, University of North Carolina at Chapel Hill
- Funded by Andrew W. Mellon Foundation
- Developing open-source software to support access to disk images. Core areas of focus:
 - Tools and reusable libraries to support web access services for disk images
 - Analyzing contents of file systems and associated metadata
 - Redacting complex born-digital objects (disk images)
 - Emulated access to data from disk images

BitCurator Access Redaction Tools

- Software to redact strings and byte sequences identified in disk images
- Three types of redaction actions:
 - □ SCRUB (overwrite the bytes in the target with zeroes),
 - □ FILL (overwrite by filling with a given character),
 - □ FUZZ (altering the content of a binary, so it can no longer run).
- Best used through a command-line interface but also include a graphic user interface (GUI) that supports the same functions
- Python API allowing institutions to develop custom redaction facilities using open-source tools including lightgrep

https://github.com/bitcurator/bitcurator-access-redaction

BitCurater NLP

- Funded by Andrew W. Mellon Foundation: October 1, 2016
 September 30, 2018
- Develop software for collecting institutions to extract, analyze, and produce reports on features of interest in text extracted from born-digital materials
- Use existing natural language processing software libraries to identify and report on those items likely to be relevant to ongoing preservation, information organization, and access activities
- May include entities (e.g. persons, places, and organizations), potential relationships among entities (e.g. appear together within documents or set of documents), and topic models to provide insight into how concepts are naturally clustered within the documents.







BitCurator Access Webtools + x	Kam
$\leftrightarrow \rightarrow \mathbb{C}$ (i) dogwood.ils.unc.edu:8080	☆ 🖬 🗆 🖸 🗄
Home Images Status Search text.	. Q.
Explore raw and forensically-packaged (.E01 and .AFF) disk images in a web browser. Supported file systems include FAT, EXFAT, NTFS, HFS+, EXT2/3/4, ISO 9660 (CD-ROM), and YAFFS2 (Android). Groups of images currently registered with the system are listed below.	
Image Groups All Images All Images included recursively. 12	
ISO test Set of ISO test disk images.	
Mixed test Images: 10 Set of mixed-format test disk images.	

C U dogwood.ils.unc.edu:8080/g	group/3/					x 🖬
Home Images Status					Search text	٩
Images						
Show 50 A entries					Search:	
Name	Sizo	MIME	SHA_1		Indexed	Download
	5120		514-1	v	Indexed	Download
charlie-work-usb-2009-12-11.E01	8.8MB	application/octet-stream	e49bf6048856570cc3d49b1485d6d87aaab6al	00a	2018-02-01 00:27:56	±
ext3.raw	8.0MB	application/octet-stream	a777aaf5426d2ea9bfb51d56a9edad0e8cd356	c 9	2018-02-01 00:27:56	<u>+</u>
fat12-floppy.raw	1.4MB	application/x-ima	1c5080ed2bba3b7e6d76696d9a53dbf2a68c5f	75	2018-02-01 00:29:25	±
fourpartusb1.E01	41.4MB	application/octet-stream	dfae935194f186807a3fec3260d769a212f30c5	a	2018-02-01 00:29:25	±
fpminisampler.E01	85.2MB	application/octet-stream	962721c27ccbc49e190cad576fe78327105754	26	2018-02-01 00:28:56	±
gutenbergsampler.E01	2.0MB	application/octet-stream	9629291561dbec56e10259b36c29758c23d4e	ed1	2018-02-01 00:32:02	±
hfs-plus.raw	8.0MB	application/octet-stream	39372cd3b01583ec7bb26fca9d2e4865df4965	01	2018-02-01 00:27:56	±
iso9660-joliet.iso	256.0KB	application/x-iso9660-image	de81bd1e6a43dcebf6daa45d44daa57ba7e3e	Bf5	2018-02-01 00:32:03	<u>+</u>
iso9660-rockridge.iso	256.0KB	application/x-iso9660-image	dafc241319fbc525582959238fde4958b0751f6	9	2018-02-01 00:32:03	±
nps-2010-emails.E01	506.5KB	application/octet-stream	7da1b0d8aaa1b14312830f26e2d75de47f1c47	df	2018-02-01 00:32:00	±
nps-2013-canon1.E01	5.7MB	application/octet-stream	c40dc3f87f6d902ec7355348d85c52668ddced	e5	2018-02-01 00:28:56	Ŧ
terry-work-usb-2009-12-11.E01	31.9MB	application/octet-stream	7709eca151daa2baa1db258ddb74432d54079	3ad	2018-02-01 00:31:59	Ŧ
Showing 1 to 12 of 12 entries						
					Previous 1	Next

C O dogwood.lis.unc.edu:8080/image/6/ ★ A Horne Images Status Search text Q							
More VICE 24/03/06/05/76/06/97/06/05/76/06/97/06/05/76/06/97/06/05/76/06/97/06/05/70/05/76/05/76/05/70/05/76/05/70/05/76/05/70/05/76/05/76/05/76/05/76/05/70/05/76/05/76/05/70/05/76/05/70/05/76/05/70/05/76/05/70/00/00/05/70/00/00/00/05/70/00/00/05/70/00/00/00/00/00/00/00/00/00/00/00/00/	C 🛈 dog	wood.ils.unc.edu:8080/image/	6/				* 🖬 🗆
fourpartusb1.E01 Image: State S	Home	Images Status				Search text	Q
fourpartusb1.E01 Image: Imag							
Format: Enclassifier Encl	fourr	artush1 E01					
Format: Enclass 0 Sectors T221312 MDS: Cetting 1326 Size: 370B Biocks/Sector 512 Size: Cetting 1326 Partitions: Show so t entries Search: Cetting 12230059 Id Name File System Search: Cetting 12230059 9 fourpartusb1.E01 Win95 FAT32 (0x0b) 2 Cetting 12230059 10 fourpartusb1.E01 Win95 FAT32 (0x0b) 2 Cetting 12230059 11 fourpartusb1.E01 Win95 FAT32 (0x0b) 2 Cetting 12330059 2 11 fourpartusb1.E01 Mac OS X HFS (0xaf) 3910660 3910660 3910660 3910660 3910660 3910660 12 fourpartusb1.E01 Linux (0x83) 5867520 Showing 1 to 4 of 4 entries Previous 1 Next	IOUIP						
State State State Control Cont	Format:	EnCase 6	Sectors:	7821312	MD5:	24f518cb5f95bcb6657a8e39f8ea135	54
Partitons Show 50 entries Id Name 9 fourpartusb1.E01 Win95 FAT32 (0x0b) 2 10 fourpartusb1.E01 Win95 FAT32 (0x0b) 11 fourpartusb1.E01 Mac OS X HFS (0xaf) 12 fourpartusb1.E01 Linux (0x83) Showing 1 to 4 of 4 entries	Size:	3.7GB	Blocks/Sector:	512	SHA-1:	dfae935194f186807a3fec3260d769a212f30c5	5a
Download: 1 Show 50 t entries Search:							
Show 50 t entries Search: Id Name File System Start 9 fourpartusb1.E01 Win95 FAT32 (0x0b) 2 10 fourpartusb1.E01 Win95 FAT32 (0x0b) 2 11 fourpartusb1.E01 Mac OS X HFS (0xaf) 3910660 12 fourpartusb1.E01 Mac OS X HFS (0xaf) 3910660 12 fourpartusb1.E01 Linux (0x83) 5867520 Showing 1 to 4 of 4 entries	Downloa	nd: 🔽					
Id Name Search:							
Partitions Show 50 t entries Search: Id Name File System Search: 9 fourpartusb1.E01 Win95 FAT32 (0x0b) 2 10 fourpartusb1.E01 Win95 FAT32 (0x0b) 1955331 11 fourpartusb1.E01 Mac OS X HFS (0xaf) 3910660 12 fourpartusb1.E01 Linux (0x83) 5867520 Showing 1 to 4 of 4 entries							
Show 50 + entries Id Name File System 9 fourpartusb1.E01 Win95 FAT32 (0x0b) 10 fourpartusb1.E01 Win95 FAT32 (0x0b) 11 fourpartusb1.E01 Mac OS X HFS (0xaf) 12 fourpartusb1.E01 Linux (0x83)	Doutit	<u></u>					
Id Name File System Start Image: Start <	Partito	ons				_	
9 fourpartusb1.E01 Win95 FAT32 (0x0b) 2 10 fourpartusb1.E01 Win95 FAT32 (0x0b) 1955331 11 fourpartusb1.E01 Mac OS X HFS (0xaf) 3910660 12 fourpartusb1.E01 Linux (0x83) 5867520	Show 50 ¢	ons entries				Search:	
10 fourpartusb1.E01 Win95 FAT32 (0x0b) 1955331 11 fourpartusb1.E01 Mac OS X HFS (0xaf) 3910660 12 fourpartusb1.E01 Linux (0x83) 5867520	Show 50 ¢	entries Name		♦ File Sys	tem	Search:	\$
11 fourpartusb1.E01 Mac OS X HFS (0xaf) 3910660 12 fourpartusb1.E01 Linux (0x83) 5867520	Partite show 50 ¢ Id 9	entries Name fourpartusb1.E01		File Sys Win95 FA	tem T32 (0x0b)	Search: Start 2	\$
12 fourpartusb1.E01 Linux (0x83) 5867520 Showing 1 to 4 of 4 entries	Partite show 50 ¢ Id 9 10	ONS entries Name fourpartusb1.E01 fourpartusb1.E01		 File Sys Win95 FA Win95 FA 	tem T32 (0x0b) T32 (0x0b)	Search: Start 2 1955331	\$
Showing 1 to 4 of 4 entries Previous 1 Next	Partite show 50 ; Id 9 10 11	entries Name fourpartusb1.E01 fourpartusb1.E01 fourpartusb1.E01		 File Sys Win95 FA Win95 FA Mac OS > 	tem T32 (0x0b) T32 (0x0b) X HFS (0xaf)	Search: Start 2 1955331 3910660	\$
Previous 1 Next	Partite show 50 \$ Id 9 10 11 12	entries entries Name fourpartusb1.E01 fourpartusb1.E01 fourpartusb1.E01 fourpartusb1.E01 fourpartusb1.E01		 File Sys Win95 FA Win95 FA Mac OS 3 Linux (0x4) 	tem T32 (0x0b) T32 (0x0b) X HFS (0xaf) 33)	Search: Start 2 1955331 3910660 5867520	\$
	Partite Show 50 \$ Id 9 10 11 12 Showing 1 to	ONS entries Name fourpartusb1.E01 fourpartusb1.E01 fourpartusb1.E01 ourpartusb1.E01		 File Sys Win95 FA Win95 FA Mac OS J Linux (0xi 	tem T32 (0x0b) T32 (0x0b) (HFS (0xaf) 33)	Search: Start 2 1955331 3910660 5867520	\$
	Partite Show 50 ¢ Id 9 10 11 12 Showing 1 to	entries entries Name fourpartusb1.E01 fourpartusb1.E01 fourpartusb1.E01 ourpartusb1.E01 fourpartusb1.E01 ourpartusb1.E01		 File Sys Win95 FA Win95 FA Mac OS X Linux (0xi) 	tem T32 (0x0b) T32 (0x0b) X HFS (0xaf) 33)	Search: Start 2 1955331 3910660 5867520 Previous	1 Next
	Partite Show 50 ¢ Id 9 10 11 12 Showing 1 to	entries Name fourpartusb1.E01 fourpartusb1.E01 fourpartusb1.E01 fourpartusb1.E01 fourpartusb1.E01 fourpartusb1.E01 of 4 of 4 entries		 File Sys Win95 FA Win95 FA Mac OS J Linux (0xi 	tem T32 (0x0b) T32 (0x0b) K HFS (0xaf) 33)	Search: Search: Start 2 1955331 3910660 5867520 5867520	1 Next

-lome	Images Status			· · · · · · · · · · · · · · · · · · ·	Search text	Q
Direct	ory Listing				Search:	
Туре	Filename	Bytes	Created	Modified	Download	
•	TESTFAT (Volume Label Entry)	0.0B	N/A	2013-05-02T16:01:26	±	
L.	Trashes	4.0KB	N/A	2013-05-02T14:11:00	±	
*	_RASHE~1.YEN	0.0B	N/A	2013-05-02T14:11:00	â	
*	.Trashes	4.0KB	N/A	2013-05-02T14:11:00		
*	.Spotlight-V100	4.0KB	N/A	2013-05-02T14:11:00		
L.	06311397.pdf	2.2MB	N/A	2013-05-02T15:56:06	¥	
L.	2013-02-20_AAFS.pdf	6.2MB	N/A	2013-05-02T15:56:36	<u>*</u>	
*	.fseventsd	4.0KB	N/A	2013-05-02T16:01:26		
Ŀ	\$MBR	512.0B	N/A	N/A	±	
L.	\$FAT1	953.0KB	N/A	N/A	Ŧ	
L.	\$FAT2	953.0KB	N/A	N/A	Ŧ	
_	\$OrphanFiles	0.0B	N/A	N/A		

BitCurator Access Webtools	
C O dogwood.ils.unc.edu:8080/image/6/9/2013-02-20_AAFS.pdf/	* 😭 🗆
Home Images Status	Search text Q
File Analysis for 2013-02-20_AAFS.pdf	
File Details	
Extension: .pdf	
Size: 6476327	
SHA1: 0364598548ca19deb1d4f89990a4f21e8f44e5b9	
MIME: application/pdf	
Full Text	
AAFS Digital & Multimedia Sciences Section	
Thursday, February 21, 2013 / 3:45 p.m 4:05 p.m.	
Bulk Data Analysis With Optimistic	
Decompression and Sector Hashing	
Simson L. Garnkel, Kristina Foster, Joel Young	
Naval Postgraduate School	
http://simson.net/	
1	
Bulk Data Analysis With Optimistic	
Decompression and Sector Hashing	

BitCurator Access Webtools	
C (dogwood.ils.unc.edu:8080/image/6/9/2013-02-20_AAFS	3.pdf/ 🎓 😭 🗖
Home Images Status	Search text Q
File Analysis for 2013-02-20 A	AFS.pdf
Extension: .pdf	
Size: 6476327	
SHA1: 0364598548ca19deb1d4f89990a4f21e8f44e5b9	
MIME: application/pdf	
Full Text	
AAFS Digital & Multimedia Sciences Section	AAFS Digital & Multimedia Sciences Section ORG Thursday DATE , February 21, 2013 /
Thursday, February 21, 2013 / 3:45 p.m 4:05 p.m.	DATE 3:45 p.m 4:05 p.m. TIME
Bulk Data Analysis With Optimistic	Bulk Data Analysis With Optimistic Decompression and Sector Hashing ORG ! GPE !
Decompression and Sector Hashing	GPE Simson L. Garnkel PERSON , Kristina Foster PERSON , Joel Young
!	Kevin Fairbanks PERSON , Johns Hopkins ORG Applied Physics Lab
Simson L. Garnkel, Kristina Foster, Joel Young	http://simson.net/
Naval Postgraduate School Kevin Fairbanks Johns Honkins Applied Physics Lab	
http://simson.net/	1
1	Bulk Data Analysis With Optimistic Decompression and Sector Hashing ORG ! GPE !
Bulk Data Analysis With Ontimistic	GPE Simson L. Garnkel ORG Associate Professor, Naval Postgraduate School ORG
Decompression and Sector Hashing	http://simson.net/

Topic Modeling in bitcurator-nlp-gentm (using pyLDAvis



Forensic Feature Extraction and Cross-Drive Analysis



Source: Simson L. Garfinkel, "Forensic Feature Extraction and Cross-Drive Analysis," Digital Forensics Research Workshop, August 15, 2006.

Incorporating digital forensics into archival workflows
Storage Media Acquisition and Handling Profile for Digital Repositories*



*Woods, Kam, Christopher A. Lee, and Simson Garfinkel. "Extending Digital Repository Architectures to Support Disk Image Preservation and Access." In *JCDL '11: Proceeding of the 11th Annual International ACM/IEEE Joint Conference on Digital Libraries*, 57-66. New York, NY: ACM Press, 2011. **253**

BitCurator-Supported Workflow



See: https://bitcurator.net/

Five Sources of Workflow Examples

Martin J. Gengenbach, "The Way We Do it Here': Mapping Digital Forensics Workflows in Collecting Institutions," A Master's Paper for the M.S. in L.S degree. August 2012.

https://web.archive.org/web/20170526011942/http://digitalcurationexcha nge.org/system/files/gengenbach-forensic-workflows-2012.pdf

AIMS Work Group, "AIMS Born-Digital Collections: An Inter-Institutional Model for Stewardship," January 2012. https://dcs.library.virginia.edu/files/2013/02/AIMS_final.pdf

Digital Sustainability Lab – Massachusetts Institute of Technology https://web.archive.org/web/20160408225012/http://www.dpworkshop.or g/sites/default/files/DCM-Pipeline_28Apr2015.pdf

Workflows, BitCurator Consortium https://bitcuratorconsortium.org/workflows

OSSArcFlow Project - https://educopia.org/research/ossarcflow



Figure 1. Beinecke Rare Book and Manuscript Library, Yale University

Martin J. Gengenbach, "'The Way We Do it Here': Mapping Digital Forensics Workflows in Collecting Institutions," A Master's Paper for the M.S. in L.S degree. August, 2012. 256



AIMS Work Group, "AIMS Born-Digital Collections: An Inter-Institutional Model for Stewardship," January 2012.

257



Kari Smith, Massachusetts Institute of Technology.

https://web.archive.org/web/20160408225012/http://www.dpworkshop.org/sites/default/files/DCM-Pipeline_28Apr2 258 015.pdf

BitCurater CONSORTIUM

Member Login

Using BitCurator - Get Involved

Search

ed - Events

Q

Using BitCurator

Getting Started

Documentation

Workflows

Videos

Not a member?

Much of the content on BitCuratorConsortium.org is accessible to members only. Learn more about the benefits of joining the BCC. The following workflows depict the step-by-step processes BitCurator Consortium members follow to acquire, process, describe, and store the born-digital materials in their collections. Most of these resources are only accessible to members. Learn more about the benefits of membership.

Why Digital Forensics -

If you are interested in adding a workflow to our listing, please contact us.

About Us -

Workflow

Title	Contributor	Release Date	
Processing Workflow	The University of Maryland, Libraries	2016 March 22	
Princeton University Archives (Members Only)	Princeton University	2015 June 30	
Penn State Born Digital (Members Only)	Penn State University	2014 July 29	
Duke University Archives	Duke University	2012 August 12	
Beineke Rare Books and Manuscripts Library	Yale University	2012 August 12	
Maryland Institute for Technology in the Humanities	The University of Maryland, MITH	2012 August 12	
University of North Carolina, Chapel Hill, Archives	University of North Carolina Chapel Hill, SILS	2012 August 12	
University of Virginia Libraries	University of Virginia	2012 August 12	
Yale University, Manuscripts and Archives	Yale University	2012 August 12	

https://bitcuratorconsortium.org/workflows

EDUCOPIA INSTITUTE



Research

Continuing Education

Nexus

Mapping the Landscapes

Digital Preservation

Aligning National Approaches to Digital Preservation (ANADP)

Chronicles

Distributed Digital Preservation (DDP)

Electronic Theses and Dissertations

Identifying Continuing Opportunities for National Collaboration (ICONC)

OSSArcFlow

News on the Margins

Scholarly Communication

Chrysalis

Developing A Curriculum to Advance Library-Based Publishing

Incubating Programs and Ideas

Digital Preservation | OSSArcFlow

OSSArcFlow



Contact: Katherine Skinner

Additional Documents:

Investigating, Synchronizing, and Modeling a Range of Archival Workflows for Born-Digital Content

Project Abstract

The Educopia Institute, in collaboration with the University of North Carolina at Chapel Hill School of Information and Library Science (UNC SILS), LYRASIS, and Artefactual, Inc., are investigating, synchronizing, and modeling a range of workflows to increase the capacity of libraries and archives to curate born digital content. These archival workflows will incorporate three leading open source software (OSS) platforms—BitCurator, Archivematica, and ArchivesSpace—and the project will be designed to generate findings that can be generalizable to settings that are using other platforms and applications.

This project will significantly impact curation practices by increasing our understanding of how institutions of different sizes and types may engage in OSS tool integration and workflow development. Our findings will be used to support a broad range of libraries and archives actively collecting and curating digital content. The knowledge gained by working with multiple institutions of different types and sizes will also broaden field-wide understanding of curation approaches and priorities, and how those impact the use of tools and capabilities in Archivematica, ArchivesSpace, and BitCurator. We expect the empirical findings about institutional needs, as well as formal workflow models, to contribute to digital curation research literature.

This project has been generously funded by the Institute of Museum and Library Services.

Project Outputs

Digital Dossiers

https://educopia.org/research/ossarcflow

Challenges

- Incorporation into LAM workflows,
 e.g. metadata conventions, connections to collection management systems
- Obsolete storage media and filesystems
- Dealing with large, internally complex data files
- Provision of public access
- Defining and implementing ethical commitments

SWAT (Software and Workstations for Antiquated Technology) Sites

"A community-based approach would use SWAT sites wherein a few self selected institutions acquire and maintain the gear and expertise to read data and transfer content from particular types of obsolete media. The SWAT sites would provide transfer services for institutions that don't have the capacity to read a particular medium (or the SWAT sites might become the likely places to deposit particular types of media)."

Erway, Ricky. "Swatting the Long Tail of Digital Media: A Call for Collaboration." Dublin, Ohio: OCLC Research, 2012. http://www.oclc.org/research/publications/library/2012/2012-08.pdf

See also:

Ricky Erway and Ben Goldman, "Agreement Elements for Outsourcing Transfer of Born Digital Content," August 2014,

http://oclc.org/research/publications/library/2012/2012-06r.html#agreement

Legal and Ethical Issues

Ethical Dilemmas

- What ethical dilemmas related to borndigital materials have you faced or do you expect to face?
- What would the competing interests or values be?
- How would you decide?

Donor Agreements

- Donor agreements (as of 2012) tend not to address the kinds of issues raised in this class*
- What are the most important issues to resolve with creators/donors?
- What's the right level of detail to include in donor agreements and discussions with potential donors?

*Matthew J. Farrell, "Born-Digital Objects in the Deeds of Gift of Collecting Repositories: A Latent Content Analysis," A Master's Paper for the M.S. in L.S degree. July 2012, <u>https://cdr.lib.unc.edu/indexablecontent?id=uuid:385c4fd9-a403-4ba3-85ac-2ea128400ddb&ds=DATA_FILE</u>

Specific Guidance Documents

- Redwine, Gabriela, Megan Barnard, Kate Donovan, Erika Farr, Michael Forstrom, Will Hansen, Jeremy Leighton John, Nancy Kuhl, Seth Shaw, and Susan Thomas. "Born Digital: Guidance for Donors, Dealers, and Archival Repositories." Washington, DC: Council on Library and Information Resources, 2013.
- Nelson, Naomi L, et al. "Gift/Purchase Agreements." In Managing Born-Digital Special Collections and Archival Materials, 122-126. SPEC Kit 329. Washington, DC: Association of Research Libraries, 2012. [Includes donor agreements and policies from Duke University, Bentley Historical Library, and Beinecke Rare Book and Manuscript Library]
- Pyatt, Timothy D. "Deed of Gift Addenda for Collections with Electronic Records." Pennsylvania State University. 2012.

A Guide to Deeds of Gift – Society of American Archivists*

Text added in 2013:

"Be aware that any digital materials that you donate, including computers, computer disks, and other digital storage media, may contain passwords, web browsing history, other users' files, and copies of seemingly deleted files. Whether or not these files are apparent to researchers will depend on the initial method of transfer and on the repository's access policies and procedures for handling digital material, which may change over time as technology evolves. Discuss any concerns about deleted content with the archivist or curator."

*https://www2.archivists.org/publications/brochures/deeds-of-gift

What does it mean for electronically stored information (ESI) to be "accessible"?

"The person responding need not provide discovery of electronically stored information from sources that the person identifies as **not reasonably accessible because of undue burden or cost**." (Rule 45 (d)(1)(D)) (emphasis added)

Judge Shira Scheindlin:

"[t]he more information there is to discover, the more expensive it is to discover all the relevant information until, in the end, 'discovery is not just about uncovering the truth, but also about how much of the truth the parties can afford to disinter.' "

(Zubulake I, 217 F.R.D. at 311 (quoting Rowe Entm't, Inc. v. William Morris Agency, Inc., 205 F.R.D. 421, 423 (S.D.N.Y. 2002)).



Seven-Factor Test from Zubulake v. UBS Warburg

- 1. extent to which the request is specifically tailored to discover relevant information
- 2. availability of such information from other sources
- 3. total cost of production, compared to the amount in controversy
- 4. total cost of production, compared to the resources available to each party
- 5. relative ability of each party to control costs and its incentive to do so
- 6. importance of the issues at stake
- 7. relative benefits to the parties of obtaining the information

(217 F.R.D. at 322)

Zubulake's Five Categories of ESI (Most to Least Accessible)*

- Active, online data
- Near-line data
- Offline storage
- Backup tapes
- Erased, fragmented or damaged data

*See: Lange, Michele C. S., and Kristin M. Nimsger. *Electronic Evidence and Discovery: What Every Lawyer Should Know Now*. 2nd ed. Chicago, IL: Section of Science & Technology Law American Bar Association, 2009. p.75.

Magistrate Judge John Facciola:

"...I am anything but certain that I should permit a party who has failed to preserve accessible information without cause to then complain about the inaccessibility of the only electronically stored information that remains"

(*Disability Rights Council of Greater Wash. v. Wash. Metro. Transit Auth.,* 242 F.R.D. 139 (D.D.C. 2007)).

Rights to Control Information

- Most frequently discussed in library lit is copyright
- Claims can extend far beyond intellectual property rights, as defined by law
- Cultural property, replevin and repatriation
- Right to privacy
- Protection of human subjects in research
- Privileged or protected information (e.g. client-attorney, healthcare, social services, library circulation, source – journalist)
- Right to publicity individual's protection from unauthorized commercial use of her name, persona, or likeness
- Prevention of misappropriation (including plagiarism)

"If a forensic examiner has complete confidence in his/her conclusions, this is usually an indication that he/she is missing something – there is always uncertainty and all assertions should be qualified accordingly..."

Casey, Eoghan. "Error, Uncertainty, and Loss in Digital Evidence." *International Journal of Digital Evidence* 1, no. 2 (2002).

"Investigators cannot, in general, directly observe digital data and instead they can only observe the data displayed on a monitor or other output device, which is driven by various types of hardware and software. Because the **observation of the data is** indirect, a hypothesis must be formulated that the actual data is equal to the observed data. Testing this hypothesis requires that the hardware and software being used are accurate and reliable. Hypotheses also need to be formulated about the data abstractions that exist and the previous states and events that occurred."*

*Carrier, Brian D. "A Hypothesis-Based Approach to Digital Forensic Investigations." Doctoral Dissertation, Purdue University, 2006. p.11 (emphasis added).

Examples of Potentially Useful Inferences (that could be wrong)

- Name embedded in a MS Word file is the document's author
- Given IP address identifies an individual
- Presence of email addresses on different hard drives indicate correspondence patterns between individuals
- Many common MD5 values across storage locations indicate sharing of files across those locations (contextbased filtering can help to address this)
- Last modified date indicates when a document was finalized
- Parts of a page available through the WayBack Machine for a given date represent the parts of the page as available on that date

The "Keyboard Dilemma"

"Even if a document can be traced to a particular computer and/or IP address, how can we identify who was actually at the keyboard composing the document? It is a particular problem in environments where multiple users may have access to the same computer or when users do not have to authenticate themselves to access a particular account."

Chaski, Carole. "The Keyboard Dilemma and Authorship Identification." In Advances in Digital Forensics III: IFIP International Conference on Digital Forensics, National Center for Forensic Science, Orlando, Florida, January 28-January 31, 2007, edited by Philip Craiger and Sujeet Shenoi. New York, NY: Springer, 2007. p.133.

Shared Computer Use in the Home*

*Frohlich, David, and Robert Kraut. "The Social Context of Home Computing." In *Inside the Smart Home*, edited by Richard Harper, 127-62. London: Springer, 2003.

Ethics Questions To Consider

- When acquiring a disk as part of a collection, should you create a bit-level image of the disk, in order to ensure the potential to recreate not only the "payload" data of files but also various forms of information within and below the filesystem?
- 2. Should you retain "hidden" data in a Word document or only retain what you assume to be the text that the author intended?
- 3. You're responsible for managing a Microsoft Outlook .pst file over time (including saved and sent messages, calendar items, draft and deleted messages, address book, and viruses). Should you retain the whole .pst file or extract messages and attachments that were sent and received?
- 4. If a collection documents the life of an individual, how would you determine the appropriate scope for collecting information associated with that person's online presence (e.g. postings, affiliations, profiles, micro-contributions)?
- 5. If your institution routinely "normalizes" submitted files into designated file formats, are you obligated to ensure that the normalization doesn't violate the intentions of creator or other interested stakeholders? If so, what does this obligate you to do specifically?
- 6. Someone cropped a set of images in order to remove sensitive parts, but the images still have pixel information and embedded thumbnail reflecting the "removed" parts. How should you approach the management of the images?

Lessons and Insights

- Digital forensics has arrived for archival processing
- Introduction of digital forensics doesn't dictate specific policies or practices
- The disk image is the cornerstone of many forensics methods
- "Taking bitstreams seriously" can have major advantages
- Disk images afford new access scenarios

To Learn More About Available Software

Forensics Wiki. <u>https://forensicswiki.xyz/page/Main_Page</u>

BitCurator Environment. <u>https://bitcurator.net</u>

BitCurator Software Overview. https://bitcurator.github.io/

Community Owned digital Preservation Tool Registry (COPTR) <u>https://coptr.digipres.org/Main_Page</u>

Information Guides on Tools for Electronic Records. Minnesota State Archives. <u>https://www.mnhs.org/preserve/records/electronicrecords/erpreserveres.php</u>

Lifecycle Tools for Archival Email Stewardship. https://docs.google.com/spreadsheets/d/1V1N22xnr5e0EbDIZWx58bjYO6rkrMr YH9wGX9-CK8c4/

Tools for processing, managing, and preserving electronic records. University of Minnesota. https://www.lib.umn.edu/dp/guides

Online Forums

BitCurator User Group

https://groups.google.com/forum/#!forum/bitcurator-users



Digital Curation List

https://groups.google.com/forum/#!forum/digital-curation

Further Education

		ATA [.]		00	9	SUPPORT	CONTACT US
	SOLUTIONS V	SERVICES & TRAINING -	RESOURCES ~	PARTN	ERS ~	сом	PANY ~
-	Training Overvi	ew					
AccessData Training is conducted by Syntricate , a trusted training partner championed by a					€	CONTACT SYNTRI	CATE
		hrough Synfricate, AccessData is abl d eDiscovery Training, Choose from i d training videos, or Custom Training			•	MEET THE TRAINE	ERS



Syntricate will educate you in technology and prepare you with innovative ideas and workflows to improve and strengthen your skills to identify, respond, investigate, prosecute, and adjudicate cases. The Syntricate computer forensic team focuses on how to properly collect, process, review and report case data toward successful case resolution. View Syntricate's Computer Forensic <u>courses</u> and <u>syllabl</u>, <u>course</u> <u>calendar</u>, <u>on-demand training videos</u>, <u>annual training passes</u>, <u>custom</u> <u>training</u>, and <u>certifications</u>. Or, simply <u>contact Syntricate</u> and we'll set up a call to discuss our offerings and which type of training and delivery method suits your specific needs.

GUIDANCE		COMPANY PRODUCTS & SERVICES RESOURCES NEWS & EVENTS SUPPORT	PARTNERS
	Courses		
	Series: Any Series Loc	Latien: Any Location Date: Any Date	Methods RESET
	Level	Title	Delivery Method
		EnCase v7 OnDemand Forensics and Enterprise Fundamentals Training	OnDemand
		EnCase⊗ v7 OnDemand Forensics and Enterprise Fundamentals Training	OnDemand
		EnCase® OnDemand eDiscovery for the Legal Team	OnDemand
		Digital Media – Acquisition and Triage	Classroom
		EnCase® v7 Computer Forensics I	Classroom
	Introductory	EnCase® v7 vClass Computer Forensics I	vClass



Digital Archivist Boot Camp \$1795.00 REGISTER / ORDER

Thank you!

Go forth and curate the bits!



BitCuratorEdu

Advancing the adoption of digital forensics tools and methods in libraries and archives through professional education efforts



This resource was released by the BitCuratorEdu project and is licensed under a <u>Creative Commons</u> <u>Attribution 4.0 International License</u>.

Most resources from the BitCuratorEdu project are intentionally left with basic formatting and without project branding. We encourage educators, practitioners, and students to adapt these materials as much as needed and share them widely.

The <u>BitCuratorEdu project</u> is a three-year effort funded by the <u>Institute of Museum and Library</u> <u>Services (IMLS)</u> to study and advance the adoption of digital forensics tools and methods in libraries and archives through professional education efforts. This project is a partnership between <u>Educopia Institute</u> and the <u>School of Information and Library Science at</u> <u>the University of North Carolina at Chapel Hill</u>, along with the <u>Council of State Archivists (CoSA)</u> and several Masters-level programs in library and information science.