Managing privacy and access with digital forensics tools and techniques

Creighton Barrett

Right to Know: Balancing Access and Privacy

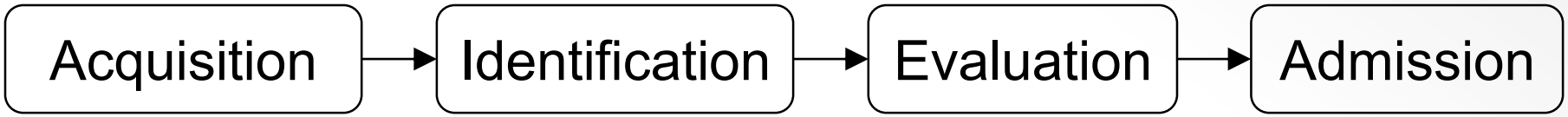
Symposium at Dalhousie University

September 28, 2017

DALHOUSIE UNIVERSITY 1818 2018

What is digital forensics?

- Forensic science – recovery and investigation of data found in digital storage devices
- Primarily used by specially trained professionals in criminal investigations, corporate investigations, etc.
- Archives are adopting digital forensics techniques to support acquisition, accessioning, appraisal, preservation, and **access**

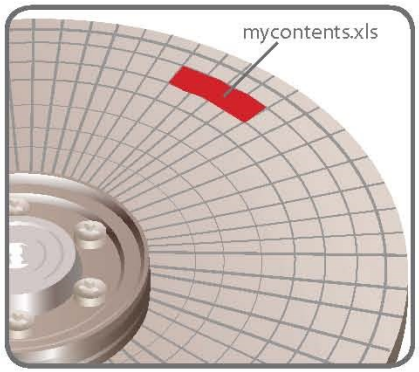


Source: Infosec Institute, Digital Forensic Models (January 25, 2016): http://resources.infosecinstitute.com/digital-forensics-models/

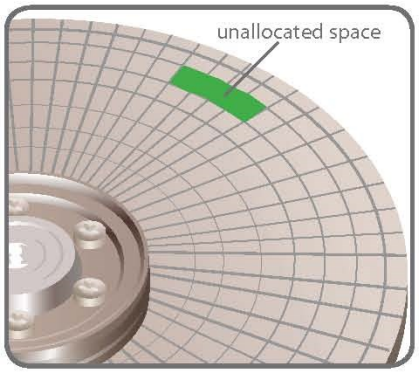
How are Deleted Files and Data Recovered?

Computers Don't Immediately Remove Data that is Deleted

Original Data

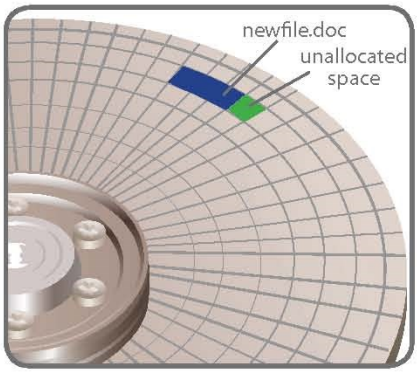


Deleted Data



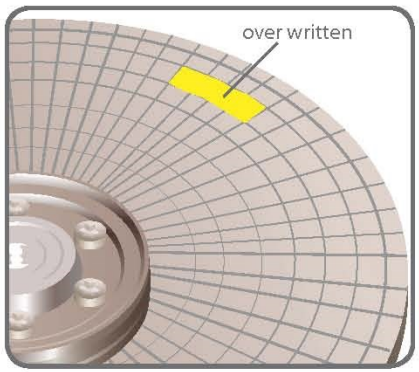
The original data is still present, but marked as unallocated space.

Partially Overwritten Data



Over time, some or all of the data can be over written. The remaining data can still be "carved" and reviewed.

Data Wiped Clean or Shreaded



The data can be wiped clean or shreaded using privacy software.

What is unallocated space?

Unallocated Space is available disk space that is not allocated to any volume. The type of volume that you can create on unallocated space depends on the disk type. On basic disks, you can use unallocated space to create primary or extended partitions. On dynamic disks, you can use unallocated space to create dynamic volumes.

PINPOINT LABORATORIES

www.pinpointlabs.com

Preserve information about the operating system and file system

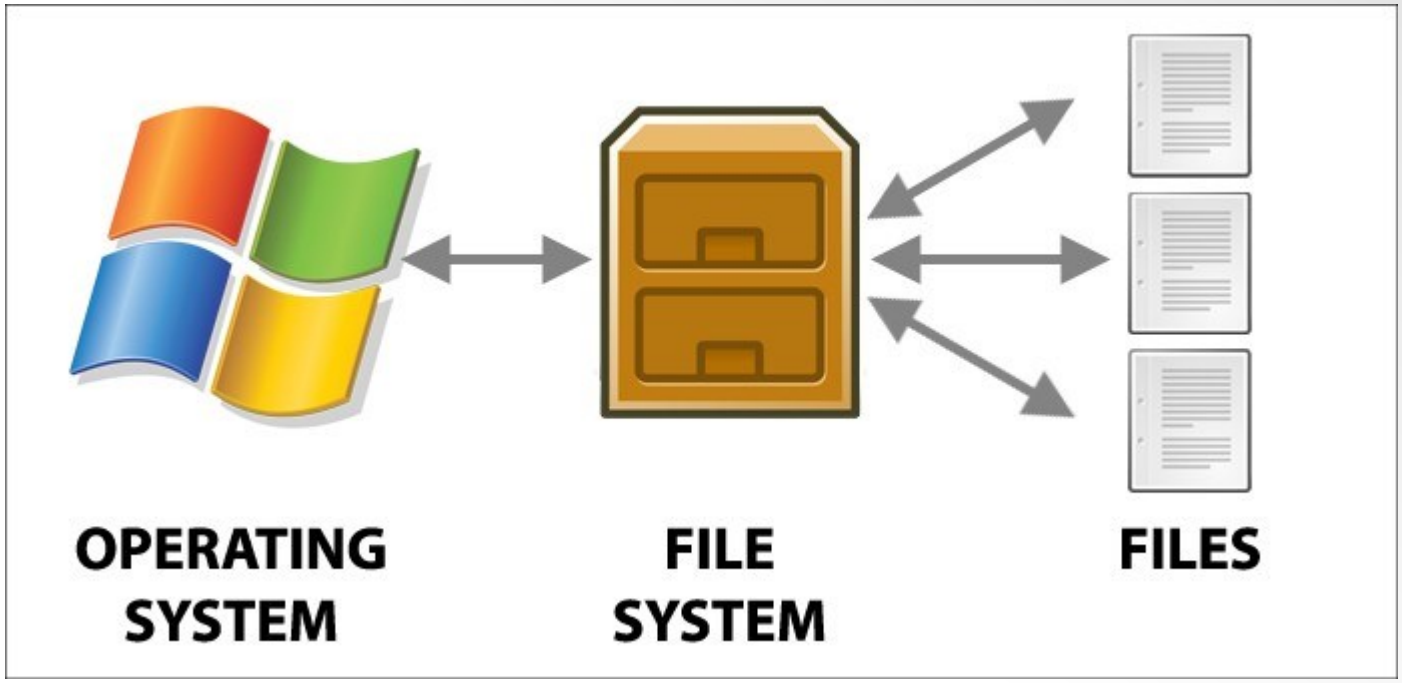


Image source: Power Data Recovery: https://www.powerdatarecovery.com/hard-drive-recovery/volume-not-contain-recognized-file-system.html

Digital forensics concepts in context

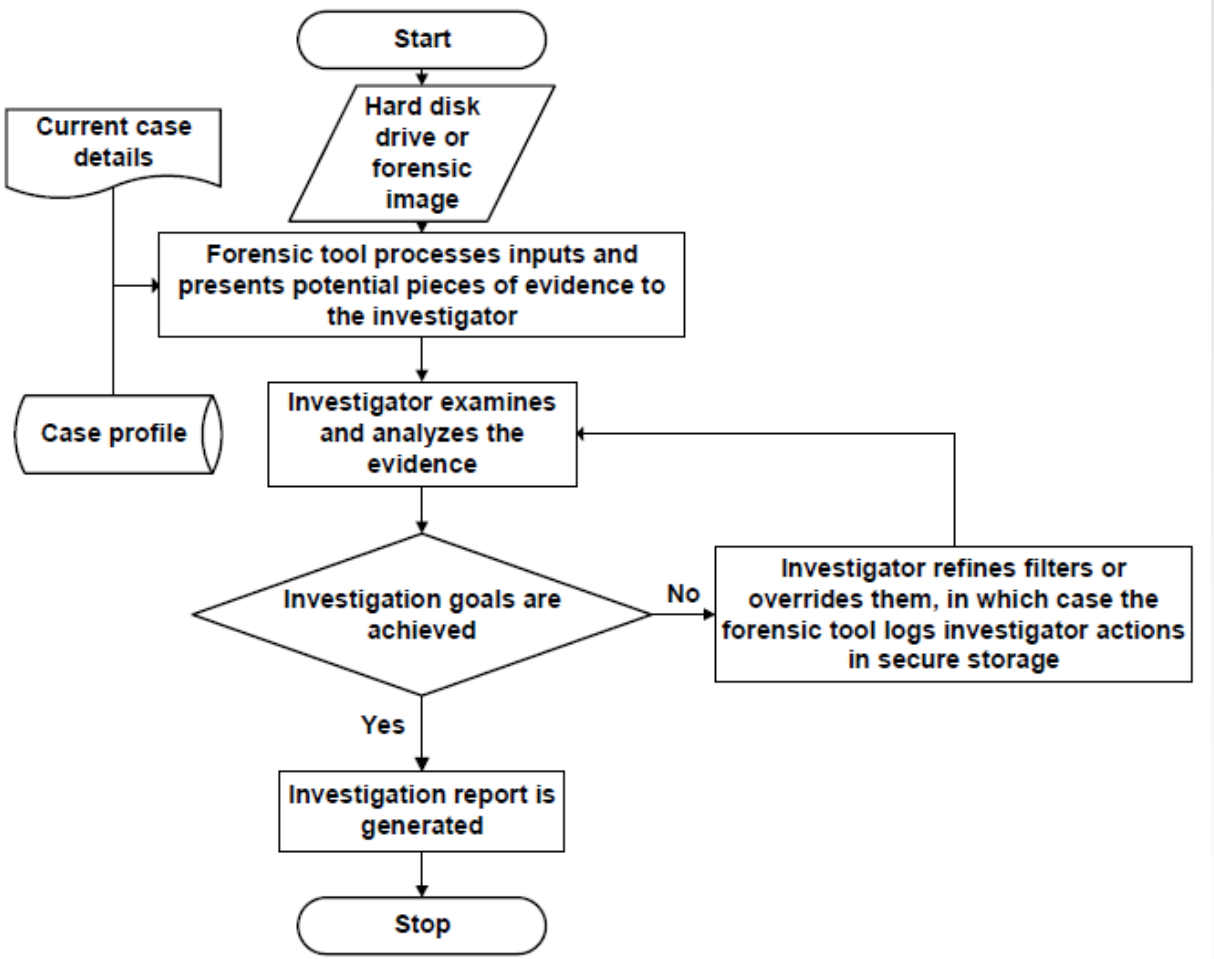
Concept	Law enforcement	Libraries, archives, and museums
Chain of custody	Incident response, investigation	Acquire, appraise, and preserve data in accordance with accepted archival practice
Private data / personal information	Exploited for prosecution	Identified and redacted to protect privacy and comply with legislation
Storage	Secure storage of digital evidence	Long-term storage of digital objects in accordance with TDR standards
Access to information	Who is entitled to see the evidence?	Who is entitled to see the digital heritage collections?
Constraints	Time to trial	Collections backlogs

Adapted from: Kam Woods, Preservation, Privacy, and Access: Enhancing Digital Curation Workflows with Forensic Analysis (March 21, 2017): http://wiki.bitcurator.net/downloads/kwoods-unc-digpres-v12.pdf

Privacy concerns in digital forensics

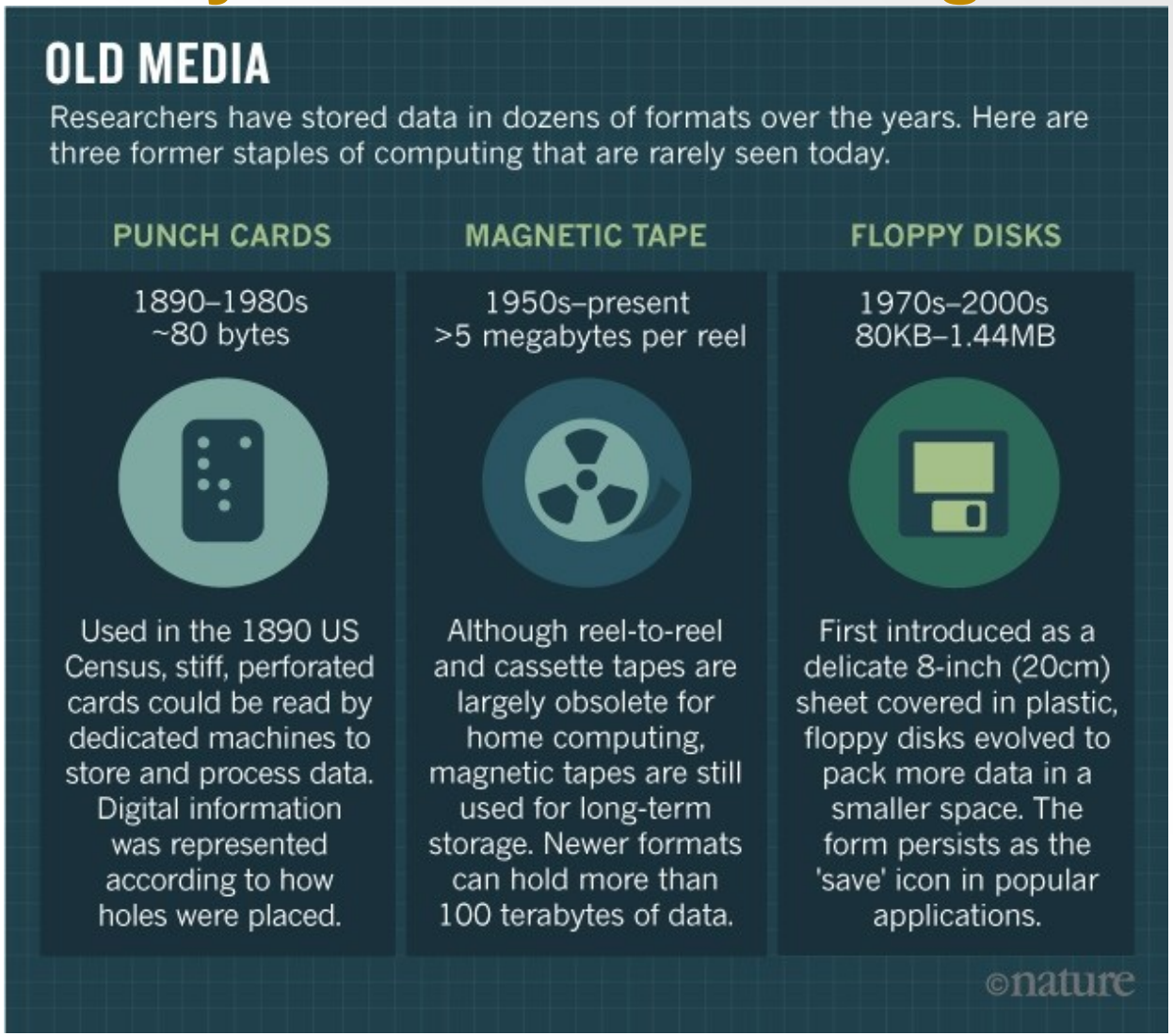
- Investigations reveal passwords, encryption keys, images, personally identifying information, etc.
- Scope of investigation
- Legal requirements
- Balance between protecting privacy and conducting complete investigations
- Lack of awareness or concern about privacy
- Lack of standardized ethical principles

Privacy solutions in digital forensics



Verma, Govindaraj, and Gupta. "Data privacy perceptions about digital forensic investigations in India." Published in *Advances in Digital Forensics XII,* edited by Gilbert Peterson and Sujeet Shenoi. Springer: 2016. https://link.springer.com/book/10.1007%2F978-3-319-46279-0

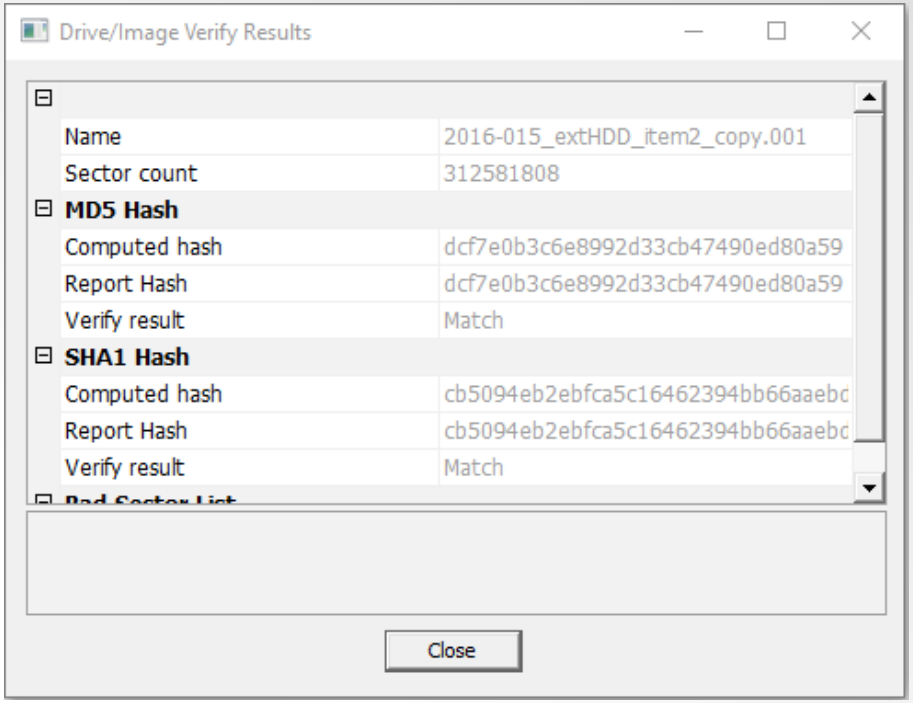
Why have a digital forensics lab in a library or archives setting?



Source: Baker, M. (2017, May 2). Disks back from the dead. *Nature, 545* (7652), 117–118. https://doi.org/10.1038/545117a

How are archivists doing digital forensics work?

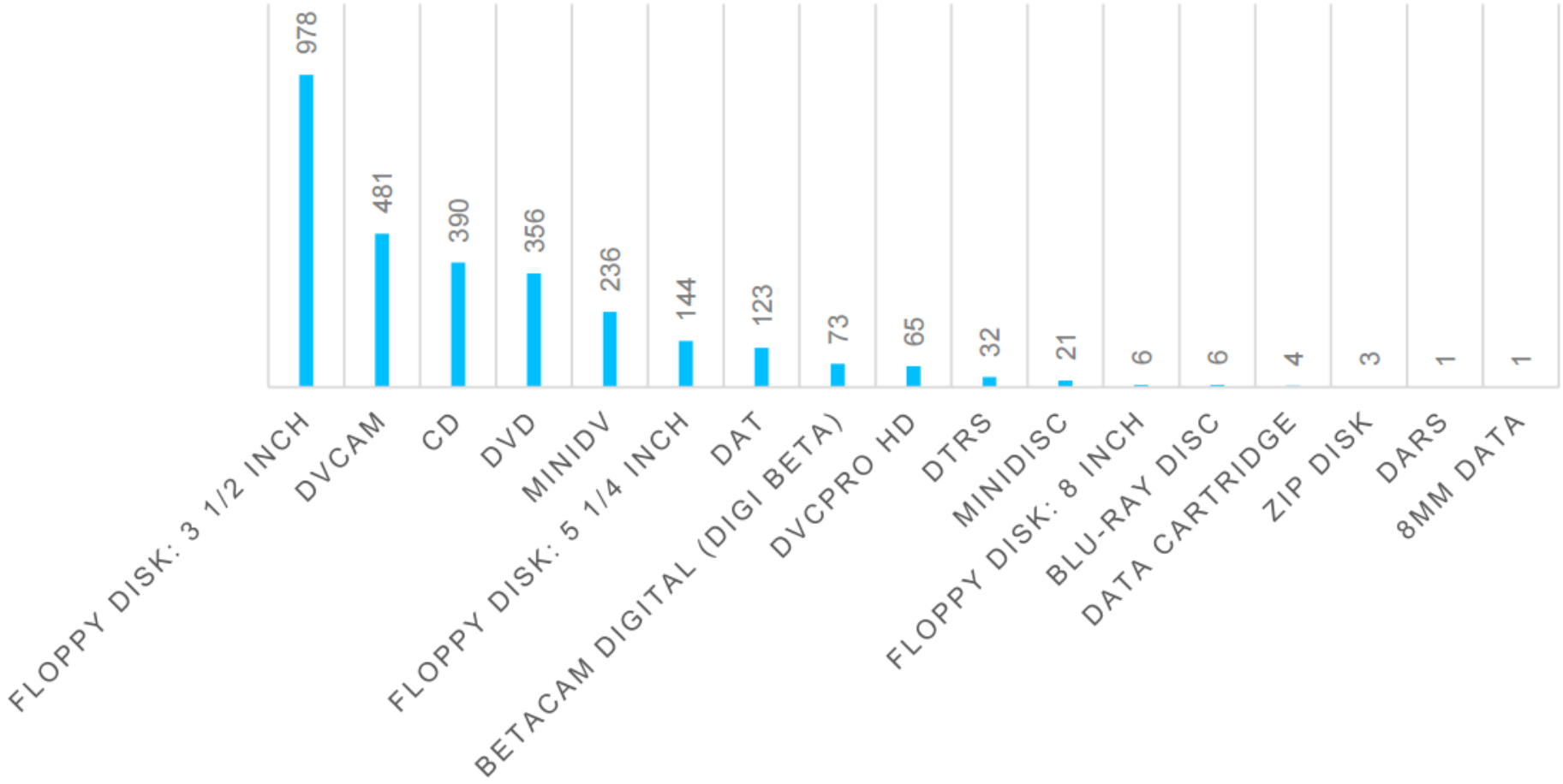
- Use write-blockers to create forensic images
- Adopt forensic software
- Incorporate digital forensics into workflows
- New policy decisions (e.g., preserve forensic image or extract files?)



Timeline at Dalhousie

- February 2016 – Acquire forensic workstation
- May – November 2016 – Digital archives collection assessment project: http://hdl.handle.net/10222/72663
- January 2017 – Install BitCurator and Forensic Toolkit (FTK) software
- February 2017 – Advanced computer forensics training
- May 2017 – Launch digital forensics lab
- April 2017 – Dal's first time at BitCurator Users Forum

DIGITAL MEDIA CARRIER FORMATS



Policy, ethical, and legal questions for libraries and archives

- Recovery of deleted files if they appear to be archival?
- Decryption of EFS files? Other encryption methods?
- Use of Password Recovery Toolkit?
- Use of registry information, browser history, etc. to support archival appraisal?
- Modifications to standard deed of gift template?
- Monetary appraisal of born-digital archival material?

Privacy at Dalhousie

- FOIPOP Act
- Policy for the protection of personal information from access outside Canada (approved in 2007)
- Data classification schema (approved in 2013)
- LOTS of personal and confidential information in various locations
- Need to consider access requirements of the Archives Permanent Collection when policies are developed and updated

Dalhousie data classification schema

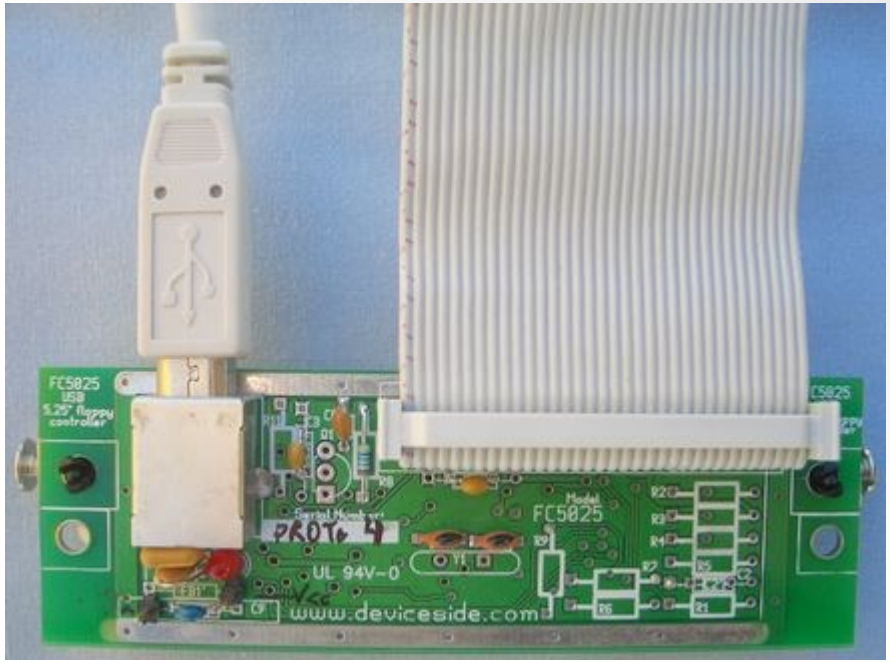
Level Number	Description
Highly sensitive	Information which may result in significant and substantial harm to the university or members of the university community, or which may violate legal or contractual requirements, if it were to be released
Sensitive	Information which could have a negative impact on the university or members of the university
Internal use	Information made available to faculty or staff of the university but not necessarily appropriate for the general public. (Directory listings, minutes from non-confidential meetings, internal websites, etc.)
Public use	Information that can be made generally available to the public.

Digital forensics tools


FORENSIC COMPUTERS, INC.
Tableau Forensic Bridge
GUIDANCE
Forensic Computers
FORENSIC
COMPUTERS
READ ONLY
SDHC/SDXC
CF/UDMA
MS/DUO
READ ONLY
M-DISC
LG
LIAN LI


Tableau Forensic Bridge
T356789iu
GUIDANCE SOFTWARE
Write Block
Read/Write
Power
Device
Host
Activity
IDE
SATA/SAS
SATA Gen3
Drive Power
PCIe
FireWire
USB 3.0

Device Side Data's FC 5025 USB floppy controller



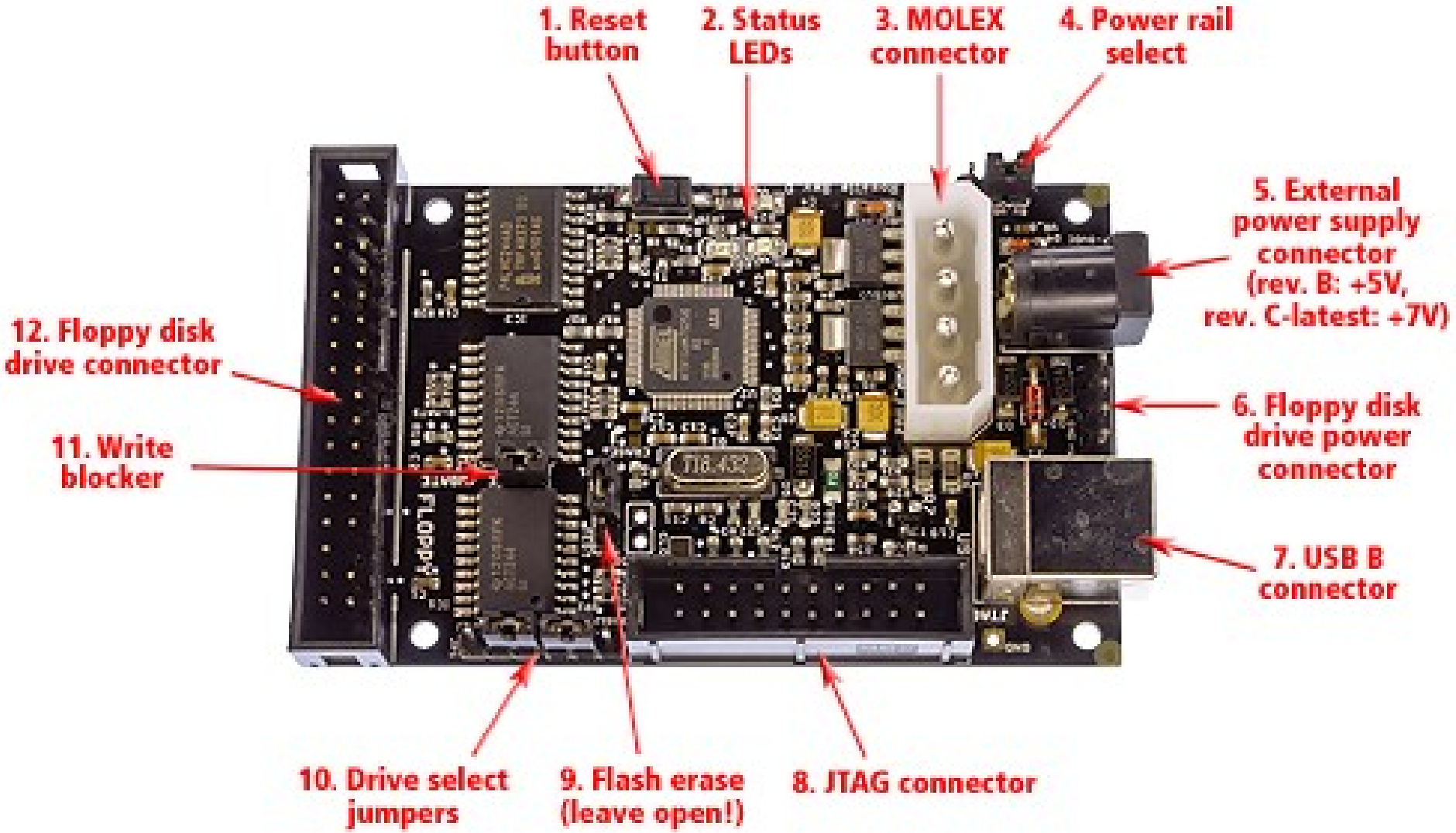


Image source: KryoFlux: https://kryoflux.com/?page=kf_tech

BitCurator

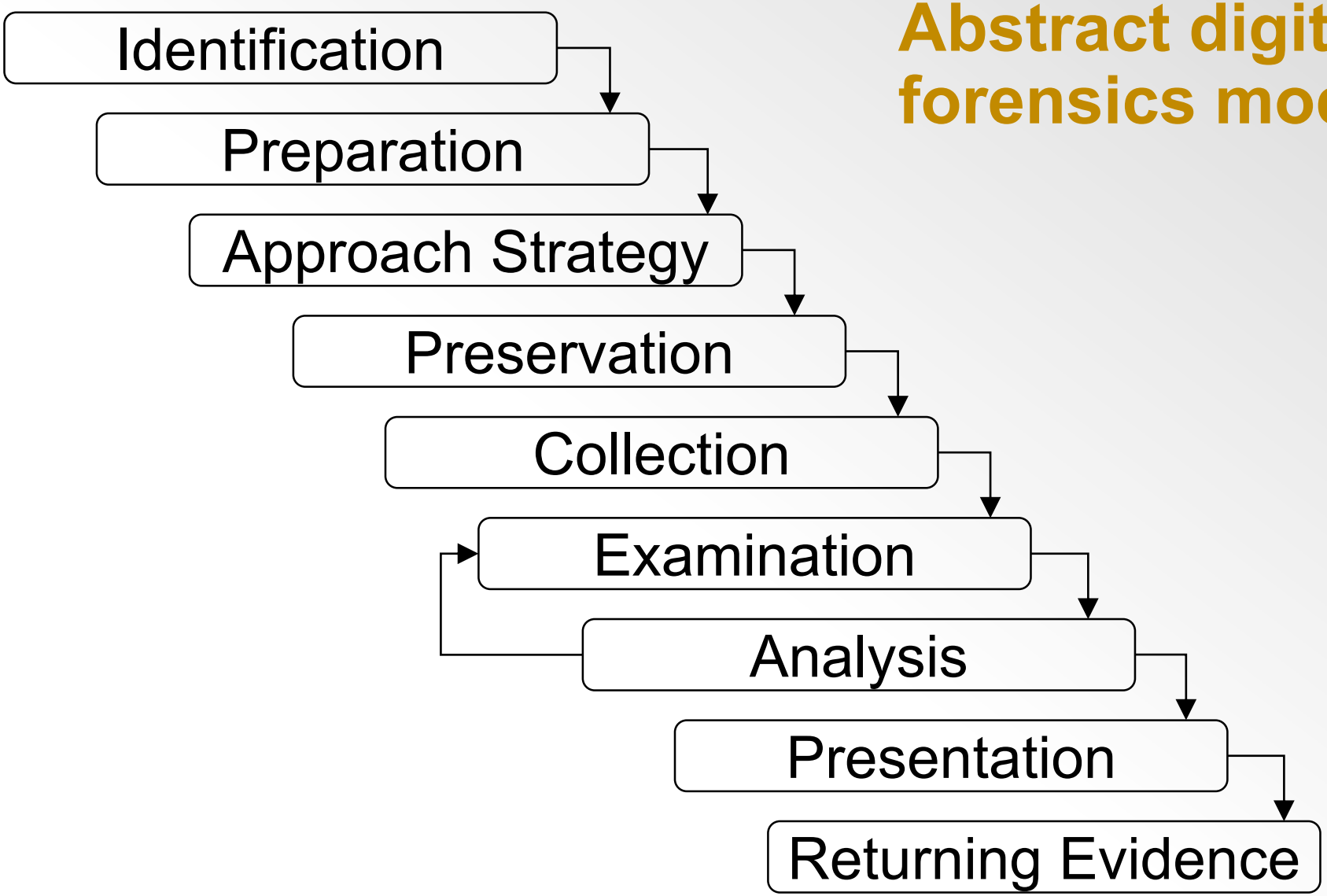
EnCase®

Forensic Toolkit (FTK)

- Three components
  - Database (Oracle, PostgreSQL, Microsoft SQL)
  - Graphical user interface (GUI)
  - Known file filter server (contains datasets with hash values for known file types)
- Indexing, live search, regular expression
- Oracle “Outside In” technology for previewing most file types
- Integration with other AccessData products (Registry Viewer and Password Recovery Toolkit)

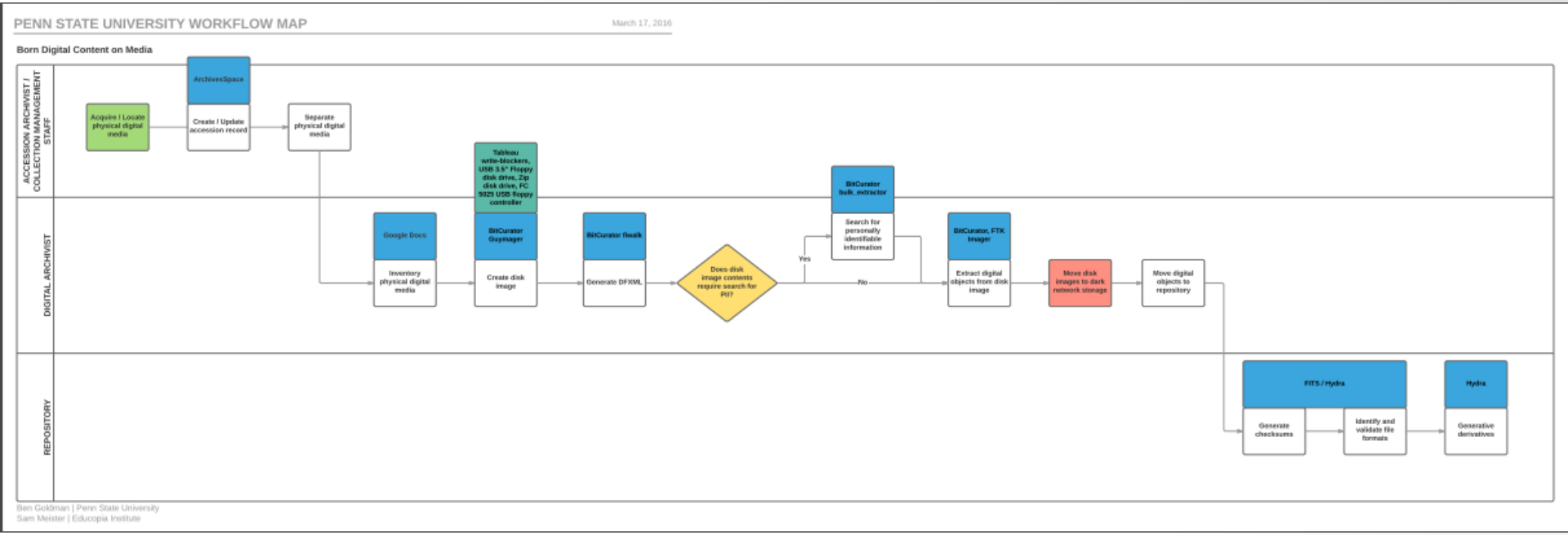
Digital forensics workflows

Abstract digital forensics model

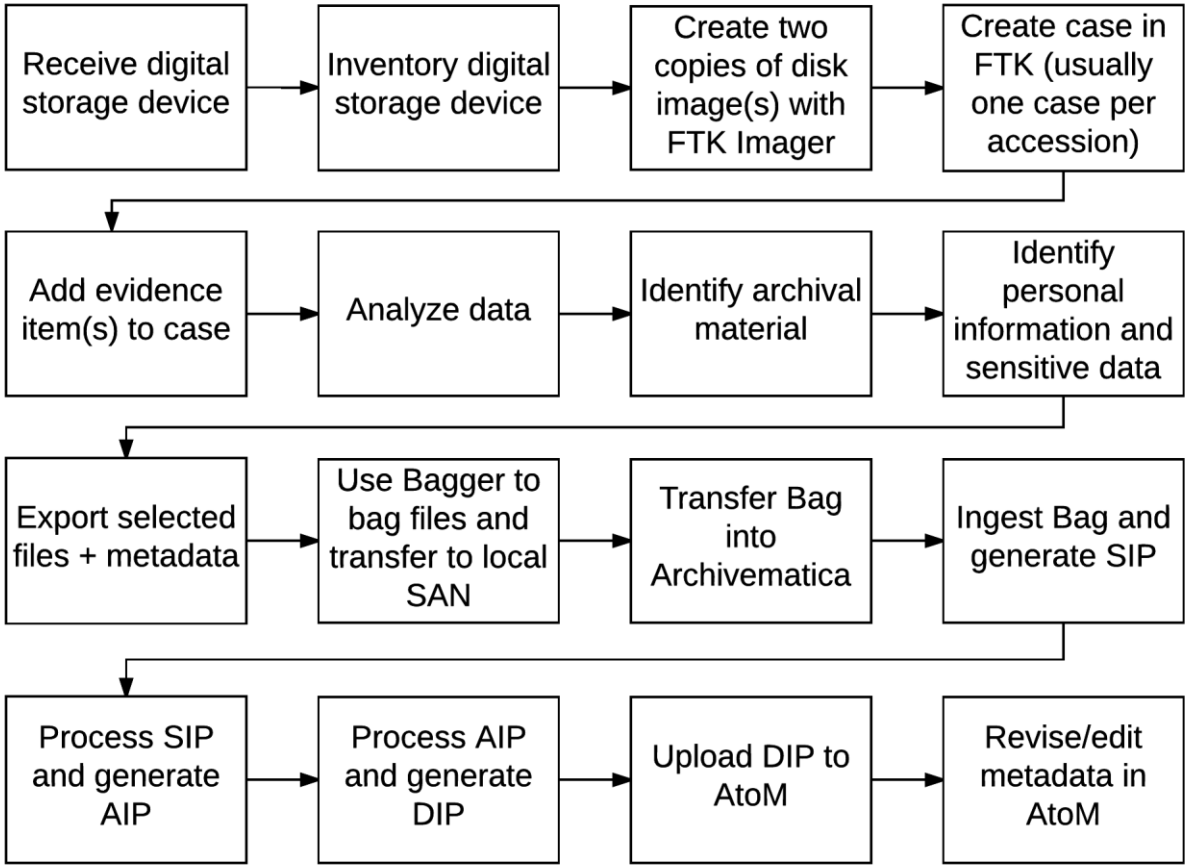


Source: Infosec Institute, Digital Forensic Models (January 25, 2016): http://resources.infosecinstitute.com/digital-forensics-models/

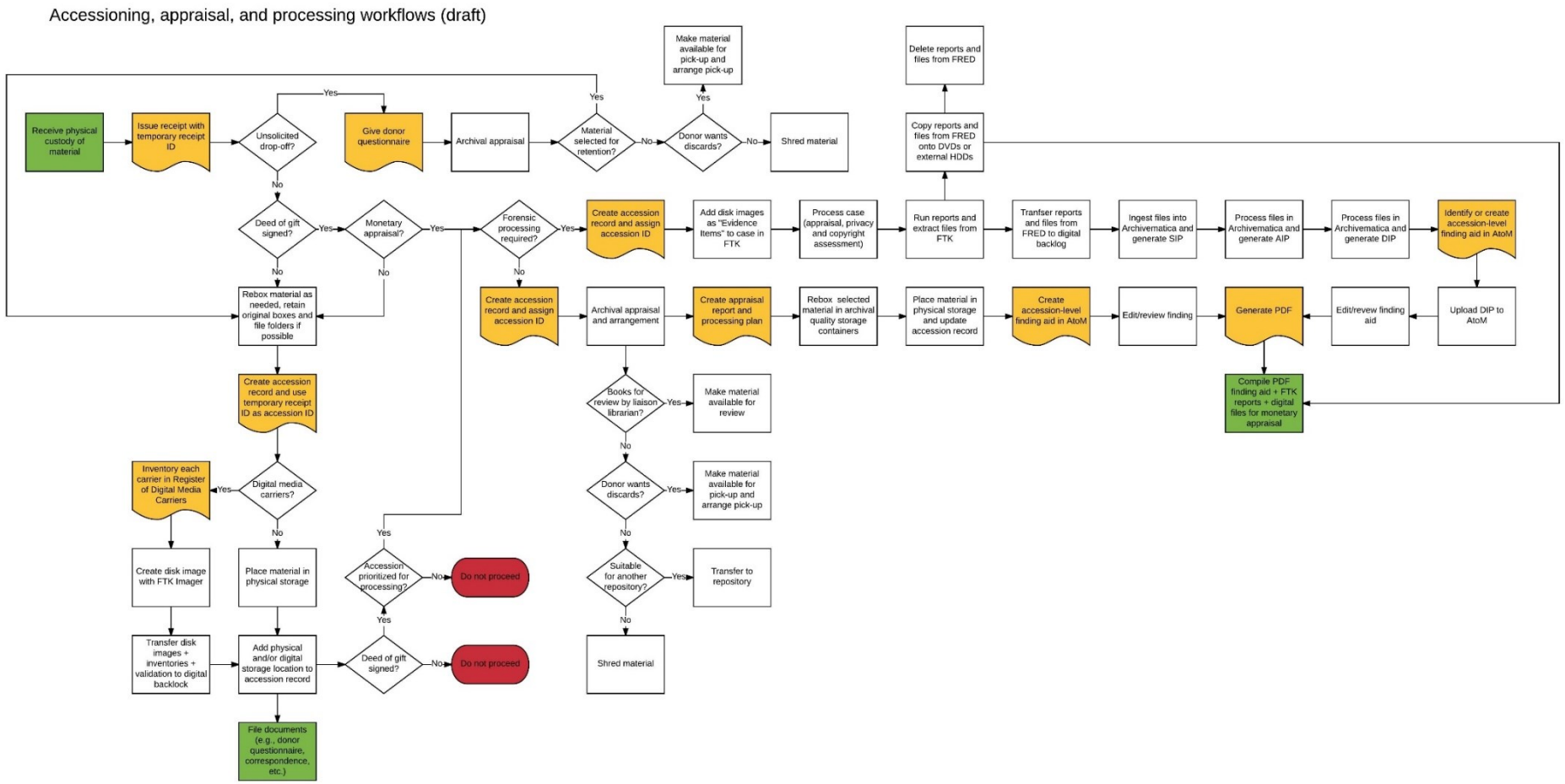
Penn State University workflow



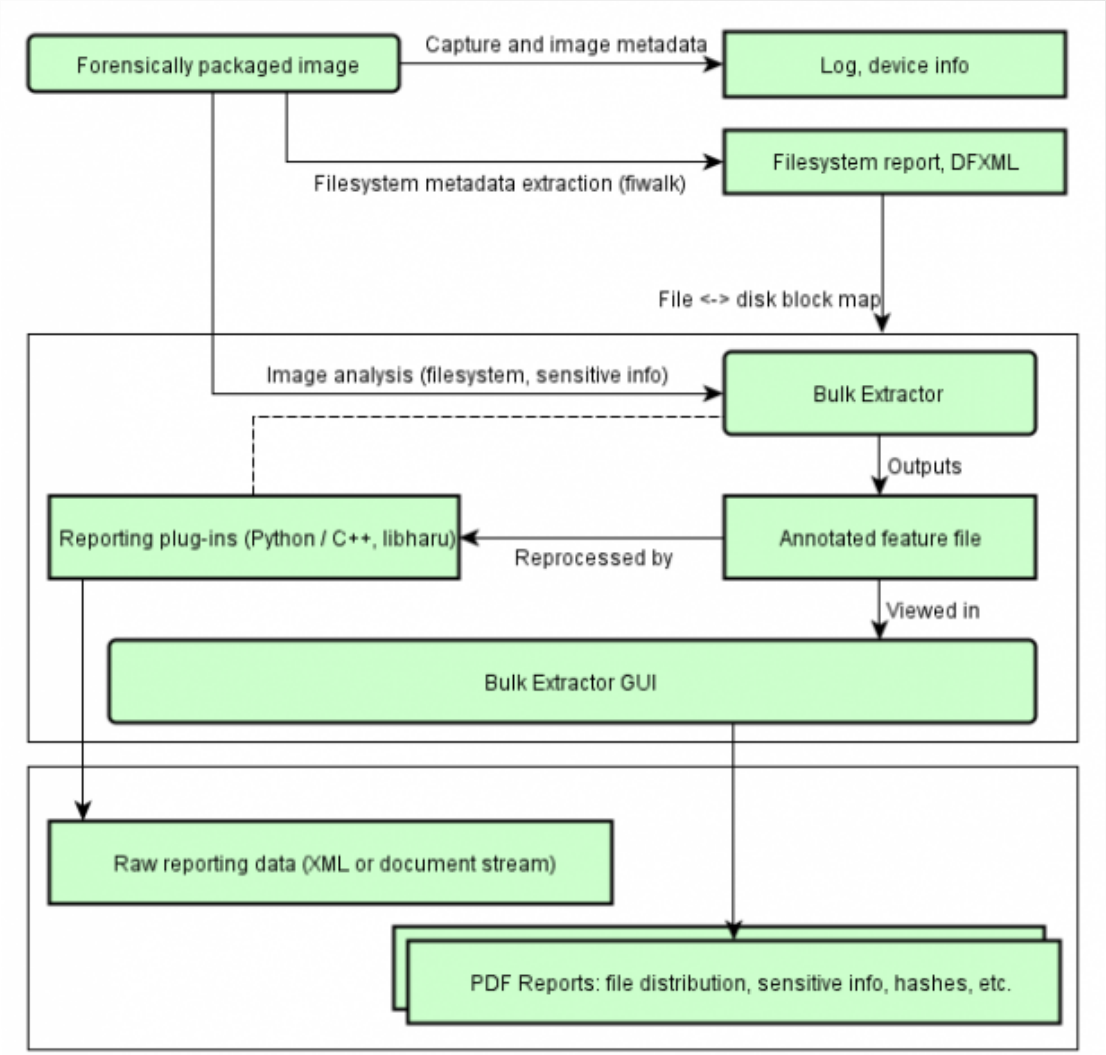
Dalhousie University workflow (draft)



Accessioning, appraisal, and processing workflow (draft)



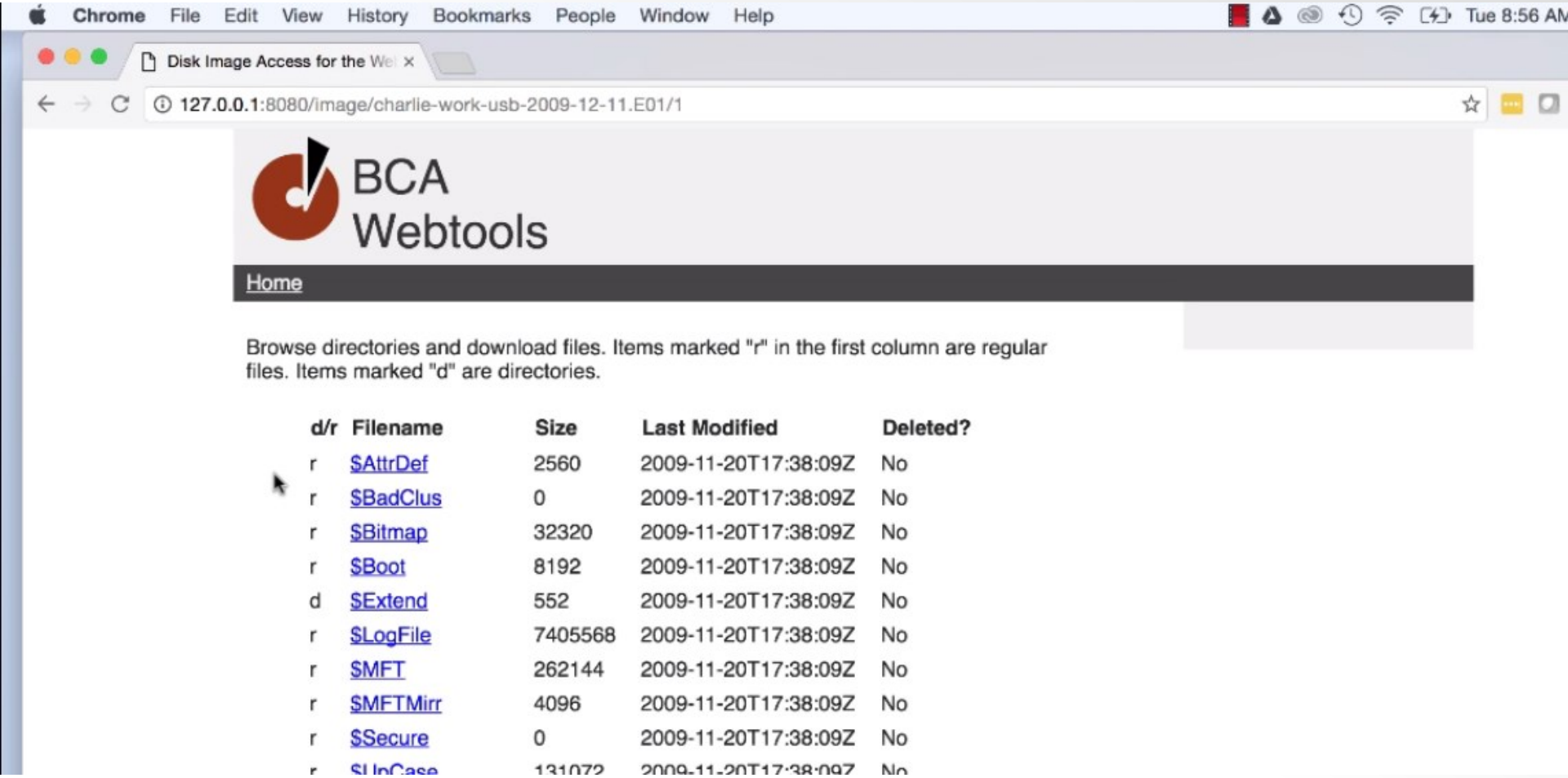
Identification of potentially sensitive information (BitCurator)



Bulk Extractor scanners – examples

Scanner name	Description
Wordlist	A list of all “words” extracted from a disk. Useful for password cracking or searching for specific terms.
Accounts	Credit card numbers, “track two” information, phone numbers, other formatted numbers. Useful for tracking how a device was used for business purposes.
Email	Discovers RFC822 email headers, HTTP cookies, hostnames, IP addresses, email addresses, and URLs. Useful for recreating email correspondence on a device.
Exif	Finds EXIF metadata in image and sound files.
Find	Returns the results of specific regular expressions.
vCard	Recovers vCards (standard electronic business cards).

BitCurator Access



11S45J9439ZVJ4YV036DCR REV:0B
lenovo
intel
CHINA
30V Serial ATA
ASAP


Forensic Computers
forensic-computers.com
Forensic Computers
T35689iu Forensic Bridge
Powered by TABLEAU
Pwr
Dev
Host
WrtBlk
Act
SAS
FireWire
USB 3.0
IDE
Insert SUBJECT Drive With Pins Outward
Connect Appropriate Cabling To T35689iu
HC/SDXC
MS/DUO
Forensic Series
EX-53
READ WRITE
M-DISC
LG

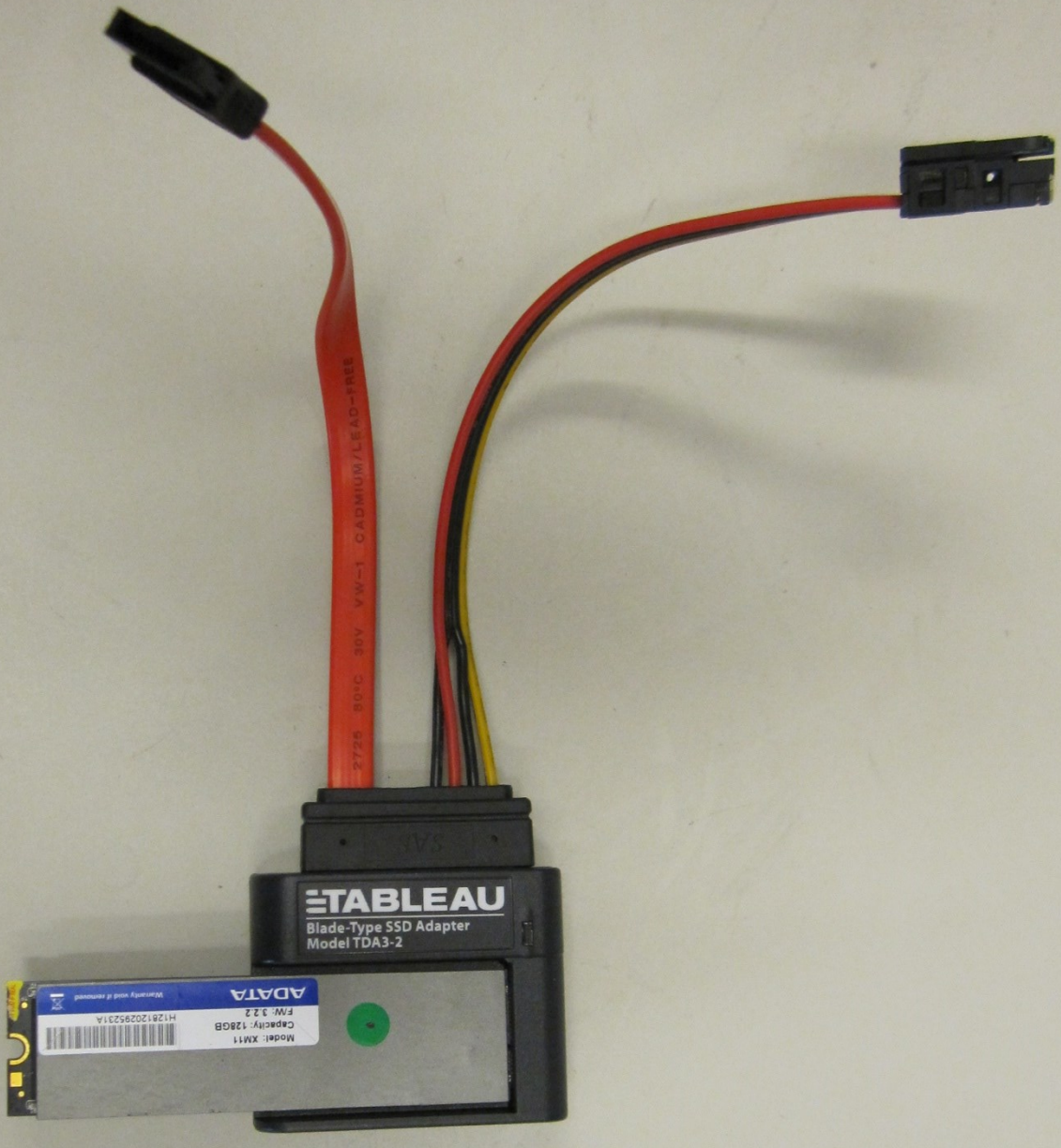
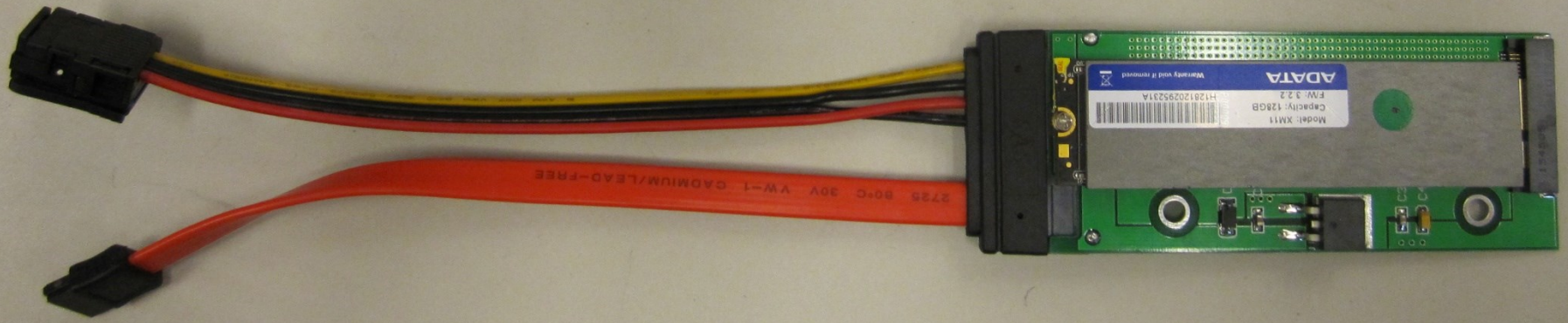
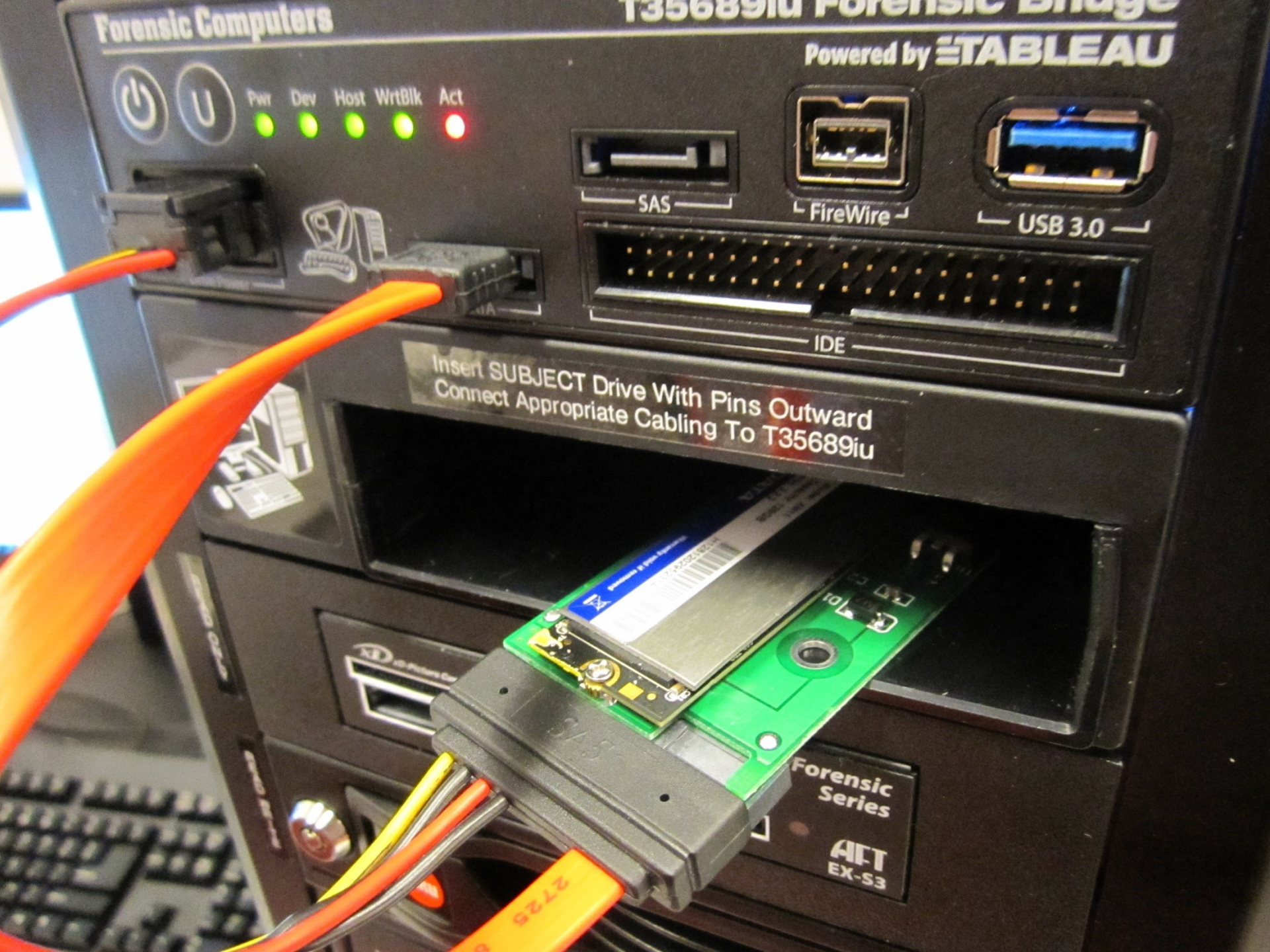
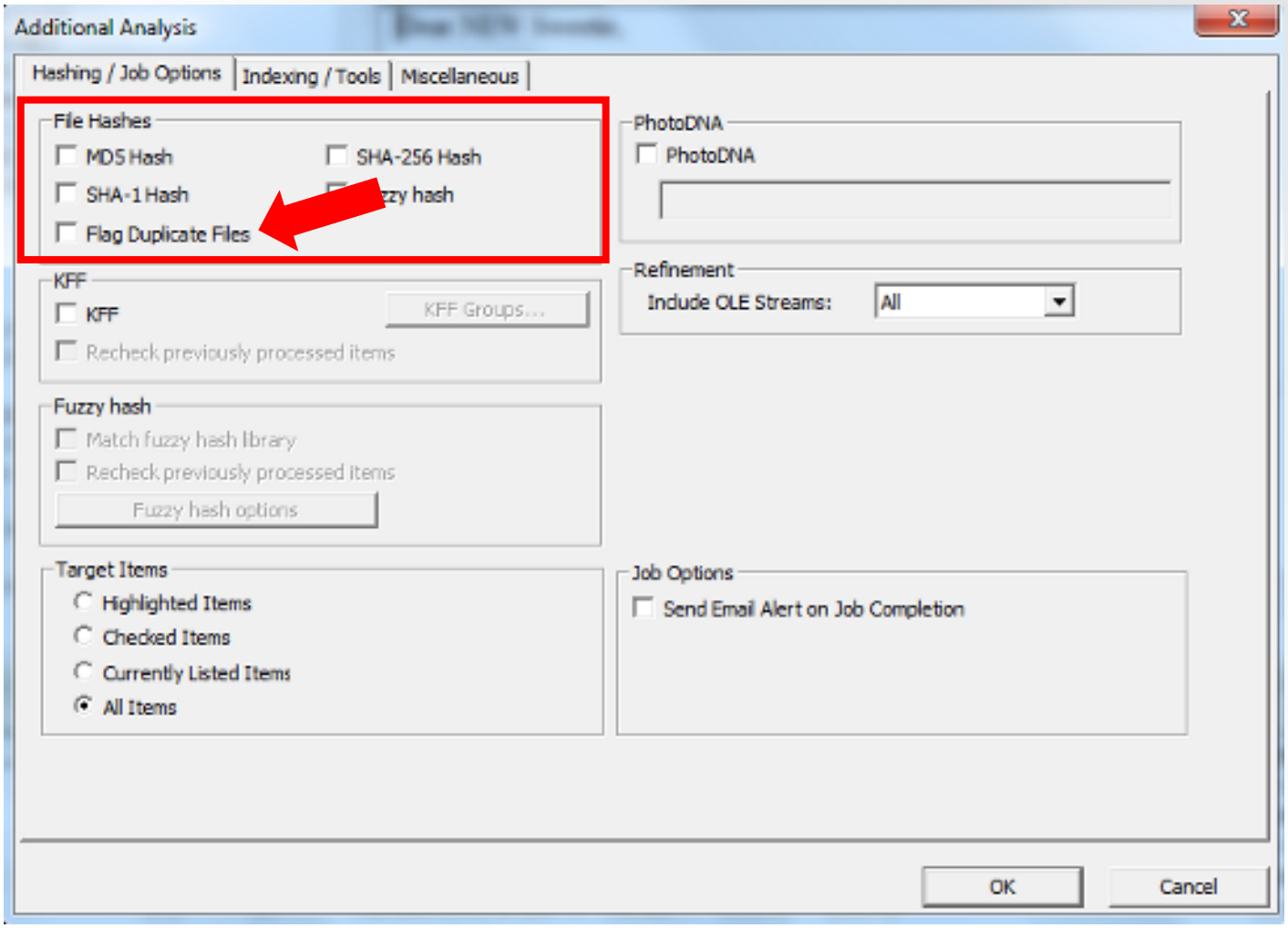


TABLEAU
Blade-Type SSD Adapter
Model TDA3-2
ADATA
Model: XM11
Capacity: 128GB
FW: 3.2.2
H12812029523 1A
Warranty void if removed
CADMIUM/LEAD-FREE
VW-1
30V
80°C
2725


ADATA
Warranty void if removed
Model: XM11
Capacity: 128GB
FW: 3.2.2
H1281202952314
2725 80°C 30V VW-1 CADMIUM/LEAD-FREE


Forensic Computers
T35689iu Forensic Bridge
Powered by TABLEAU
Pwr
Dev
Host
WrtBlk
Act
SAS
FireWire
USB 3.0
IDE
Insert SUBJECT Drive With Pins Outward
Connect Appropriate Cabling To T35689iu
Forensic Series
AFT EX-S3

FTK – Flag Duplicates



NSRL Reference Data Set (RDS)

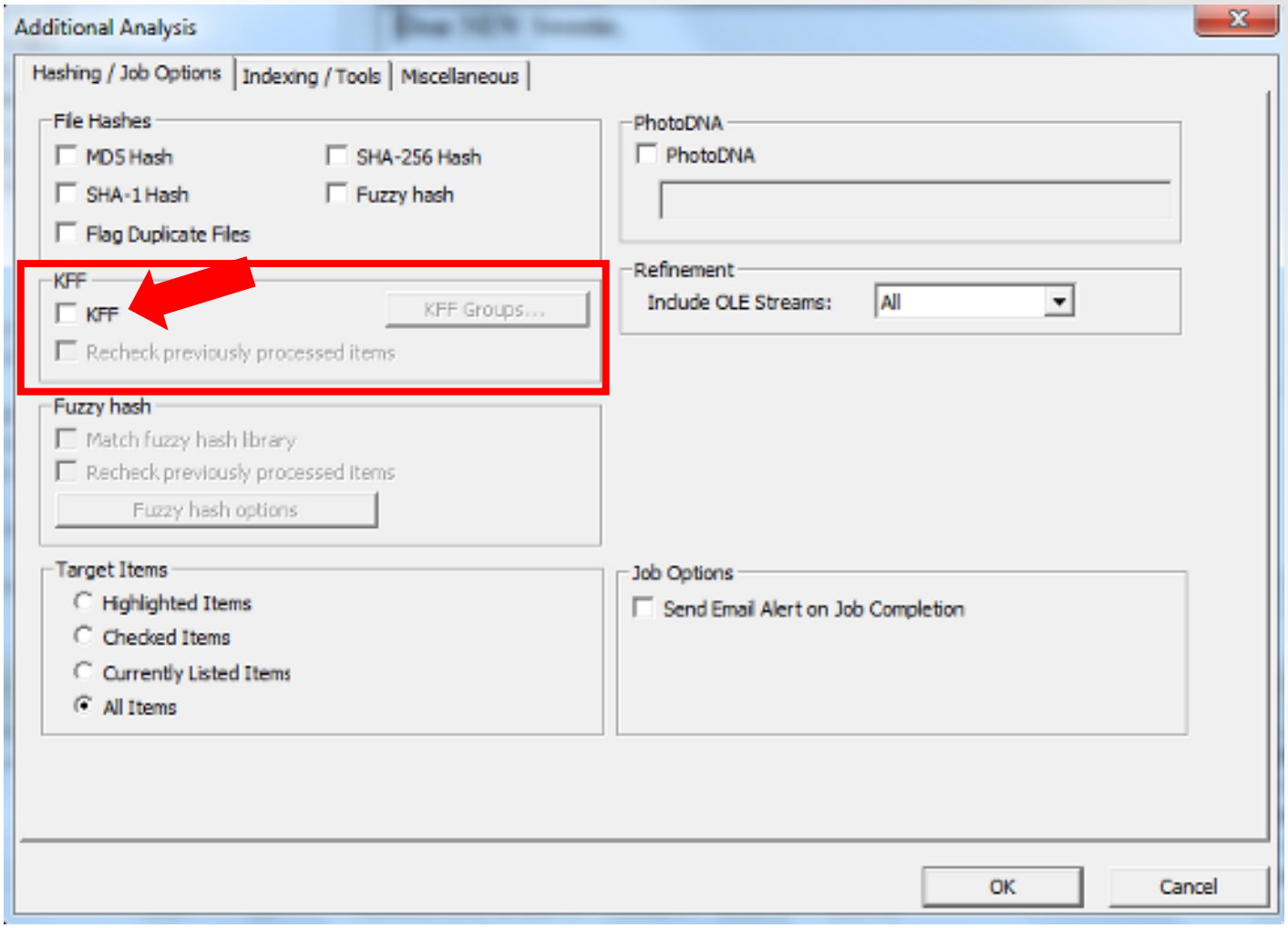
- Hashsets and metadata used in file identification
- Data can be used in third-party digital forensics tools

- RDS is updated four times each year
- As of v2.55, RDS is partitioned into four divisions:
  - Modern – applications created in or after 2000
  - Legacy – applications created in or before 1999
  - Android – Mobile apps for the Android OS
  - iOS – Mobile apps for iOS

FTK – Known File Filter (KFF)

- KFF data – hash values of known files that are compared against files in an FTK case
- KFF data can come from pre-configured libraries (e.g., NSRL RDS, DHS, ICE, etc.) or custom libraries
- FTK ships with version of NSRL RDS bifurcated into “Ignore” and “Alert” libraries
- KFF Server – used to process KFF data against evidence in an FTK case
- KFF Import Utility – used to import and index KFF data

FTK – Known File Filter (KFF)



Bill Freedman fonds filtered in FTK

Filter	Description	# of files	Size
Unfiltered	All files in case	26,651,084	3,568 GB
Primary status	Duplicate File indicator IS “Primary”	731,417	83.48 GB
Secondary status	Duplicate File indicator IS “Secondary”	16,569,218	271.5 GB
KFF Ignore	Match all files where KFF status IS “Ignore”	2,548,119	44.29 GB
No KFF Ignore	Match all files where KFF status IS NOT “Ignore” + KFF status IS “Not checked”	24,102,965	3524 GB
Primary status + No KFF Ignore	Match all files where duplicate file indicator IS “Primary” + KFF status IS NOT “Ignore”	626,351	71.95 GB
Actual files + Primary status + No KFF Ignore	Match all disk-bound files where duplicate file indicator IS “Primary” + KFF status IS NOT “Ignore”	103,412	61.81 GB

Research challenges and next steps

- Finish processing Bill Freedman fonds and preparing paper and electronic records for monetary appraisal
- Develop Privacy and Confidential Information Assessment Tool
- Finish Digital Forensics Lab manual
- Create forensic images of storage media identified during Digital Archives Collection Assessment
- Finish one thing…