# Using Open Source Digital Forensics Software for Digital Archives Workshop

Mark A. Matienzo ’04
Manuscripts and Archives, Yale University Library

Society of American Archivists
University of Michigan School of Information Chapter
October 19, 2012

# Overview

- Open source digital forensics: what, why, and how
- Technical overview on storage: media, file systems, etc.
- Introduction to tools: Sleuth Kit, fiwalk, bulk_extractor
- Hands-on walkthroughs with sample data/disk images

# We’re not covering...

- 1. Hands-on disk imaging
- 2. Processing, arrangement, description, etc. - left as an exercise to the student
- 3. How to aggregate extracted (meta)data in ways most useful to archives and libraries
- 2 and 3 are left as exercises for the student - but we can discuss! :)

# Digital Forensics

# Branches of
# Digital Forensics

- File system forensics
- Incident response
- Intrusion detection
- Mobile device forensics
- Network forensics
- Database forensics

# We know how to go from this … to this

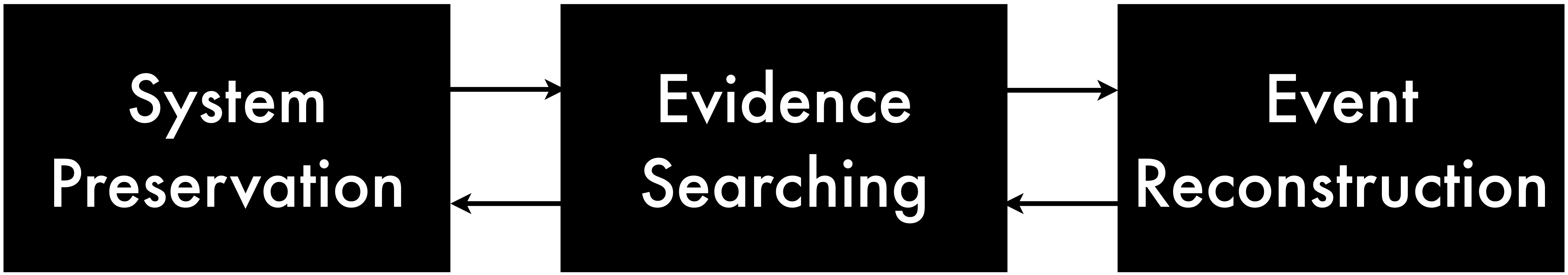
http://blogs.library.duke.edu/rubenstein/2011/09/06/all-in-a-days-processing/

# Should these be any different?

# Why Digital Forensics in Archives?

- Digital forensics is an established discipline that demands holistic capture and preservation of evidence
- Archives are faced with growing mass of digital information, with much stored on removable media
- Overlap in terms of skills and knowledge and many potential opportunities for collaboration

# Forensic Discovery Process



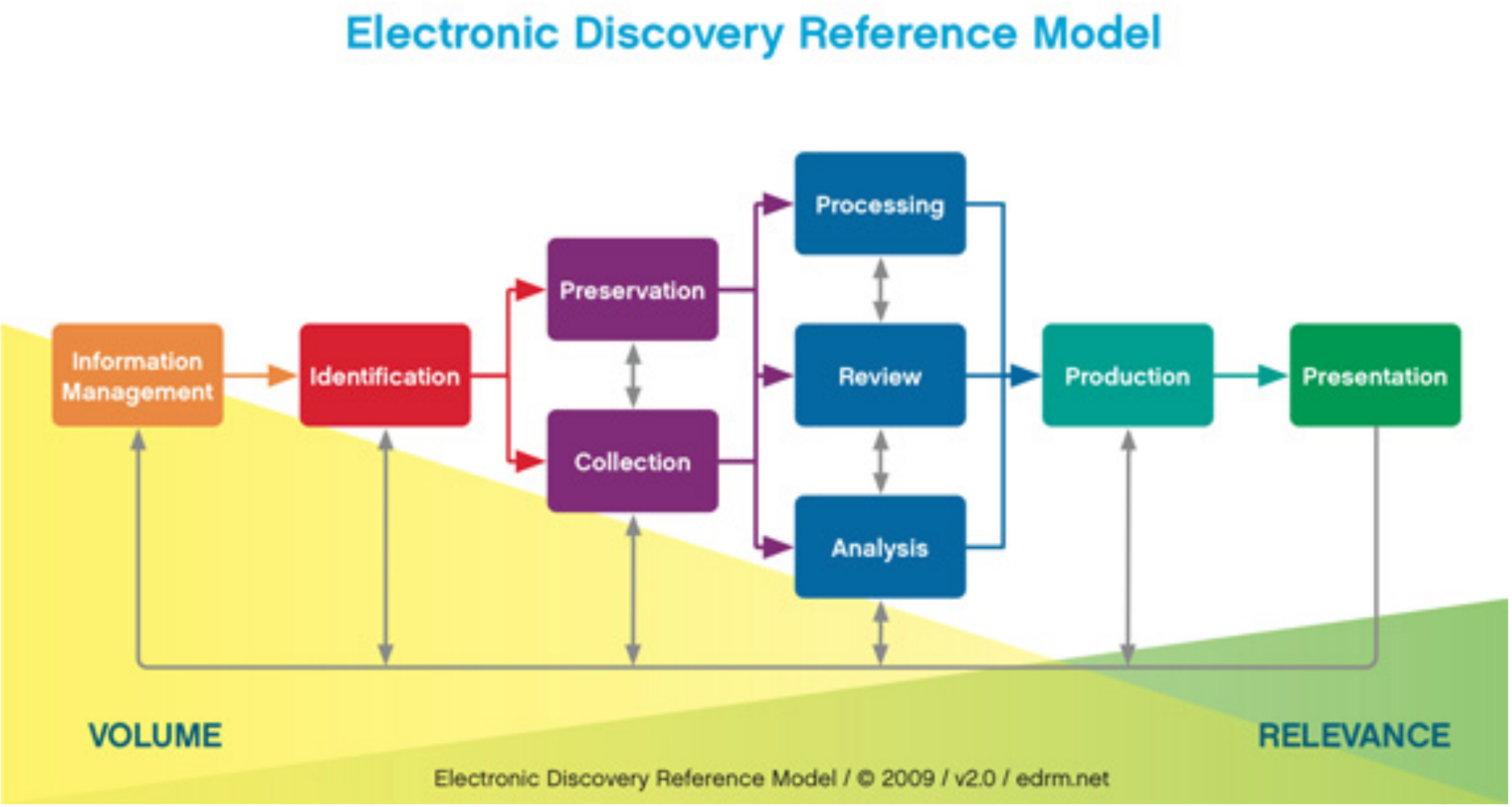
System Preservation

Evidence Searching

Event Reconstruction

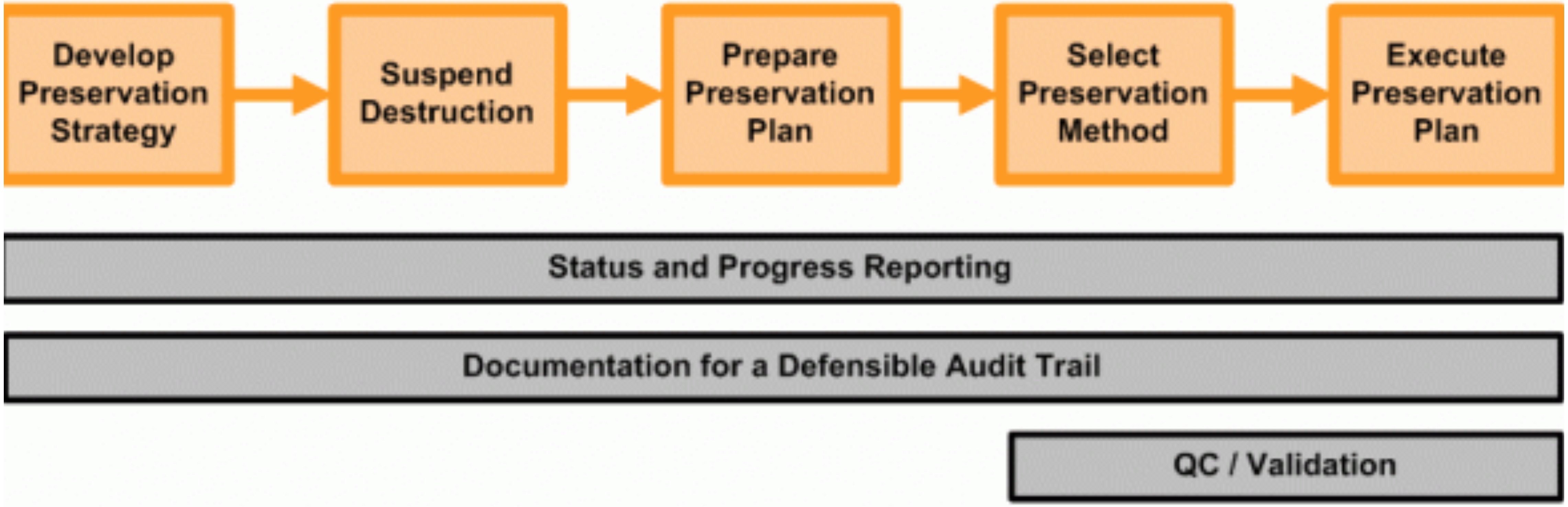
Based on Brian Carrier, *File System Forensic Analysis* (2005), p. 13

# Electronic Discovery Reference Model Stages



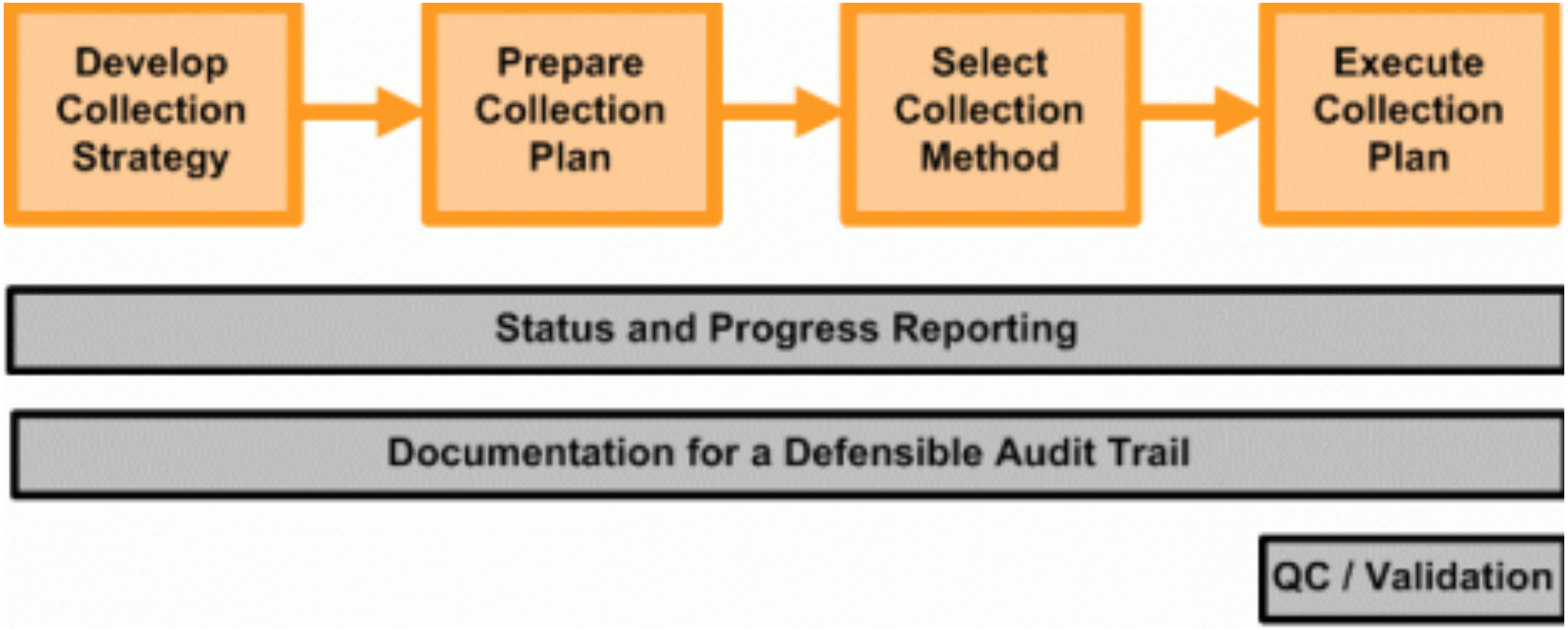
http://www.edrm.net/resources/edrm-stages-explained

# EDRM: Preservation


Develop Preservation Strategy
Suspend Destruction
Prepare Preservation Plan
Select Preservation Method
Execute Preservation Plan
Status and Progress Reporting
Documentation for a Defensible Audit Trail
QC / Validation

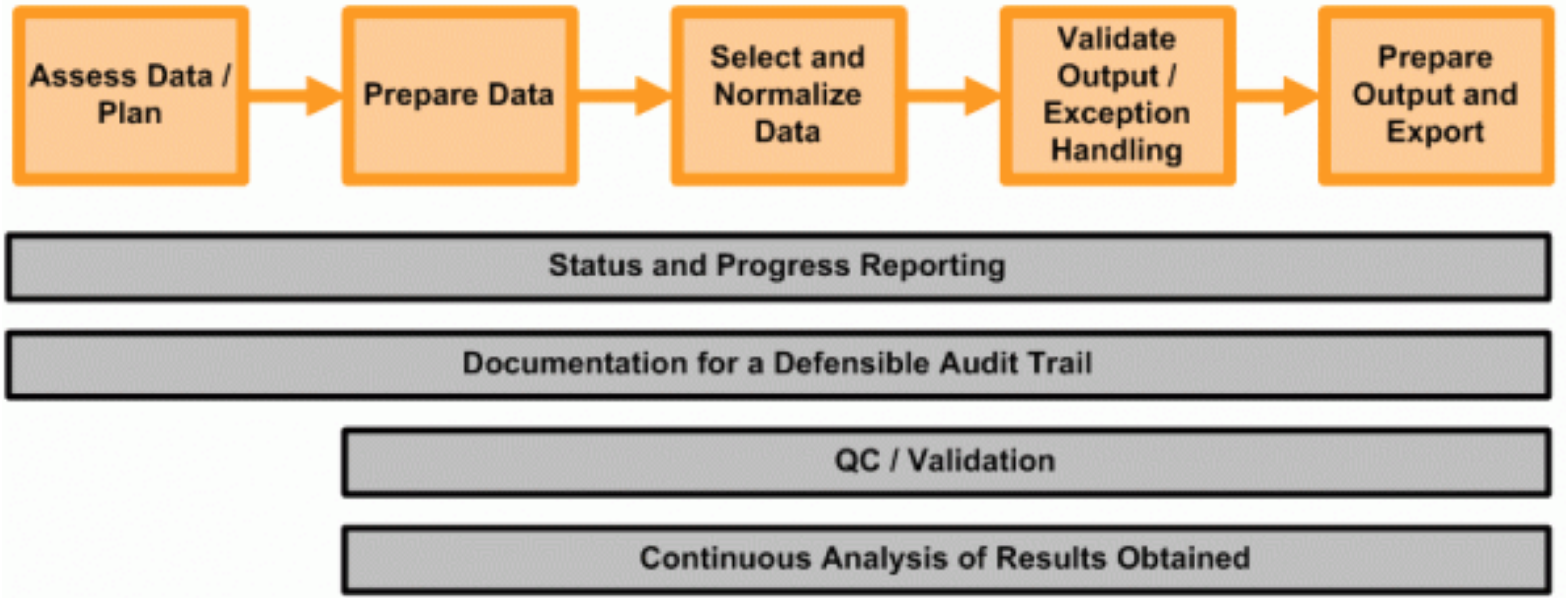
http://www.edrm.net/resources/guides/edrm-framework-guides/preservation

# EDRM: Collection


Develop Collection Strategy
Prepare Collection Plan
Select Collection Method
Execute Collection Plan
Status and Progress Reporting
Documentation for a Defensible Audit Trail
QC / Validation

http://www.edrm.net/resources/guides/edrm-framework-guides/collection

# EDRM: Processing



Assess Data / Plan
Prepare Data
Select and Normalize Data
Validate Output / Exception Handling
Prepare Output and Export
Status and Progress Reporting
Documentation for a Defensible Audit Trail
QC / Validation
Continuous Analysis of Results Obtained

http://www.edrm.net/resources/guides/edrm-framework-guides/processing

# Useful Aspects of Digital Forensics for Archives

- Obtain provenance information about context of creation, and record provenance information about processes of transfer and analysis
- Document original order: relationship of files in directory hierarchy, related applications, associated accounts
- Document and ensure chain of custody through proven transfer methods that maintain integrity and authenticity
- Identification of sensitive information

# Combining Workflows (1)

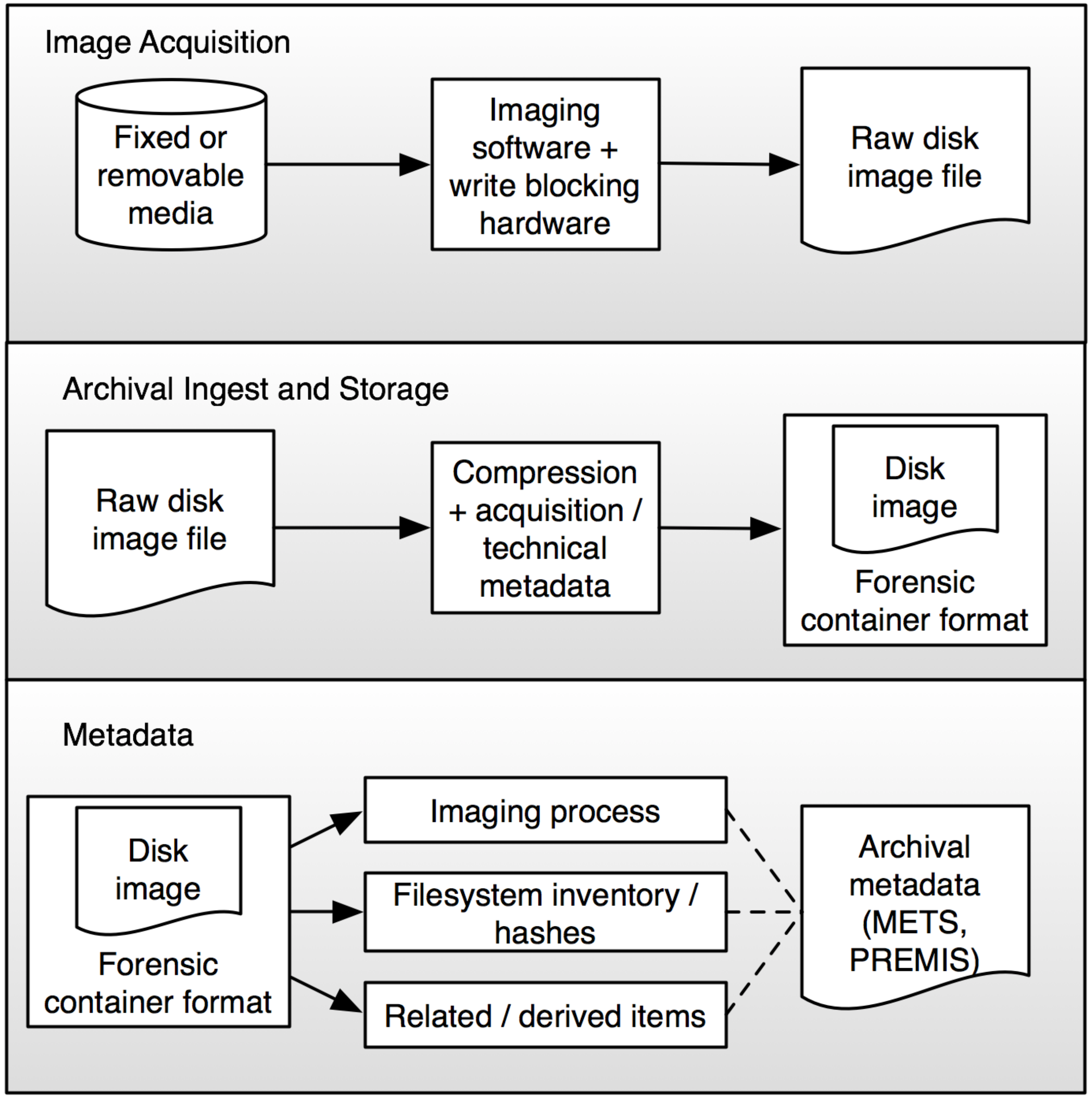
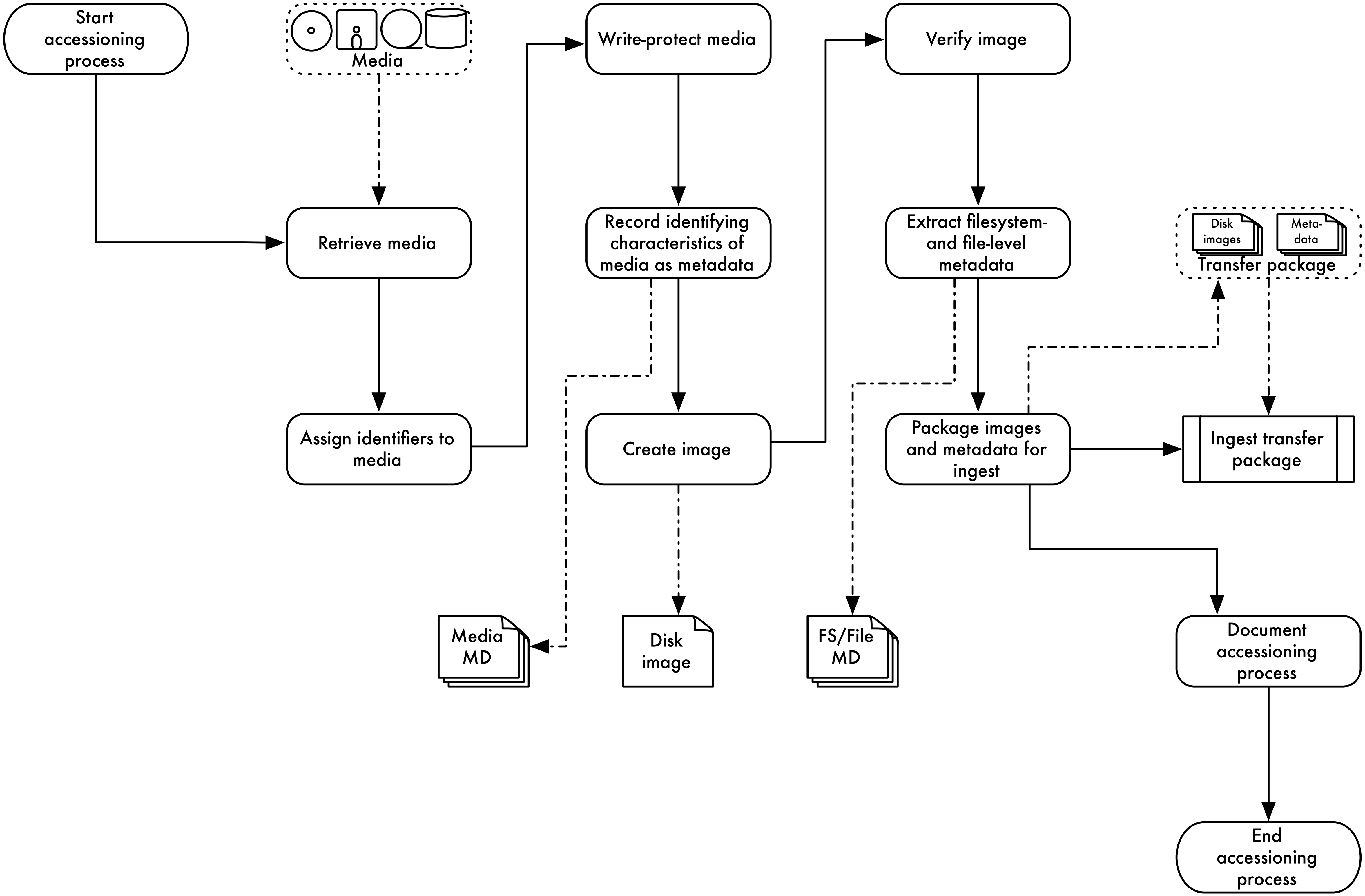


Image Acquisition
Fixed or removable media
Imaging software + write blocking hardware
Raw disk image file
Archival Ingest and Storage
Raw disk image file
Compression + acquisition / technical metadata
Disk image
Forensic container format
Metadata
Disk image
Forensic container format
Imaging process
Filesystem inventory / hashes
Related / derived items
Archival metadata (METS, PREMIS)

Woods, Lee, and Garfinkel (2011)

# Combining Workflows (2)



Start accessioning process
Media
Retrieve media
Assign identifiers to media
Write-protect media
Record identifying characteristics of media as metadata
Create image
Media MD
Disk image
Verify image
Extract filesystem- and file-level metadata
Package images and metadata for ingest
FS/File MD
Disk images
Meta-data
Transfer package
Ingest transfer package
Document accessioning process
End accessioning process

# Transfer Goals

- Obtain records/files/assets in a manner that does not threaten their integrity and authenticity
- Understand correspondence or gaps between capabilities and identified requirements

# Ensuring Integrity and Authenticity

- Use means to prevent accidental alteration of assets as received, using write protection mechanisms
- Document process, especially when you take extraordinary steps

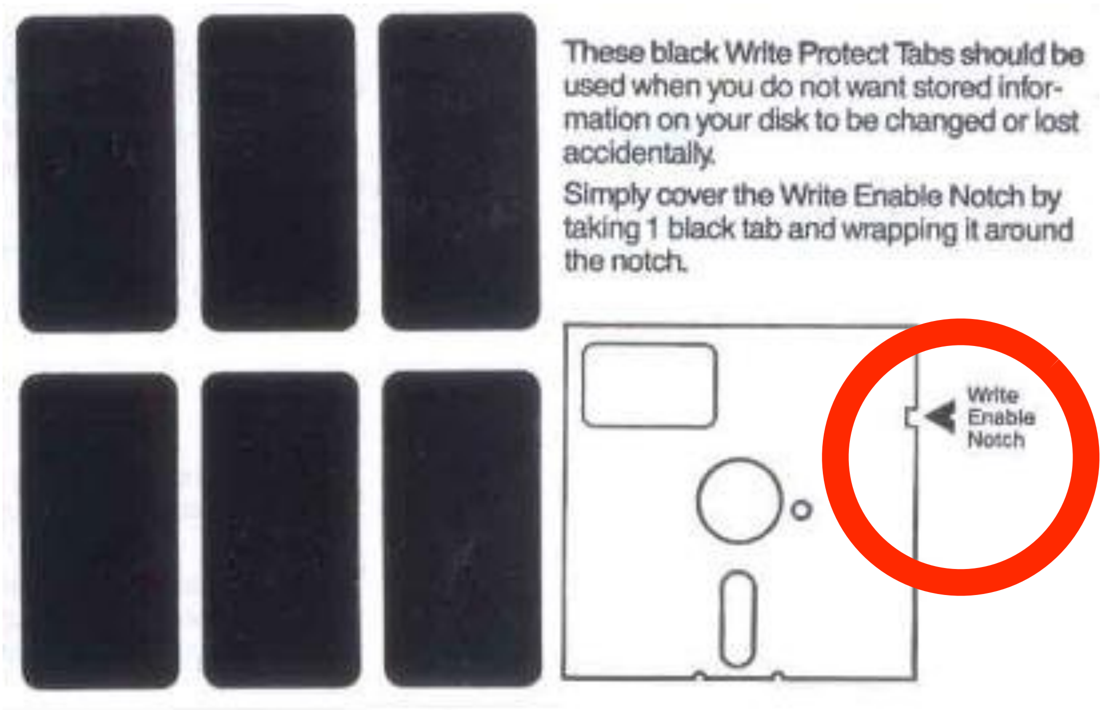
# Transfer Options

- Disk imaging the entirety of a piece of media including deleted files, errors, etc.
- Logical imaging: Selecting files directly and transferring them off
- Need to ensure that files do not get altered regardless of process

# Preventing Accidental Modification

- Write protection: some media formats have physical means (floppies) or limitations (CD-ROMs)
- Write blocking: using hardware or software mechanism to prevent write signals from being processed by computer or drive

# Write Protection

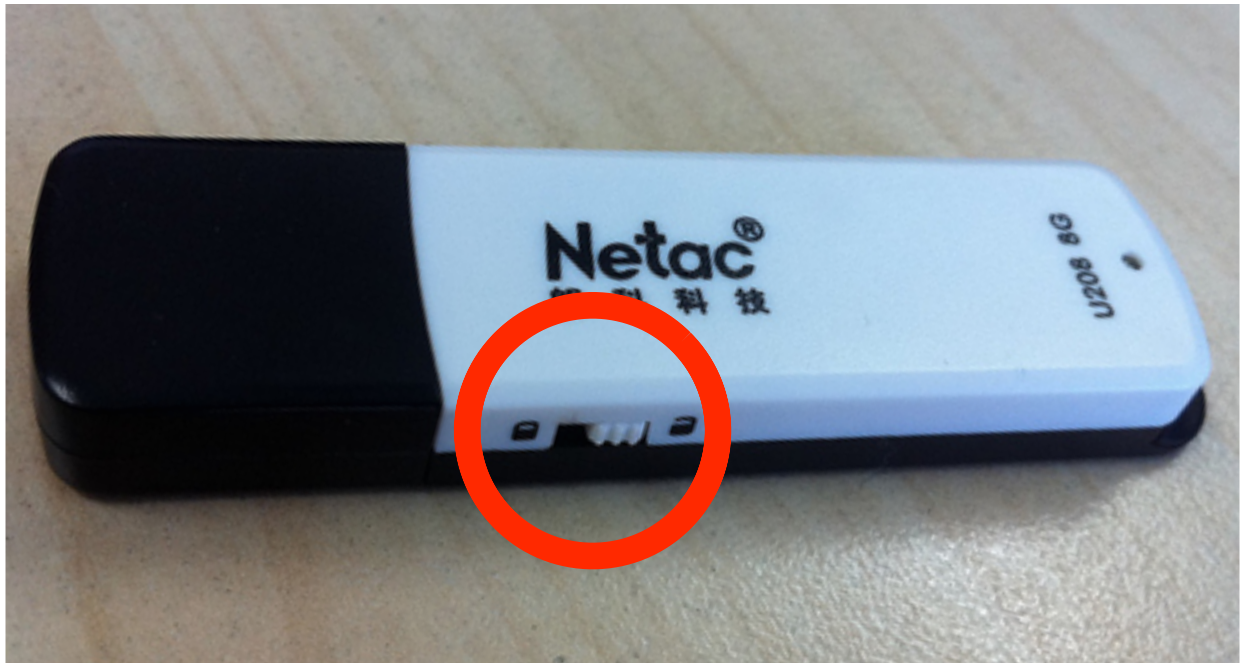


These black Write Protect Tabs should be used when you do not want stored information on your disk to be changed or lost accidentally.

Simply cover the Write Enable Notch by taking 1 black tab and wrapping it around the notch.

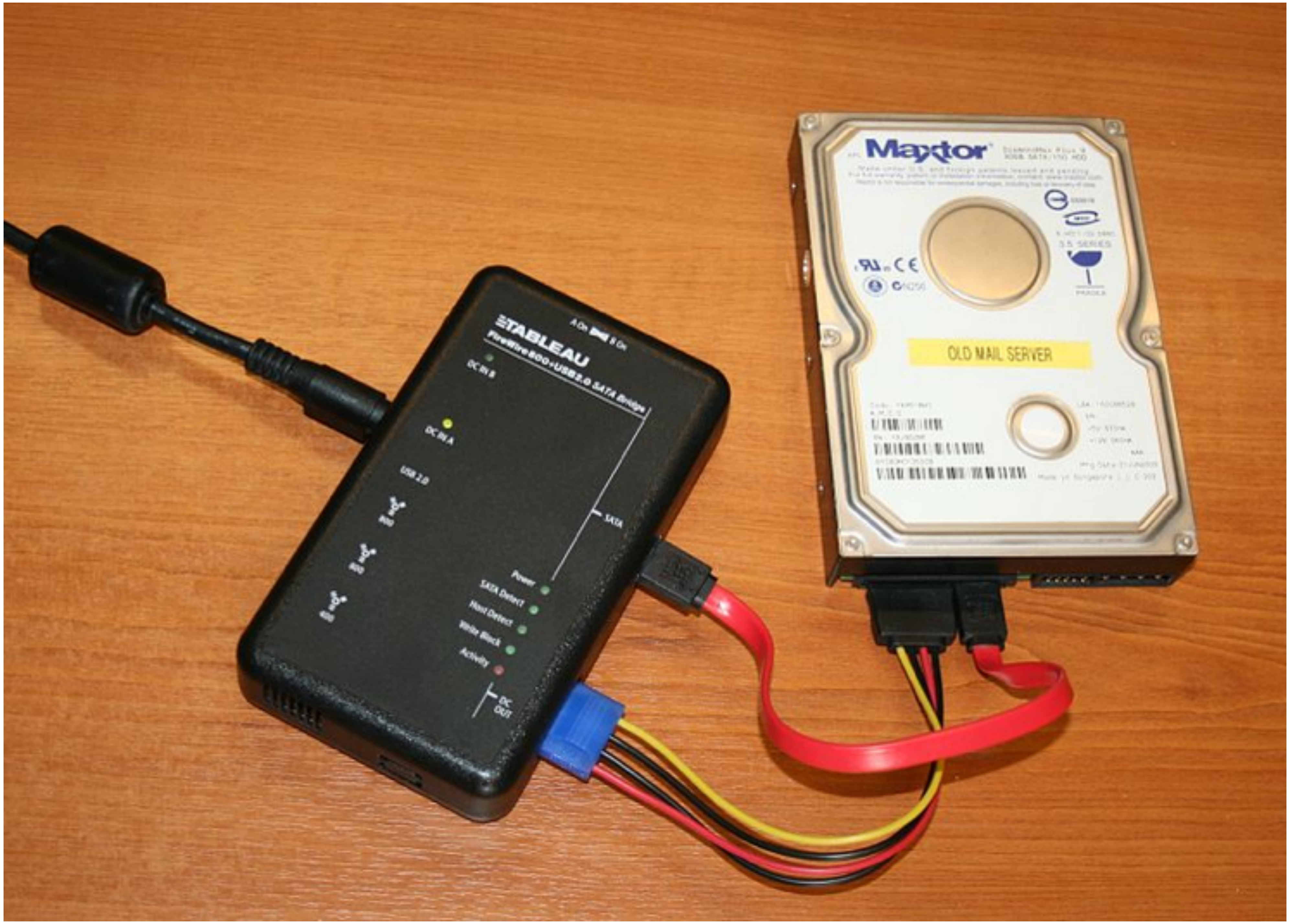
Write
Enable
Notch

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



http://www.flickr.com/photos/bfishadow/5533694844

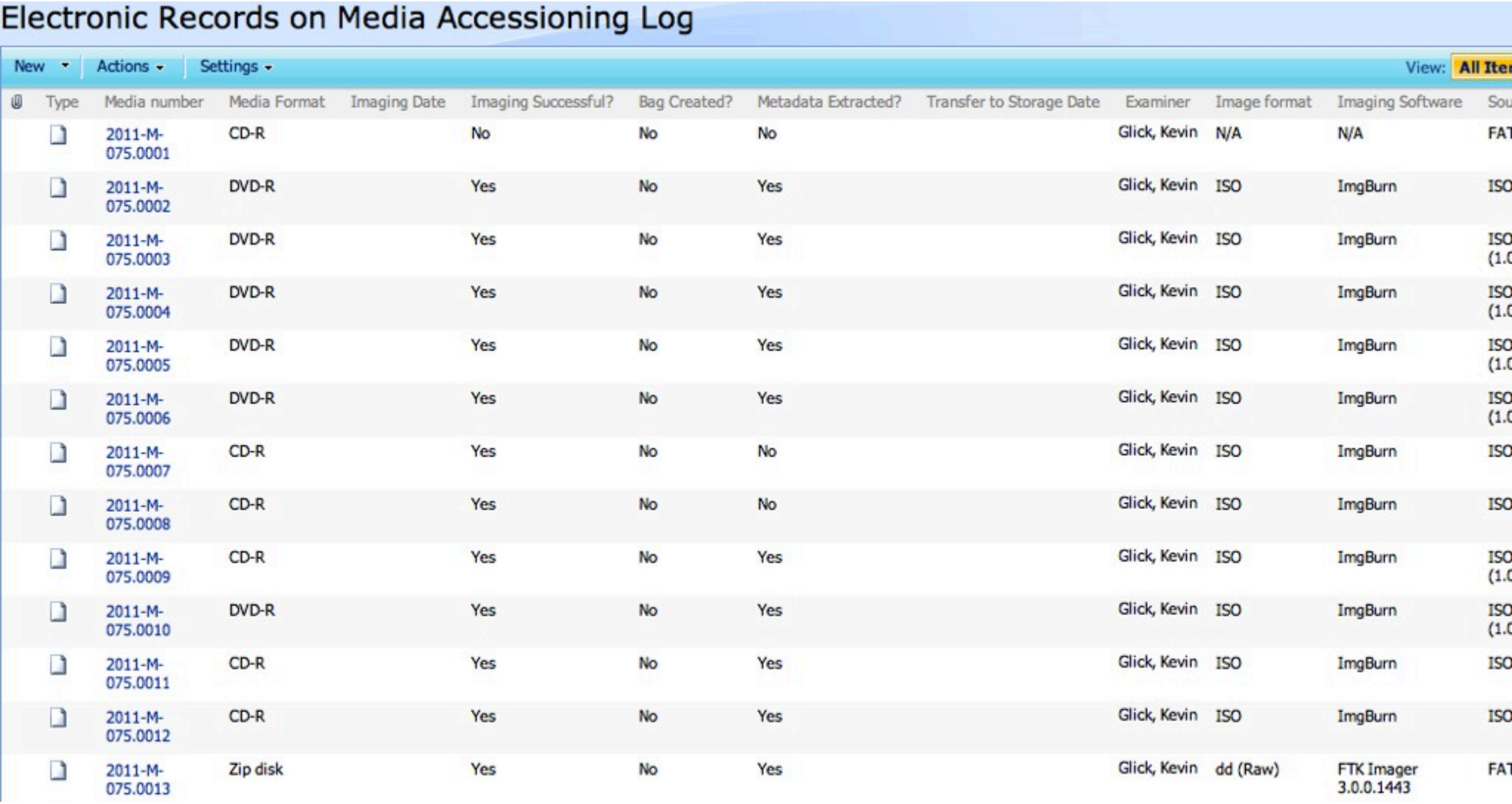
# Hardware Write Blocking



http://en.wikipedia.org/wiki/File:Portable_forensic_tableau.JPG

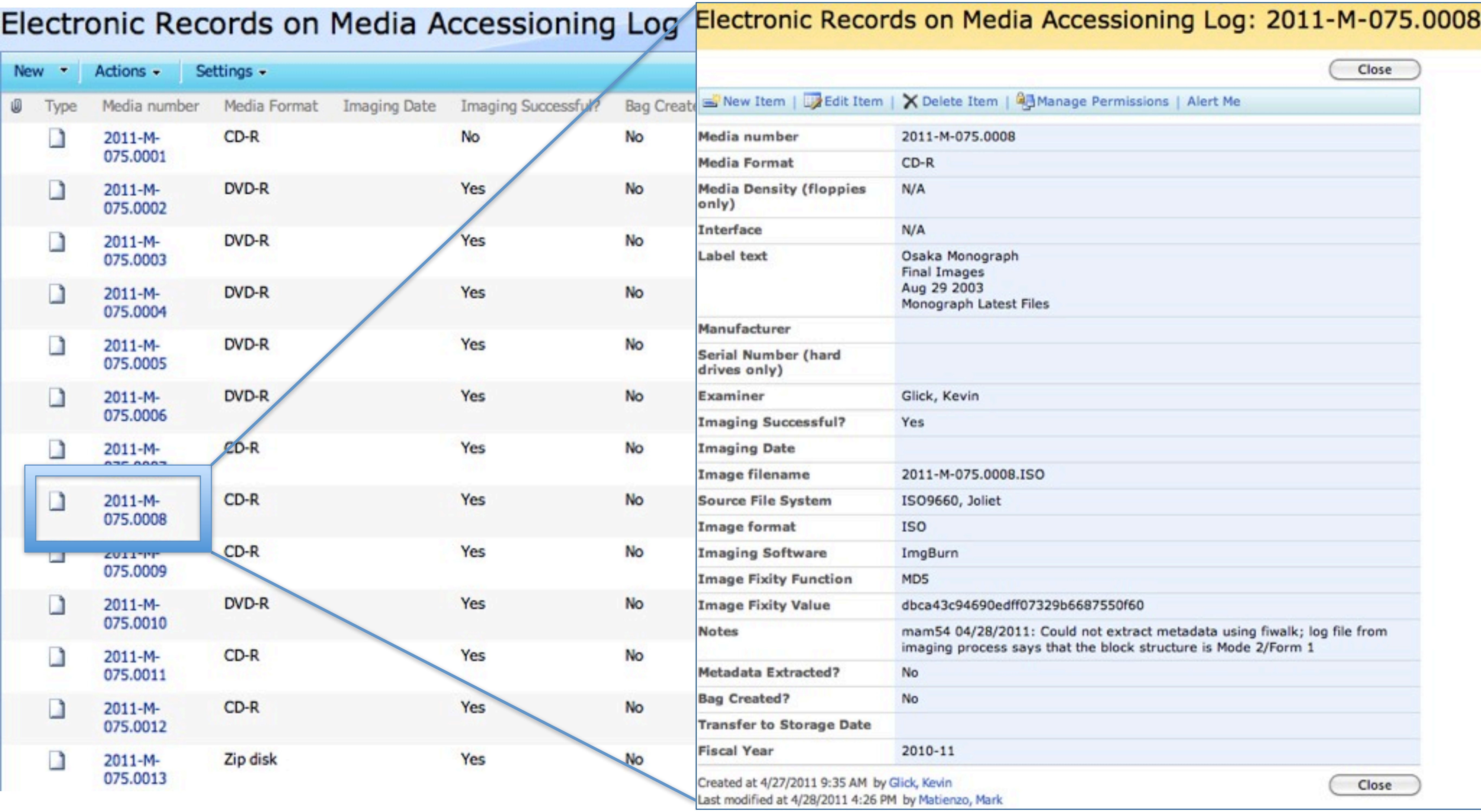
# Documentation Goals

- Identify and record characteristics of media
- Document transfer process
- Gather information about assets (descriptive metadata, technical metadata, preservation metadata...)



# Electronic Records on Media Accessioning Log

| 📎 | Type | Media number | Media Format | Imaging Date | Imaging Successful? | Bag Created? | Metadata Extracted? | Transfer to Storage Date | Examiner | Image format | Imaging Software | Sou |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 🗋 | 2011-M-075.0001 | CD-R | | No | No | No | | Glick, Kevin | N/A | N/A | FAT |
| | 🗋 | 2011-M-075.0002 | DVD-R | | Yes | No | Yes | | Glick, Kevin | ISO | ImgBurn | ISO |
| | 🗋 | 2011-M-075.0003 | DVD-R | | Yes | No | Yes | | Glick, Kevin | ISO | ImgBurn | ISO (1.0 |
| | 🗋 | 2011-M-075.0004 | DVD-R | | Yes | No | Yes | | Glick, Kevin | ISO | ImgBurn | ISO (1.0 |
| | 🗋 | 2011-M-075.0005 | DVD-R | | Yes | No | Yes | | Glick, Kevin | ISO | ImgBurn | ISO (1.0 |
| | 🗋 | 2011-M-075.0006 | DVD-R | | Yes | No | Yes | | Glick, Kevin | ISO | ImgBurn | ISO (1.0 |
| | 🗋 | 2011-M-075.0007 | CD-R | | Yes | No | No | | Glick, Kevin | ISO | ImgBurn | ISO |
| | 🗋 | 2011-M-075.0008 | CD-R | | Yes | No | No | | Glick, Kevin | ISO | ImgBurn | ISO |
| | 🗋 | 2011-M-075.0009 | CD-R | | Yes | No | Yes | | Glick, Kevin | ISO | ImgBurn | ISO (1.0 |
| | 🗋 | 2011-M-075.0010 | DVD-R | | Yes | No | Yes | | Glick, Kevin | ISO | ImgBurn | ISO (1.0 |
| | 🗋 | 2011-M-075.0011 | CD-R | | Yes | No | Yes | | Glick, Kevin | ISO | ImgBurn | ISO |
| | 🗋 | 2011-M-075.0012 | CD-R | | Yes | No | Yes | | Glick, Kevin | ISO | ImgBurn | ISO |
| | 🗋 | 2011-M-075.0013 | Zip disk | | Yes | No | Yes | | Glick, Kevin | dd (Raw) | FTK Imager 3.0.0.1443 | FAT |



# Electronic Records on Media Accessioning Log

| | Type | Media number | Media Format | Imaging Date | Imaging Successful? | Bag Create |
|---|---|---|---|---|---|---|
| | | 2011-M-075.0001 | CD-R | | No | No |
| | | 2011-M-075.0002 | DVD-R | | Yes | No |
| | | 2011-M-075.0003 | DVD-R | | Yes | No |
| | | 2011-M-075.0004 | DVD-R | | Yes | No |
| | | 2011-M-075.0005 | DVD-R | | Yes | No |
| | | 2011-M-075.0006 | DVD-R | | Yes | No |
| | | 2011-M-075.0007 | CD-R | | Yes | No |
| | | 2011-M-075.0008 | CD-R | | Yes | No |
| | | 2011-M-075.0009 | CD-R | | Yes | No |
| | | 2011-M-075.0010 | DVD-R | | Yes | No |
| | | 2011-M-075.0011 | CD-R | | Yes | No |
| | | 2011-M-075.0012 | CD-R | | Yes | No |
| | | 2011-M-075.0013 | Zip disk | | Yes | No |

# Electronic Records on Media Accessioning Log: 2011-M-075.0008

Close
New Item | Edit Item | Delete Item | Manage Permissions | Alert Me

| | |
|---|---|
| Media number | 2011-M-075.0008 |
| Media Format | CD-R |
| Media Density (floppies only) | N/A |
| Interface | N/A |
| Label text | Osaka Monograph<br>Final Images<br>Aug 29 2003<br>Monograph Latest Files |
| Manufacturer | |
| Serial Number (hard drives only) | |
| Examiner | Glick, Kevin |
| Imaging Successful? | Yes |
| Imaging Date | |
| Image filename | 2011-M-075.0008.ISO |
| Source File System | ISO9660, Joliet |
| Image format | ISO |
| Imaging Software | ImgBurn |
| Image Fixity Function | MD5 |
| Image Fixity Value | dbca43c94690edff07329b6687550f60 |
| Notes | mam54 04/28/2011: Could not extract metadata using fiwalk; log file from imaging process says that the block structure is Mode 2/Form 1 |
| Metadata Extracted? | No |
| Bag Created? | No |
| Transfer to Storage Date | |
| Fiscal Year | 2010-11 |

Created at 4/27/2011 9:35 AM by Glick, Kevin
Last modified at 4/28/2011 4:26 PM by Matienzo, Mark

# Extraction & Analysis Goals

- Desire to obtain metadata that can be repurposed:
  - Provide an inventory (listing of files, with modification dates and extents)
  - Provide more detailed technical information (file format, software used, etc.)
  - Provide context (creator information, etc.)
- Repurposing may mean translation into standards used by archives and libraries
- Extract and possibly migrate files of interest

# Why Open Source Digital Forensics?

- Cultural heritage sector is an emerging market for vendors and comparatively small to their primary market
- Allows for better collaboration and less dependence on specific individuals or companies
- Transparency of design and implementation allows for better understanding of impact on authenticity
- Potential to shape future of software

# Understanding Storage and Forensic Analysis

# Nature of Digital Objects

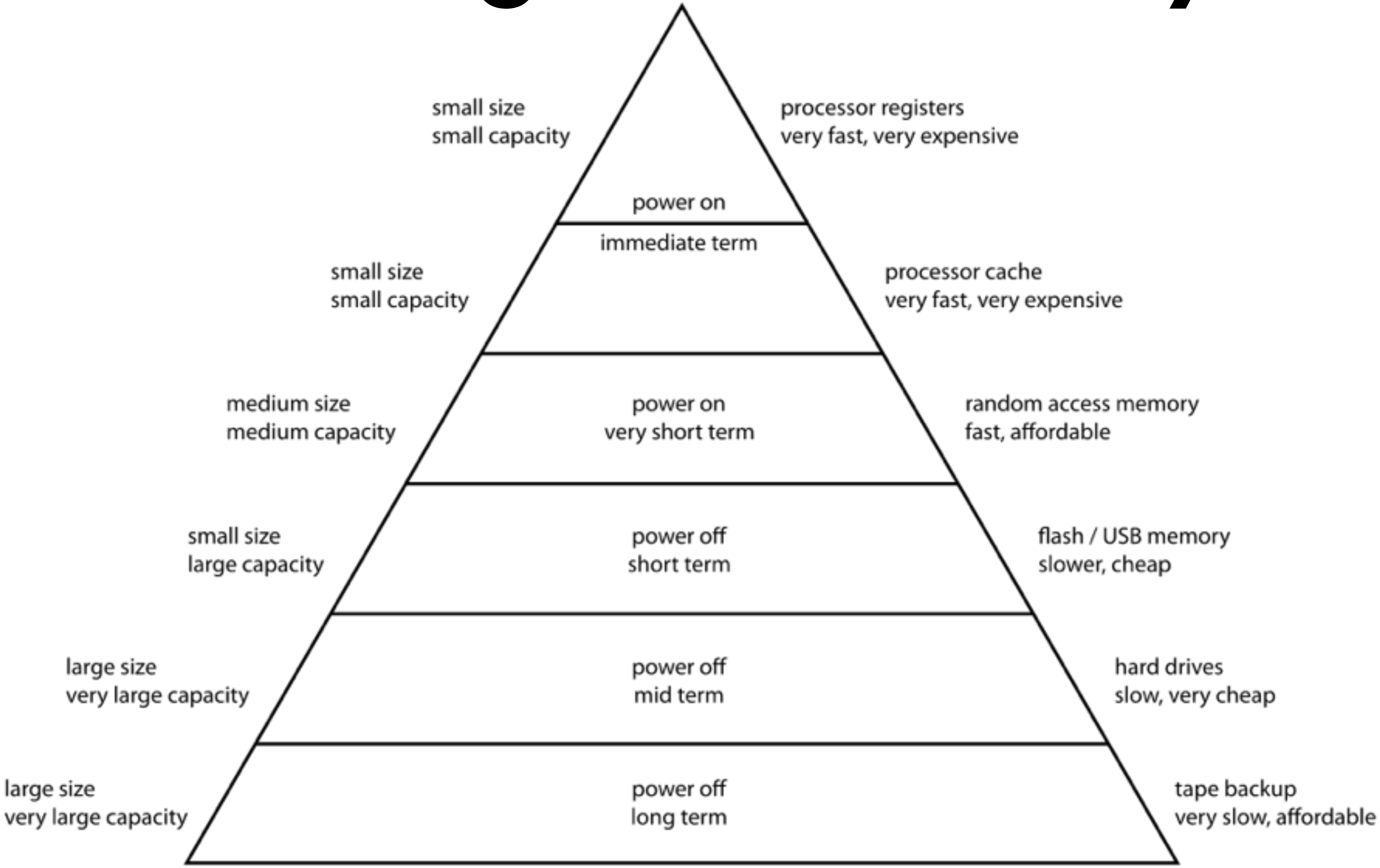
- Digital objects require mediation and depend on a complex interplay of technological systems and entities
- Like any technology, digital objects depend on layers of abstraction, e.g. OSI Model for computer networking

# Levels of Representation

| Level | Label | Description |
|---|---|---|
| 8 | Aggregation of Objects | Set of objects that form an aggregation that is meaningful when 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 equipment | Series of 1s and 0s as accessed from the storage media decoded using input/output hardware and software |
| 1 | Raw signal stream through I/O equipment | Stream of analog electronic output read from the drive without yet interpreting the signal stream as a set of discrete values |
| 0 | Signal on physical medium | Physical properties of the storage medium used to encode an analog signal |

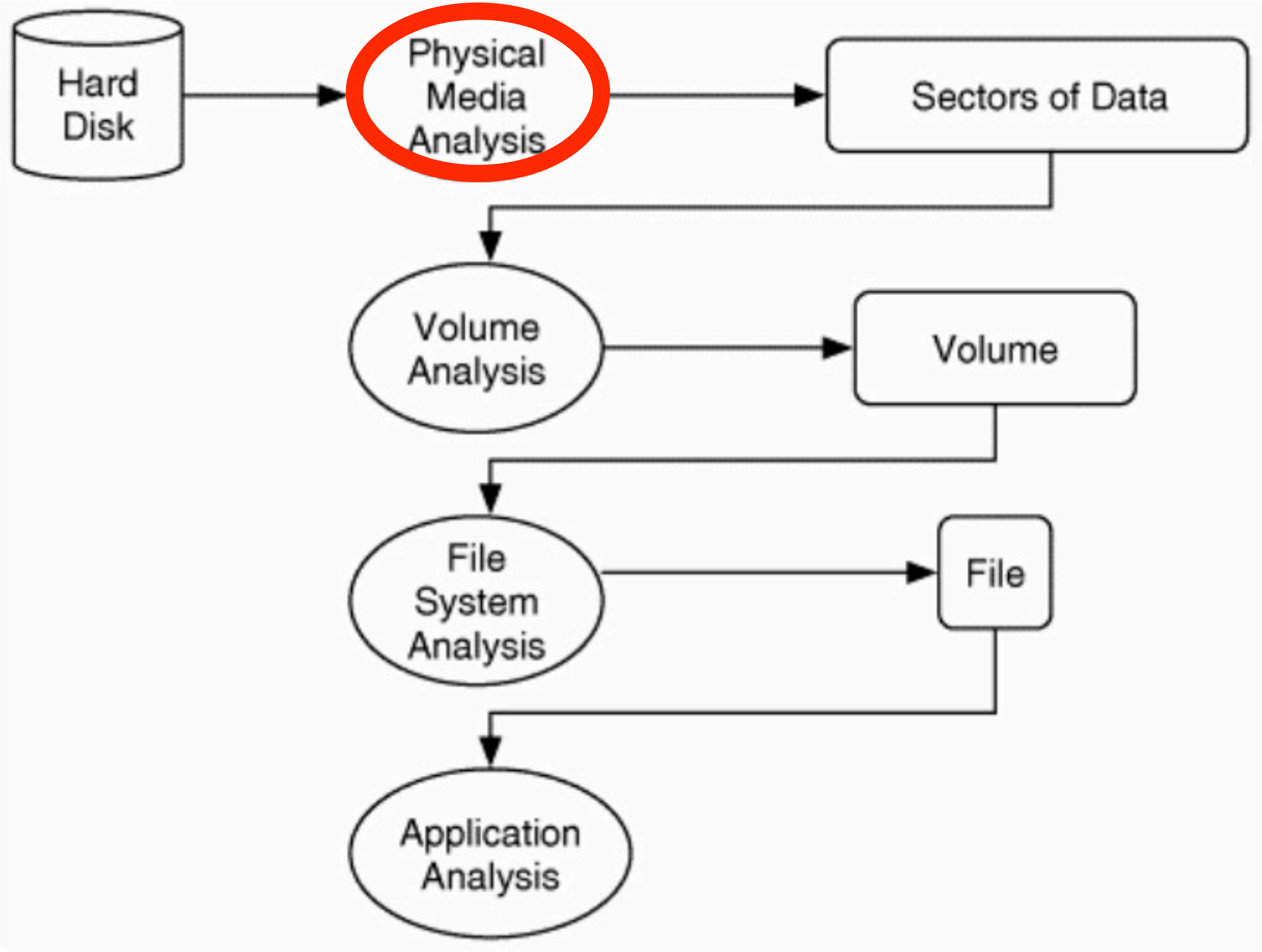
Adapted from Lee 2012

# Storage Hierarchy



http://en.wikipedia.org/wiki/File:ComputerMemoryHierarchy.svg

# Layers of Forensic Analysis

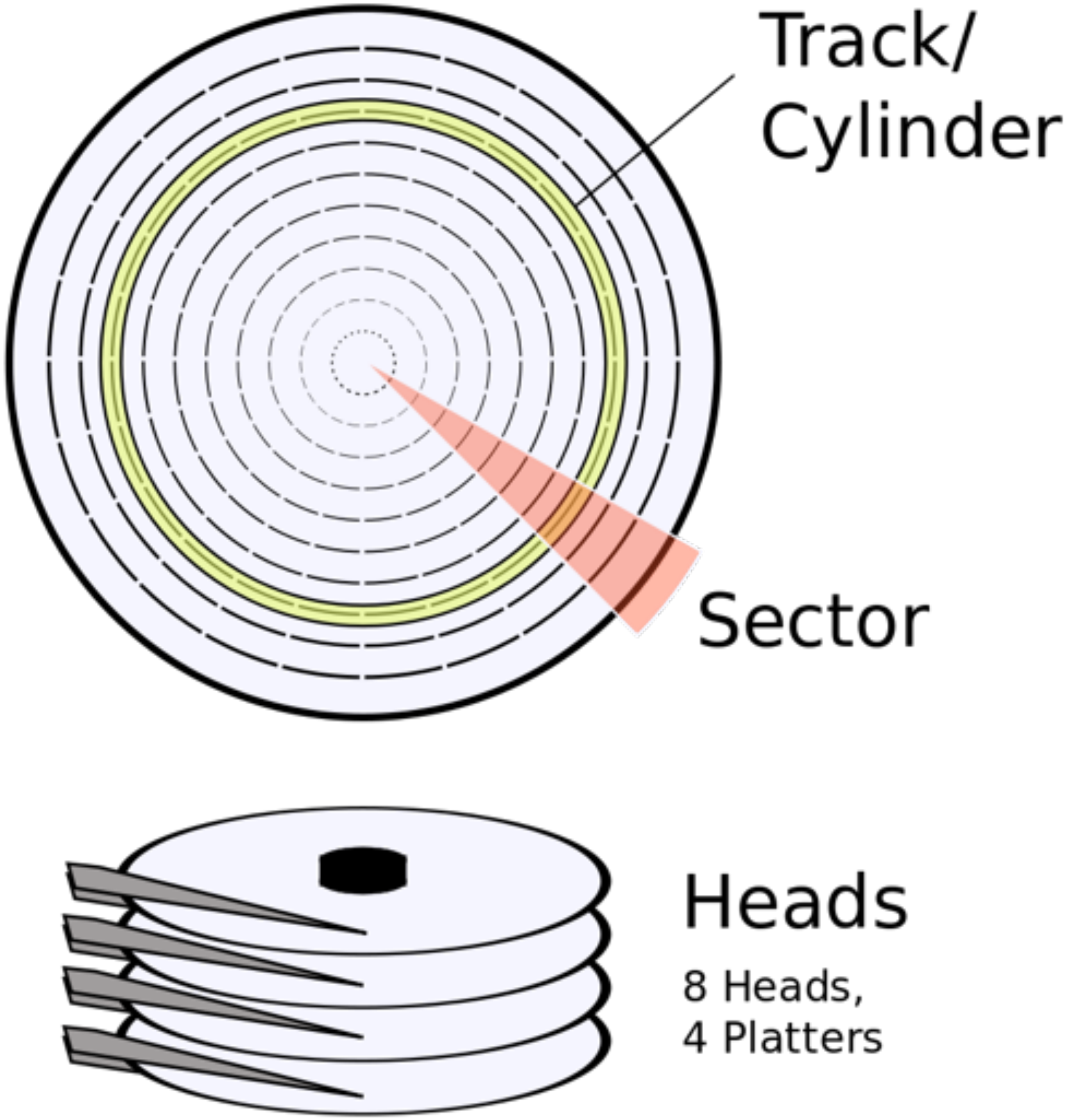


Brian Carrier, *File System Forensic Analysis* (2005), p. 18

# Physical Media & Signals

- Floppy disks and hard drives use changes in magnetic polarity (magnetic flux transitions); encoded and decoded using a particular algorithm
- Optical media (CDs/DVDs) use physically altered substrate with “pits”/“lands” that determine reflectivity of light; encoded and decoded using series of algorithms
- Flash memory uses stored amounts of electric charge

# Disk Geometry

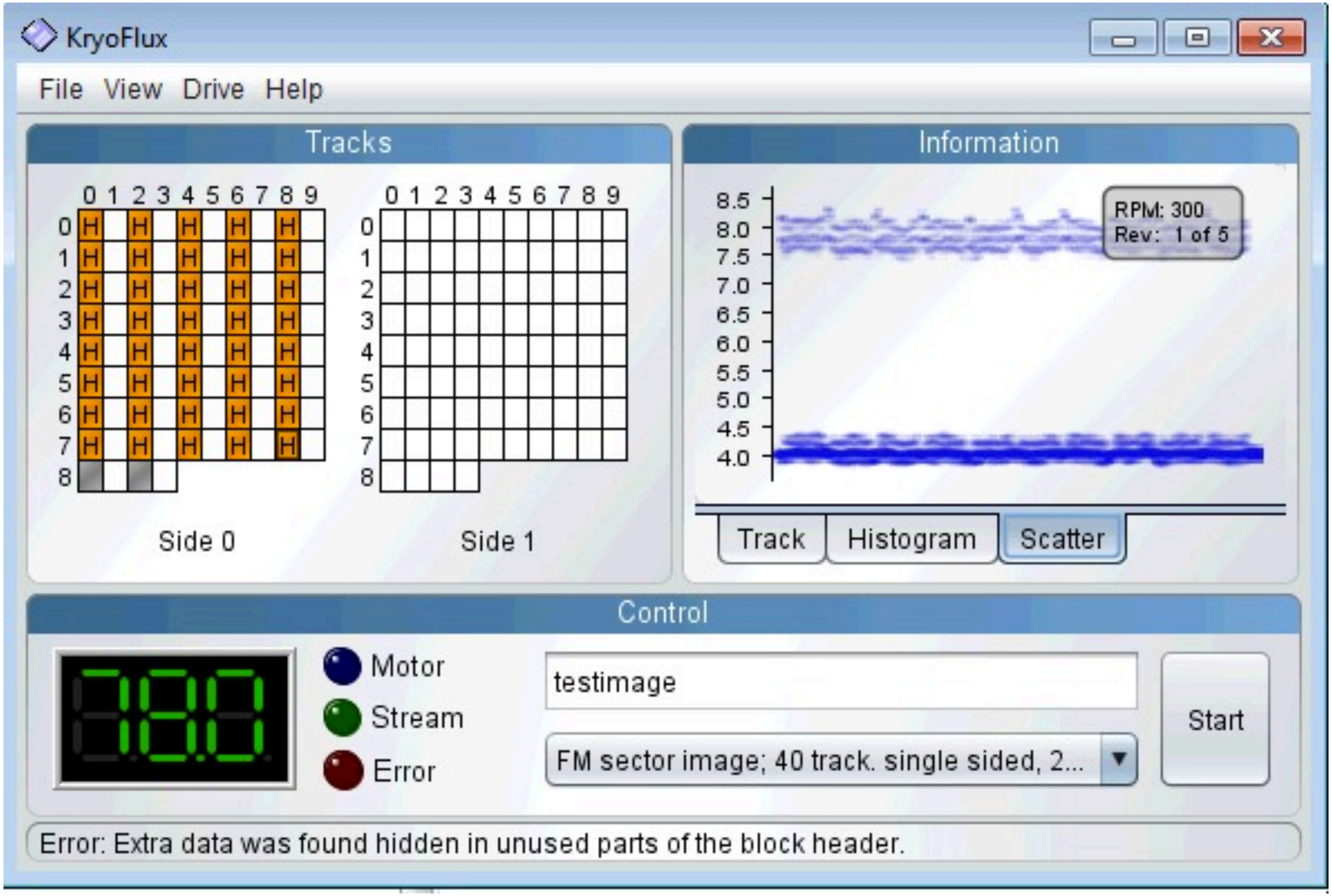


http://en.wikipedia.org/wiki/File:Cylinder_Head_Sector.svg

# Disk Imaging

- Process that runs through representation levels 0-2
- Uses drive to acquire analog signals stored on physical medium
- Those analog signals become analog electrical signals
- Hardware/software interprets those electrical signals into a bitstream using one or more algorithms

# Decoding Floppy Disks



Side 0

|   | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | H |  | H |  | H |  | H |  | H |  |
| 1 | H |  | H |  | H |  | H |  | H |  |
| 2 | H |  | H |  | H |  | H |  | H |  |
| 3 | H |  | H |  | H |  | H |  | H |  |
| 4 | H |  | H |  | H |  | H |  | H |  |
| 5 | H |  | H |  | H |  | H |  | H |  |
| 6 | H |  | H |  | H |  | H |  | H |  |
| 7 | H |  | H |  | H |  | H |  | H |  |
| 8 |  |  |  |  |  |  |  |  |  |  |

Side 1

|   | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 |  |  |  |  |  |  |  |  |  |  |
| 1 |  |  |  |  |  |  |  |  |  |  |
| 2 |  |  |  |  |  |  |  |  |  |  |
| 3 |  |  |  |  |  |  |  |  |  |  |
| 4 |  |  |  |  |  |  |  |  |  |  |
| 5 |  |  |  |  |  |  |  |  |  |  |
| 6 |  |  |  |  |  |  |  |  |  |  |
| 7 |  |  |  |  |  |  |  |  |  |  |
| 8 |  |  |  |  |  |  |  |  |  |  |

780

testimage

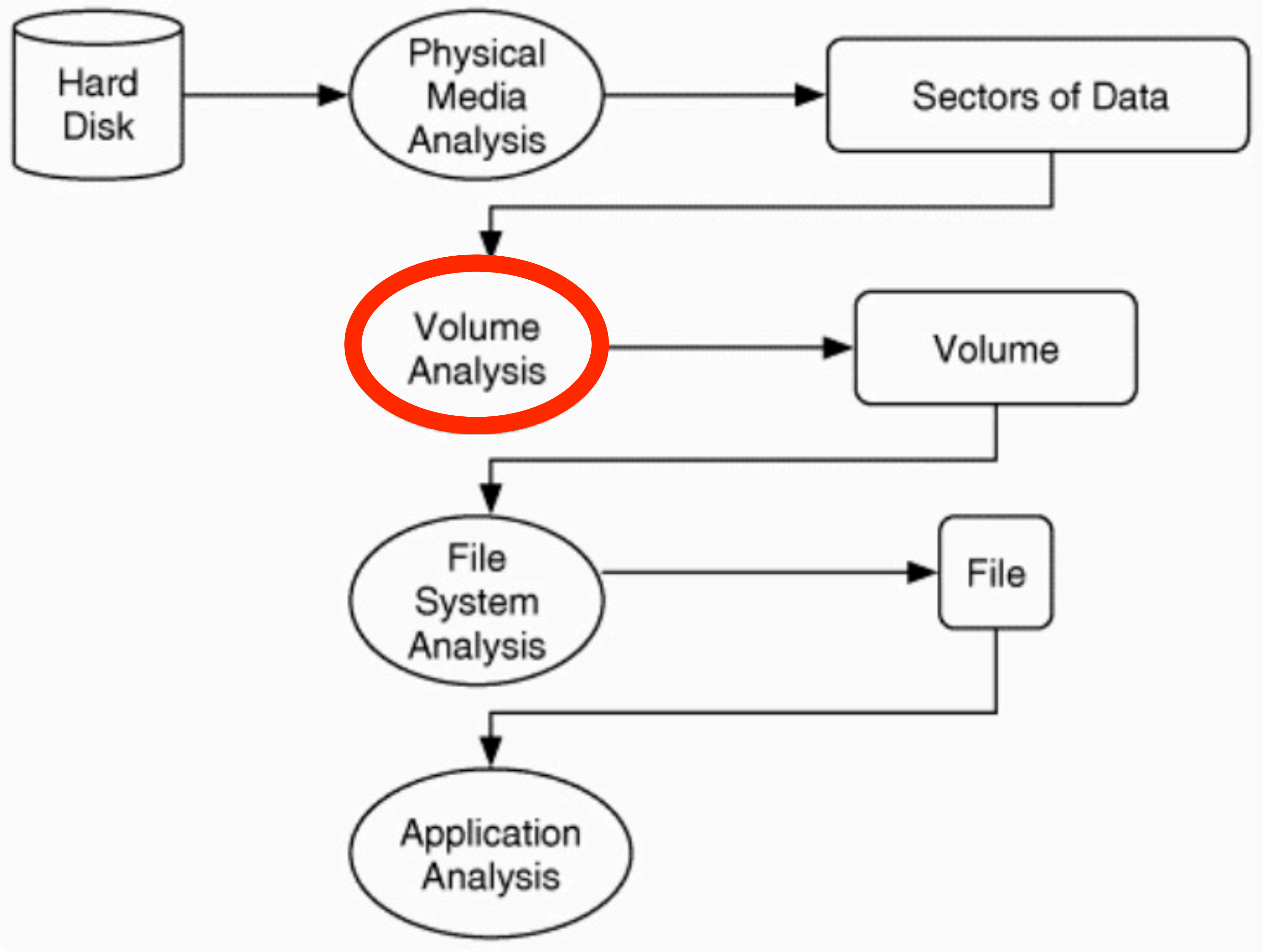
FM sector image; 40 track. single sided, 2...

Error: Extra data was found hidden in unused parts of the block header.

# Disk Image Formats

- Raw image (“dd”): decoded bitstream from media
- AFF: Open source; has embedded or external metadata
- EnCase E01: Proprietary with open source support; has embedded metadata
- Above formats can be split into multiple files
- Other formats: VMDK, DMG ...

# Layers of Forensic Analysis

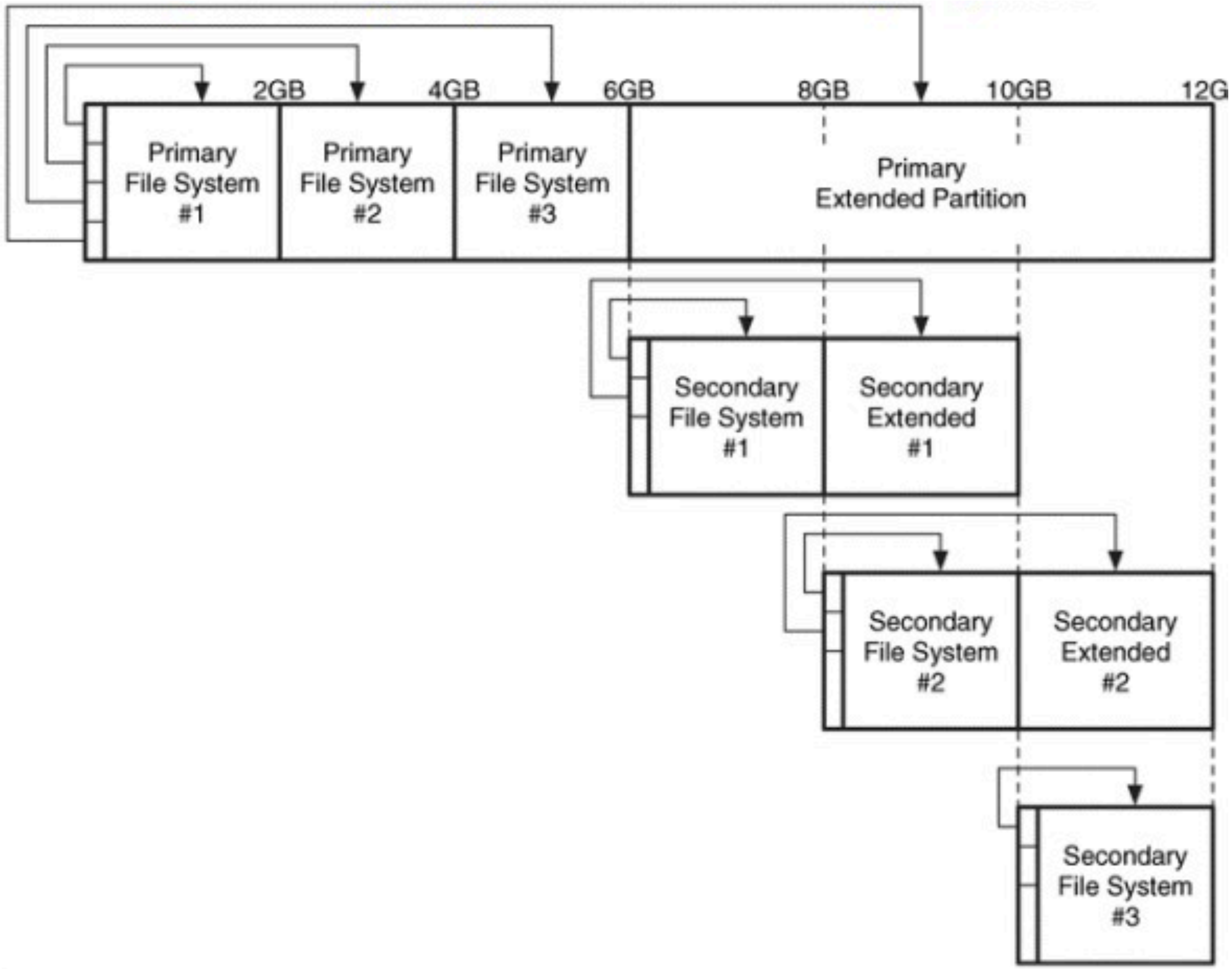


Brian Carrier, *File System Forensic Analysis* (2005), p. 18

# Volumes and Partitions

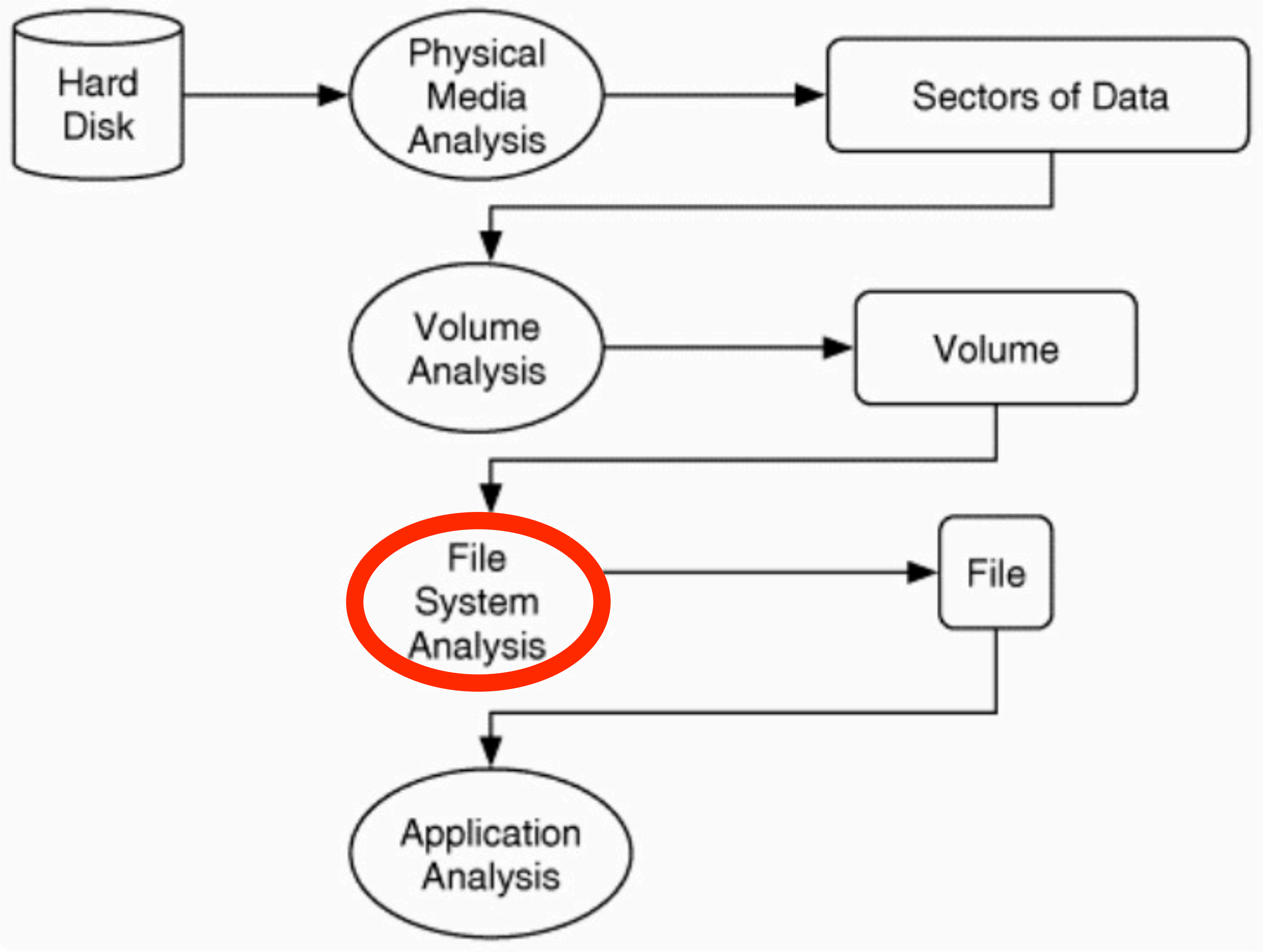
- Volume: collection of addressable sectors usable for storage
- Partition: a collection of consecutive sectors in a volume
- Partition map (or partition table): metadata structure that describes layout of partitions within a volume
- Partition formats: DOS, GPT, Apple, others

# Partition Organization



Brian Carrier, *File System Forensic Analysis* (2005), p. 69

# Layers of Forensic Analysis

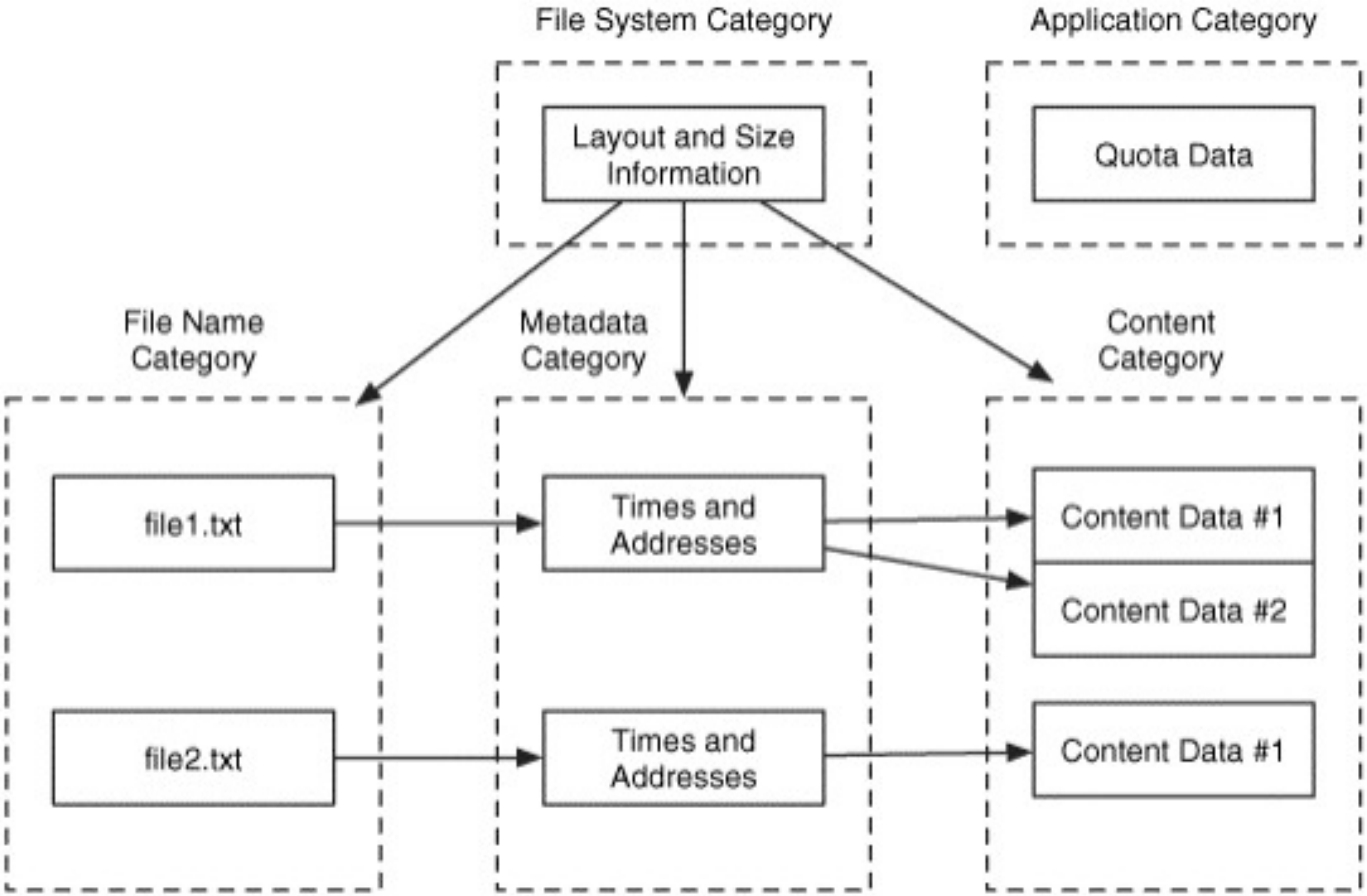


Brian Carrier, *File System Forensic Analysis* (2005), p. 18

# File Systems

- Mechanism to store data in series of files and directories with associated information about those data using unified set of procedures
- Separates information and content into “layers” or “categories”

# File System Data Categories

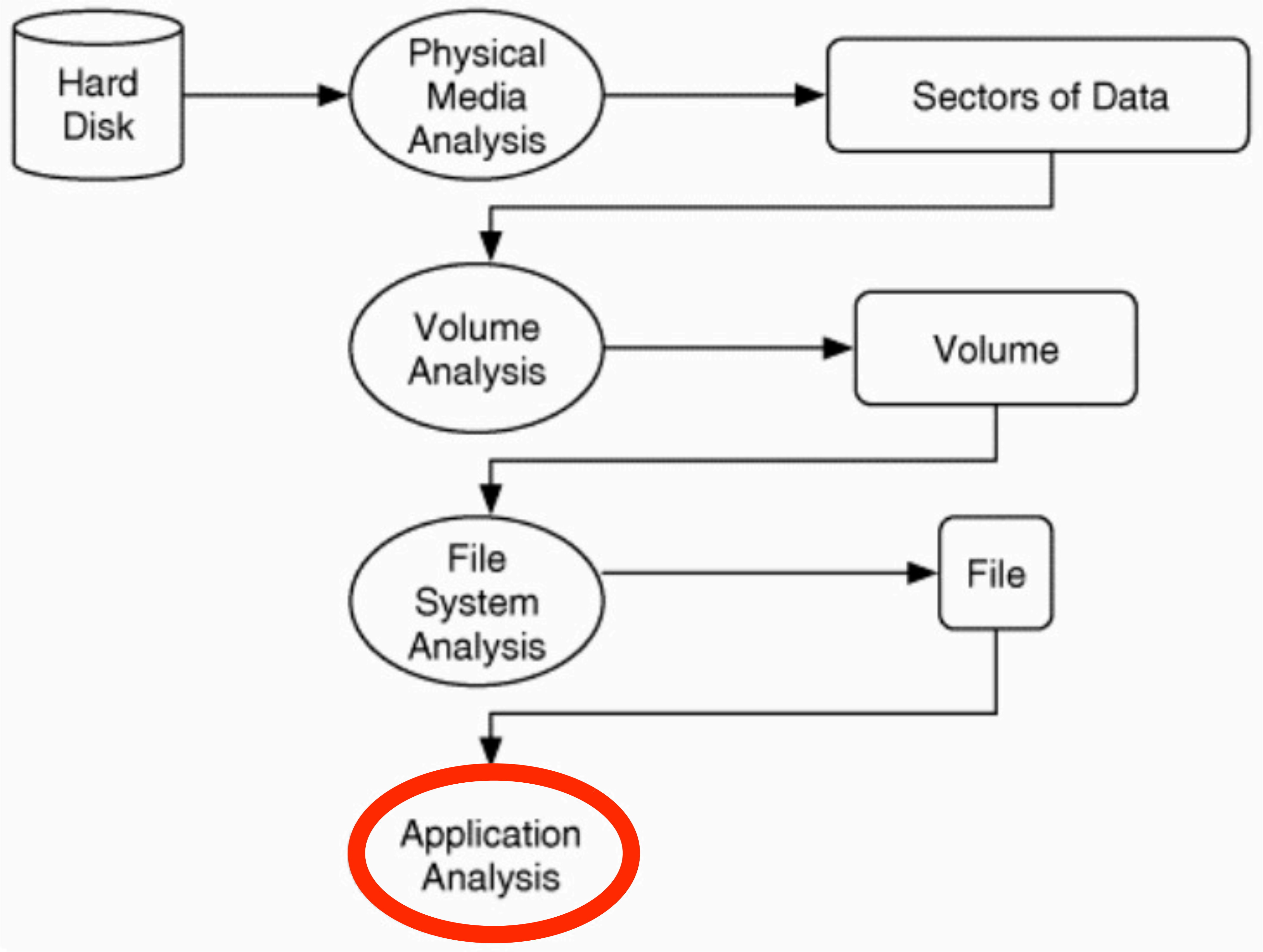


Brian Carrier, *File System Forensic Analysis* (2005), p. 130

# File System Types

- FAT (FAT12, FAT16, FAT32 ...): Early PCs (1981-) onward; still very common
- NTFS: Windows systems
- HFS+/HFSX: Mac systems (Mac OS 8.1+)
- ext2/ext3/ext4: Linux systems
- ISO9660: optical media
- Many, many others

# Layers of Forensic Analysis



Brian Carrier, *File System Forensic Analysis* (2005), p. 18

# Application-level Analysis

- File format identification
- Data recovery (“carving”)
- Checksum calculation/verification
- Searching for specific data
- Virus checking
- Searching for PII

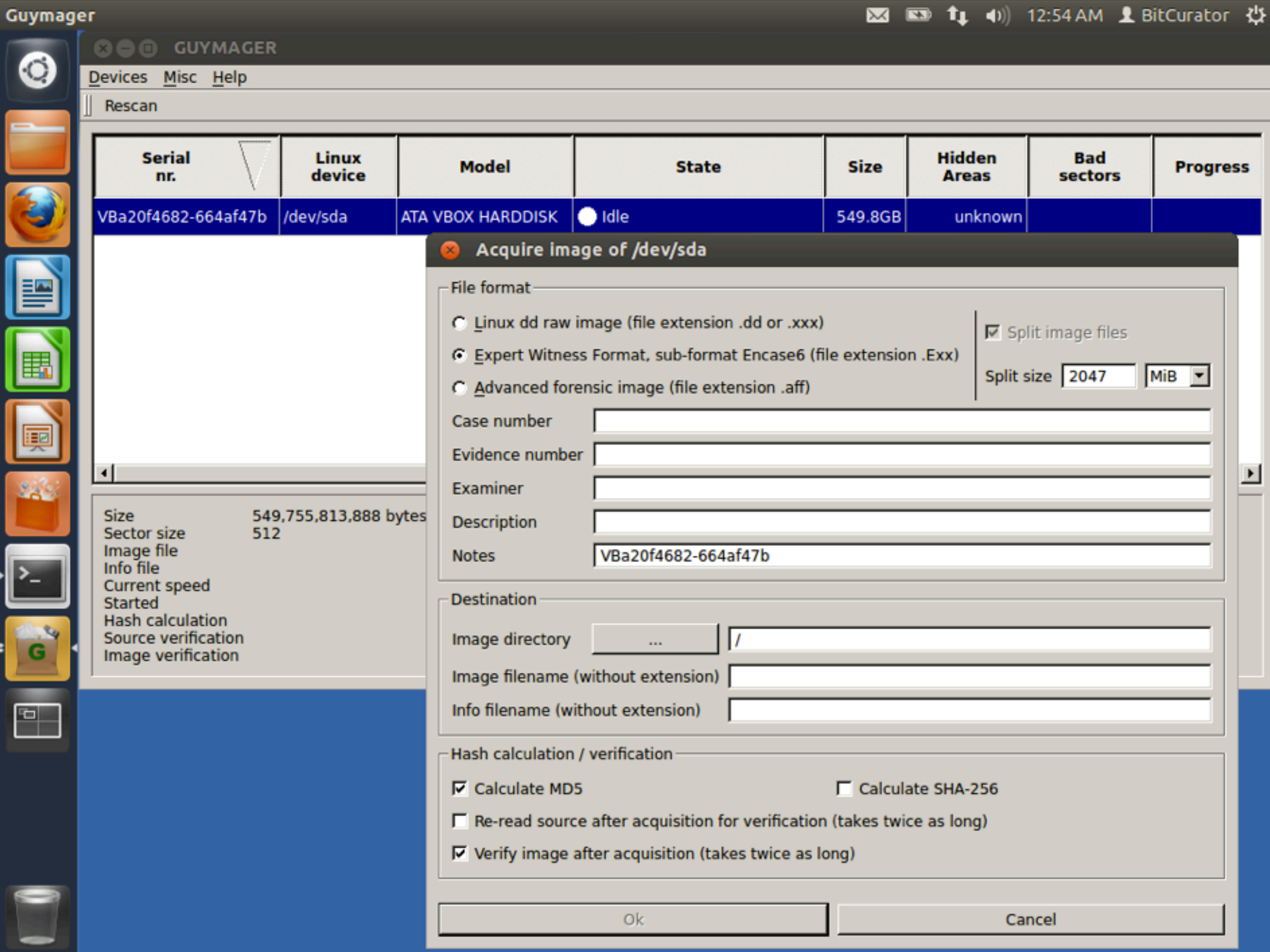
# Tools Overview

# BitCurator

- Project funded by Andrew W. Mellon Foundation
- Partners: UNC SILS and Maryland Institute for Technology in the Humanities
- Developing a system for cultural heritage sector that incorporates functionality of digital forensics tools into a common environment
- Still under development!
- http://bitcurator.net

# Guymager

- Disk imaging software
- Supports multiple imaging formats (raw, E01, AFF)
- Allows basic metadata entry and checksum calculation and verification



| Serial nr. | Linux device | Model | State | Size | Hidden Areas | Bad sectors | Progress |
|---|---|---|---|---|---|---|---|
| VBa20f4682-664af47b | /dev/sda | ATA VBOX HARDDISK | ● Idle | 549.8GB | unknown | | |

Size 549,755,813,888 bytes
Sector size 512
Image file
Info file
Current speed
Started
Hash calculation
Source verification
Image verification

Acquire image of /dev/sda

File format

○ Linux dd raw image (file extension .dd or .xxx)

● Expert Witness Format, sub-format Encase6 (file extension .Exx)

○ Advanced forensic image (file extension .aff)

☑ Split image files

Split size 2047 MiB

Case number

Evidence number

Examiner

Description

Notes VBa20f4682-664af47b

Destination

Image directory ... /

Image filename (without extension)

Info filename (without extension)

Hash calculation / verification

☑ Calculate MD5

□ Calculate SHA-256

□ Re-read source after acquisition for verification (takes twice as long)

☑ Verify image after acquisition (takes twice as long)

# The Sleuth Kit (TSK)

- Open source library, command line tools, and GUI application (Autopsy) for forensic analysis
- Supports analysis of FAT, NTFS, ISO9660, HFS+, Ext2/3, UFS1/2
- Splits tools into layers: volume system, file system, file name, metadata, data unit (“block”)
- Additional utilities to sort and post-process extracted metadata
- http://sleuthkit.org

# Image File Tools

- img_stat: Display information about a disk image
- img_cat: Dump the entire bitstream of a disk image (removes wrapper if using E01, AFF, etc.)

# Volume System Tools

- mmls: Display partition layout of a volume system
- mmstat: Display information about volume system
- mmcat: Dump the entire bitstream of a partition

# File System Layer Tools

- fsstat: Display file system details: layout, sizes, labels

# File Name Layer Tools

- fls: List allocated and unallocated file name entries
- ffind: Find allocated and unallocated file name entries that refer to a given metadata structure

# Metadata Layer Tools

- ils: List metadata structures and their contents
- ifind: Find metadata structure referred to by a specific file name entry
- istat: Display information about a specific metadata structure
- icat: Extract data units of a file specified by its metadata address

# Data Unit Layer Tools

- blkls: List details about data units, especially when unallocated
- blkstat: Display information about a specific data unit
- blkcat: Extract contents of a given data unit
- blkcalc: Calculate location of where data in unallocated space exists within a disk image

# Additional TSK Tools

- tsk_loaddb: Extract metadata into a SQLite database
- tsk_recover: Extract allocated or unallocated files from a disk image
- mactime: Create timeline of activity (using ils/fls input)
- sorter: Sorts files based on type (basic application-level analysis)

# fiwalk

- Command line program; depends on The Sleuth Kit
- Outputs results in multiple formats: Digital Forensics XML, CSV, plain text, ARFF (for data mining)
- Developed to support automated forensic processing by breaking it into three steps: extract, represent, process
- Can create plugins to allow for application-level analysis
- Faster in many cases; reads directly in sector order

# Digital Forensics XML

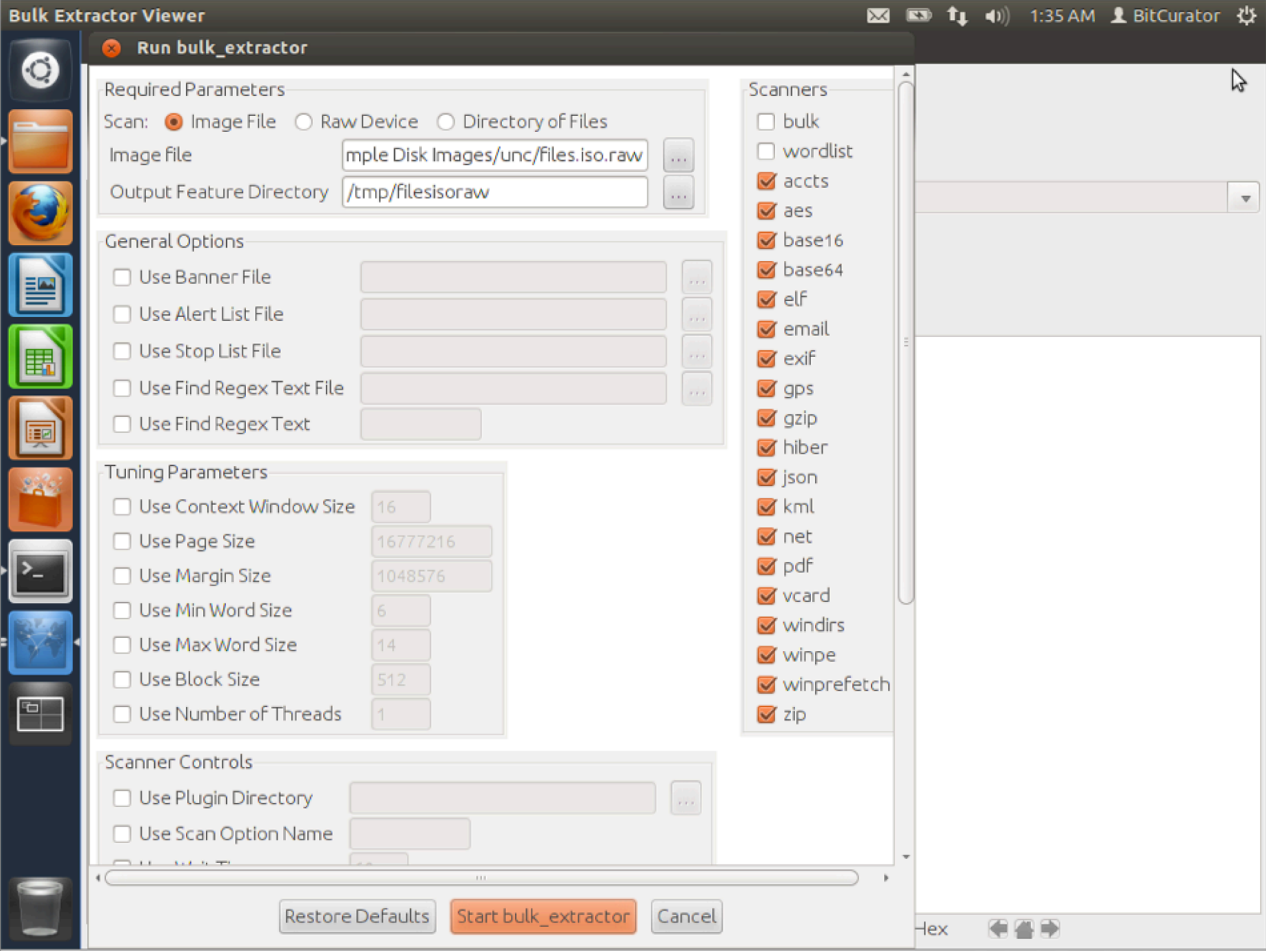
- Representation in XML of structured forensic information developed by Simson Garfinkel

- Easily extensible to incorporate additional data elements added by other tools

- Straightforward to process; has existing set of Python scripts for data processing and analysis

# bulk_extractor

- Performs bulk data analysis (reads entire bitstream in one pass instead of analyzing individual files)
- Command line program with additional GUI interface (BEViewer)
- Finds patterns: email addresses, phone numbers, URLs, SSNs, credit card numbers, GPS coordinates, EXIF metadata
- Very good for identifying potential PII issues



Run bulk_extractor

Required Parameters

Scan: ● Image File ○ Raw Device ○ Directory of Files

Image file: mple Disk Images/unc/files.iso.raw

Output Feature Directory: /tmp/filesisoraw

General Options

□ Use Banner File
□ Use Alert List File
□ Use Stop List File
□ Use Find Regex Text File
□ Use Find Regex Text

Tuning Parameters

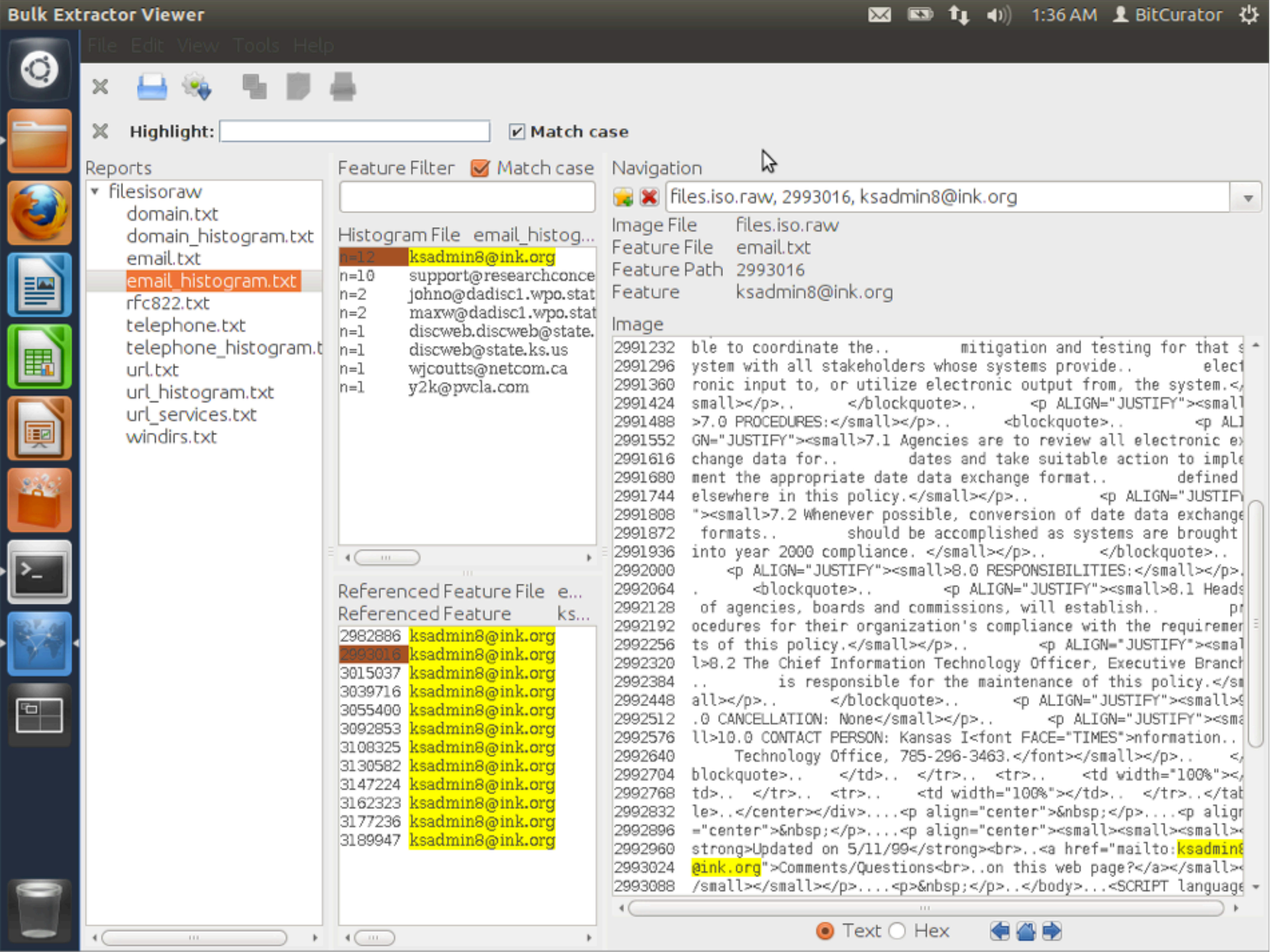
□ Use Context Window Size 16
□ Use Page Size: 16777216
□ Use Margin Size: 1048576
□ Use Min Word Size: 6
□ Use Max Word Size: 14
□ Use Block Size: 512
□ Use Number of Threads: 1

Scanner Controls

□ Use Plugin Directory
□ Use Scan Option Name

Scanners

□ bulk
□ wordlist
☑ accts
☑ aes
☑ base16
☑ base64
☑ elf
☑ email
☑ exif
☑ gps
☑ gzip
☑ hiber
☑ json
☑ kml
☑ net
☑ pdf
☑ vcard
☑ windirs
☑ winpe
☑ winprefetch
☑ zip



Highlight: ☑ Match case

Reports

- filesisoraw
  - domain.txt
  - domain_histogram.txt
  - email.txt
  - email_histogram.txt
  - rfc822.txt
  - telephone.txt
  - telephone_histogram.t
  - url.txt
  - url_histogram.txt
  - url_services.txt
  - windirs.txt

Feature Filter ☑ Match case

Histogram File email_histog...

| | |
|---|---|
| n=12 | ksadmin8@ink.org |
| n=10 | support@researchconce |
| n=2 | johno@dadisc1.wpo.stat |
| n=2 | maxw@dadisc1.wpo.stat |
| n=1 | discweb.discweb@state. |
| n=1 | discweb@state.ks.us |
| n=1 | wjcoutts@netcom.ca |
| n=1 | y2k@pvcla.com |

Referenced Feature File e...
Referenced Feature ks...

| | |
|---|---|
| 2982886 | ksadmin8@ink.org |
| 2993016 | ksadmin8@ink.org |
| 3015037 | ksadmin8@ink.org |
| 3039716 | ksadmin8@ink.org |
| 3055400 | ksadmin8@ink.org |
| 3092853 | ksadmin8@ink.org |
| 3108325 | ksadmin8@ink.org |
| 3130582 | ksadmin8@ink.org |
| 3147224 | ksadmin8@ink.org |
| 3162323 | ksadmin8@ink.org |
| 3177236 | ksadmin8@ink.org |
| 3189947 | ksadmin8@ink.org |

Navigation

files.iso.raw, 2993016, ksadmin8@ink.org

| | |
|---|---|
| Image File | files.iso.raw |
| Feature File | email.txt |
| Feature Path | 2993016 |
| Feature | ksadmin8@ink.org |

Image

```
2991232  ble to coordinate the..         mitigation and testing for that s
2991296  ystem with all stakeholders whose systems provide..         elect
2991360  ronic input to, or utilize electronic output from, the system.</
2991424  small></p>..      </blockquote>..        <p ALIGN="JUSTIFY"><small
2991488  >7.0 PROCEDURES:</small></p>..       <blockquote>..         <p ALI
2991552  GN="JUSTIFY"><small>7.1 Agencies are to review all electronic ex
2991616  change data for..         dates and take suitable action to imple
2991680  ment the appropriate date data exchange format..         defined
2991744  elsewhere in this policy.</small></p>..        <p ALIGN="JUSTIFY
2991808  "><small>7.2 Whenever possible, conversion of date data exchange
2991872   formats..        should be accomplished as systems are brought
2991936  into year 2000 compliance. </small></p>..       </blockquote>..
2992000      <p ALIGN="JUSTIFY"><small>8.0 RESPONSIBILITIES:</small></p>.
2992064  .     <blockquote>..          <p ALIGN="JUSTIFY"><small>8.1 Heads
2992128   of agencies, boards and commissions, will establish..        pr
2992192  ocedures for their organization's compliance with the requiremen
2992256  ts of this policy.</small></p>..         <p ALIGN="JUSTIFY"><smal
2992320  l>8.2 The Chief Information Technology Officer, Executive Branch
2992384  ..         is responsible for the maintenance of this policy.</sm
2992448  all></p>..      </blockquote>..        <p ALIGN="JUSTIFY"><small>9
2992512  .0 CANCELLATION: None</small></p>..       <p ALIGN="JUSTIFY"><sma
2992576  ll>10.0 CONTACT PERSON: Kansas I<font FACE="TIMES">nformation..
2992640       Technology Office, 785-296-3463.</font></small></p>..      </
2992704  blockquote>..     </td>..   </tr>..   <tr>..     <td width="100%"></
2992768  td>..  </tr>..  <tr>..     <td width="100%"></td>..  </tr>..</tab
2992832  le>..</center></div>....<p align="center">&nbsp;</p>....<p align
2992896  ="center">&nbsp;</p>....<p align="center"><small><small><small><
2992960  strong>Updated on 5/11/99</strong><br>..<a href="mailto:ksadmin8
2993024  @ink.org">Comments/Questions<br>..on this web page?</a></small><
2993088  /small></small></p>....<p>&nbsp;</p>..</body>...<SCRIPT language
```

● Text ○ Hex

# Hands-On

- Start BitCurator VM on VirtualBox
- Walkthrough of command line tools
- Walkthrough of bulk_extractor/BEViewer
- Other tools if we have time (ghex Hex Editor, sdhash)
- Username and password (if it asks) are both “bcadmin”

# Thanks!
# Questions?

Mark A. Matienzo
mark@matienzo.org
http://matienzo.org
@anarchivist

# References

- AIMS Work Group (2012). *AIMS Born-Digital Collections: An Inter-Institutional Model for Stewardship.* http://www2.lib.virginia.edu/aims/whitepaper/
- Carrier, B. (2003). "Defining Digital Forensic Examination and Analysis Tools Using Abstraction Layers." *International Journal of Digital Evidence* 1(4).
- Carrier, B. (2005). *File System Forensic Analysis*. Boston and London: Addison Wesley.
- Duranti, L. (2009). "From Digital Diplomatics to Digital Records Forensics." *Archivaria* 68, 39-66
- Garfinkel, S. (2011). "Digital Media Triage with Bulk Data Analysis and bulk_extractor." http://simson.net/ref/2011/bulk_extractor.pdf
- Garfinkel, S. (2012). "Digital Forensics XML and the DFXML Toolset." *Digital Investigation* 8, 161-174.
- Kirschenbaum, M.G., et al. (2010). *Digital Forensics and Born-Digital Content in Cultural Heritage Collections*. Washington: Council on Library and Information Resources. http://www.clir.org/pubs/reports/pub149
- Lee, C.A. (2012). "Archival Application of Digital Forensics Methods for Authenticity, Description, and Access Provision." International Council on Archives Congress, August 20-24, 2012, Brisbane, Australia. http://ils.unc.edu/callee/ica-2012-lee.pdf and http://www.ica2012.com/files/data/Full%20papers%20upload/ica12Final00290.pdf
- Lee, C.A., et al. (2012). "BitCurator: Tools and Techniques for Digital Forensics in Collecting Institutions." *D-Lib Magazine* 18(5/6).
- Lee, C.A. and Woods, K. (2011). "Digital Acquisition Learning Laboratory: A White Paper." School of Information and Library Science, University of North Carolina at Chapel Hill. http://www.ils.unc.edu/callee/dall-white-paper.pdf
- Ross, S. and Gow, A. (1999). *Digital Archaeology: Rescuing Neglected and Damaged Data Resources.* A JISC/NPO Study within the Electronic Libraries (eLib) Programme on the Preservation of Electronic Materials. http://eprints.erpanet.org/47/
- Woods, K., Lee, C.A., and Garfinkel, S. (2011). "Extending Digital Repository Architectures to Support Disk Image Preservation and Access." In *JCDL '11*.
- Xie, S.L. (2011). "Building Foundations for Digital Records Forensics: A Comparative Study of the Concept of Reproduction in Digital Records Management and Digital Forensics." *American Archivist 74*(2), *576-599*.