

# BITCURATOR FOR STATE ARCHIVES

---

Use the chat box at the right of the screen to tell us who you are, where you're from, and who is participating with you today.

(To open the chat window, click on the CHAT icon in the upper right corner.)

Connect to the audio portion of the webinar through your phone line or through VoIP.



# WELCOME!

---

## Fall 2019 State Electronic Records Initiative webinars

- October: social media
- November: scripting for archivists
- December: developing processing workflows

Spring 2020 SERI webinars will begin in January

<https://www.statearchivists.org/programs/state-electronic-records-initiative/seri-webinars/>



**Sarah Koonts**  
State Archives of  
North Carolina



INSTITUTE of  
**Museum and Library**  
SERVICES



**SERI**  
State Electronic  
Records Initiative  
COUNCIL OF STATE ARCHIVISTS

# SERI ONLINE RECORDINGS

---

- **Managing Digital Content Over Time**
  - Identify
  - Select
  - Store
  - Protect
  - Manage
  - Provide access
- <https://archives.utah.gov/community/SERI/SERI-resources.html>
- <https://www.statearchivists.org/programs/videos-and-recordings/state-electronic-records-initiative-training-recordings/>

# NEW SERI VIDEO

---

- **FAQs on Bit Rot**
  - <https://www.youtube.com/user/StateArchivists>

# TODAY'S PRESENTER

---

**Cal Lee**

School of Information and Library Science  
University of North Carolina at Chapel Hill



# Some Goals When Acquiring Materials

**Ensure integrity of materials**

**Allow users to make sense of materials and understand their context**

**Prevent inadvertent disclosure of sensitive data**

# Some Fundamental Principles

## 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”<sup>1</sup>
- Ideal recordkeeping system would provide “an unblemished line of responsible custody”<sup>2</sup>

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.

## But what about these?



Source: "Digital Forensics and creation of a narrative." *Da Blog: ULCC Digital Archives Blog*.  
<http://dablog.ulcc.ac.uk/2011/07/04/forensics/>



# 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 equipment	Series of 1s and 0s as accessed from the storage media using input/output hardware and software (e.g. controllers, drivers, ports, connectors)
1	Raw signal stream through I/O equipment	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)
0	Bitstream on physical medium	Physical properties of the storage medium that are interpreted as bitstreams at Level 1

# Interaction Examples

## Level

### Aggregation of objects

### Object or package

### In-application rendering

### File through filesystem

### File as “raw” bitstream

### Sub-file data structure

### Bitstream through I/O equipment

### Raw signal stream through I/O equipment

### Bitstream on physical medium

ContextMiner Alpha 3.0

[\[Home\]](#)[\[Publications\]](#)[\[Reports\]](#)[\[Add\]](#)[\[View\]](#)[\[Search\]](#)[\[Profile\]](#)[\[Visualize\]](#)[\[Monitor\]](#)[\[Tools\]](#)[\[Developer\]](#)

This page lists all the seed queries that are used for monitoring videos related to elections on YouTube. Clicking on a query will show all the results collected over several crawls. Total number of these results are also listed here for each query. The last column in the following table shows how many total results YouTube had for a given query during our latest crawl. Clicking on 'Setup' associated with a query will bring up an interface where the curator can specify what constitutes as a "significant" change for a video of that query.

#	Query	Setup	Total results so far	Max results on last crawl
1	<a href="#">election 2008</a>	<a href="#">Setup</a>	574	6150
2	<a href="#">US election 2008</a>	<a href="#">Setup</a>	349	795
3	<a href="#">United States election 2008</a>	<a href="#">Setup</a>	216	257
4	<a href="#">presidential election 2008</a>	<a href="#">Setup</a>	206	1820
5	<a href="#">campaign 2008</a>	<a href="#">Setup</a>	273	2530
6	<a href="#">decision 2008</a>	<a href="#">Setup</a>	168	142
7	<a href="#">Joe Biden</a>	<a href="#">Setup</a>	209	1080
8	<a href="#">Hillary Rodham Clinton</a>	<a href="#">Setup</a>	193	353
9	<a href="#">Christopher Dodd</a>	<a href="#">Setup</a>	267	815
10	<a href="#">John Edwards</a>	<a href="#">Setup</a>	902	7540
11	<a href="#">Mike Gravel</a>	<a href="#">Setup</a>	301	1210
12	<a href="#">Dennis Kucinich</a>	<a href="#">Setup</a>	229	1600
13	<a href="#">Barack Obama</a>	<a href="#">Setup</a>	861	9140
14	<a href="#">Bill Richardson</a>	<a href="#">Setup</a>	287	1100
15	<a href="#">Wesley Clark</a>	<a href="#">Setup</a>	191	375
16	<a href="#">Al Gore</a>	<a href="#">Setup</a>	613	4910
17	<a href="#">Tom Vilsack</a>	<a href="#">Setup</a>	89	68
18	<a href="#">Sam Brownback</a>	<a href="#">Setup</a>	254	404
19	<a href="#">John McCain</a>	<a href="#">Setup</a>	22	16



INSTITUTE of  
Museum and Library  
SERVICES



**SERI**  
State Electronic  
Records Initiative  
COUNCIL OF STATE ARCHIVISTS

# Interaction Examples

## Level

Aggregation of objects

Object or package

In-application rendering

File through filesystem

File as “raw” bitstream

Sub-file data structure

Bitstream through I/O equipment

Raw signal stream through equipment

Bitstream on physical medium

ContextMiner Alpha 3.0

[\[Home\]](#)[\[Publications\]](#)[\[Reports\]](#)[\[Add\]](#)[\[View\]](#)[\[Search\]](#)[\[Profile\]](#)[\[Visualize\]](#)[\[Monitor\]](#)[\[Tools\]](#)[\[Developer\]](#)

This page presents contextual information for a video captured over a number of days. Contextual information is defined as the information about a video that may change with time. Usually this information is contributed by the visitors of the video page. [See](#) the metadata information for this video. Description of various attributes displayed is given [here](#).



Query: Rudy Giuliani

[I Got A Crush On.... Giuliani](#)

Collaboration with the very talented JackDanyells, who came up with the concept for this video. Check out his channel at: <http://www.youtube.com/jackdanyells> -Lyrics by JackDanyells -Vocal melody composed and sung by me -Royalty free background music from sounddogs.com

Comedy

Crawling since 2007-07-19

Color coding for % changes

< 0.05 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0 5.0 >

Crawl #	Crawl date	Rank	Views	Ratings	Avg Rating	Comments	Links	Favorited	Honors	Change
1	2007-07-31	5	27357	301	3.74	288	5	44	0	--
2	2007-08-01	5	27452	303	3.73	290	5	44	0	--
3	2007-08-02	5	27780	307	3.72	291	5	45	0	--
4	2007-08-03	5	28048	309	3.71	291	5	45	0	--
5	2007-08-04	2	28398	310	3.71	291	5	45	0	--
6	2007-08-05	2	28443	314	3.69	294	5	45	0	--
7	2007-08-06	3	28980	314	3.69	296	5	45	0	--
8	2007-08-07	3	29265	318	3.65	298	5	45	0	--
9	2007-08-08	3	29551	319	3.65	299	5	46	0	--
10	2007-08-09	3	30094	320	3.64	300	5	47	0	--
11	2007-08-10	3	30384	323	3.61	302	5	47	0	--
12	2007-08-10	5	30419	324	3.62	303	5	48	0	--
13	2007-08-11	3	30540	324	3.62	305	5	49	0	--
14	2007-08-12	3	30697	326	3.61	306	5	49	0	--
15	2007-08-13	3	30848	326	3.61	306	5	49	0	--
16	2007-08-14	3	31036	326	3.61	306	5	49	0	--
17	2007-08-15	2	31181	326	3.61	306	5	49	0	--
18	2007-08-16	2	31321	326	3.61	307	5	51	0	--
19	2007-08-17	2	31459	327	3.61	307	5	51	0	--
20	2007-08-18	2	31662	331	3.59	308	5	51	0	--
21	2007-08-19	2	31792	332	3.58	308	5	51	0	--
22	2007-08-20	2	31937	335	3.57	310	5	51	0	--
23	2007-08-21	2	32135	335	3.57	311	5	52	0	--
24	2007-08-22	2	32404	335	3.57	311	5	54	0	--

# Interaction Examples

## Level

Aggregation of objects

Object or package

In-application rendering

File through filesystem

File as “raw” bitstream

Sub-file data structure

Bitstream through I/O  
equipment

Raw signal stream through I/O  
equipment

Bitstream on physical medium

The screenshot shows the YouTube interface for a video titled "Vote Different". The video features a woman in a white tank top with an Obama campaign logo, holding a rifle. The video player shows a progress bar at 0:16 / 1:14. Below the video, the rating is 4.5 stars (12,058 ratings) and the view count is 5,268,816. The right sidebar shows the channel "ParkRidge47" with a "Subscribe" button, and a list of related videos including "Barack Obama Hillary Clinton - Umbrella", "The Shocking Video Hillary Does NOT Want You To See! (1of2)", and "Obama Girl Returns for Iowa! (Why Obama Won)".

## Level

Aggregation of objects

Object or package

In-application rendering

File through filesystem

File as “raw” bitstream

Sub-file data structure

Bitstream through I/O  
equipment

Raw signal stream through I/O  
equipment

Bitstream on physical medium

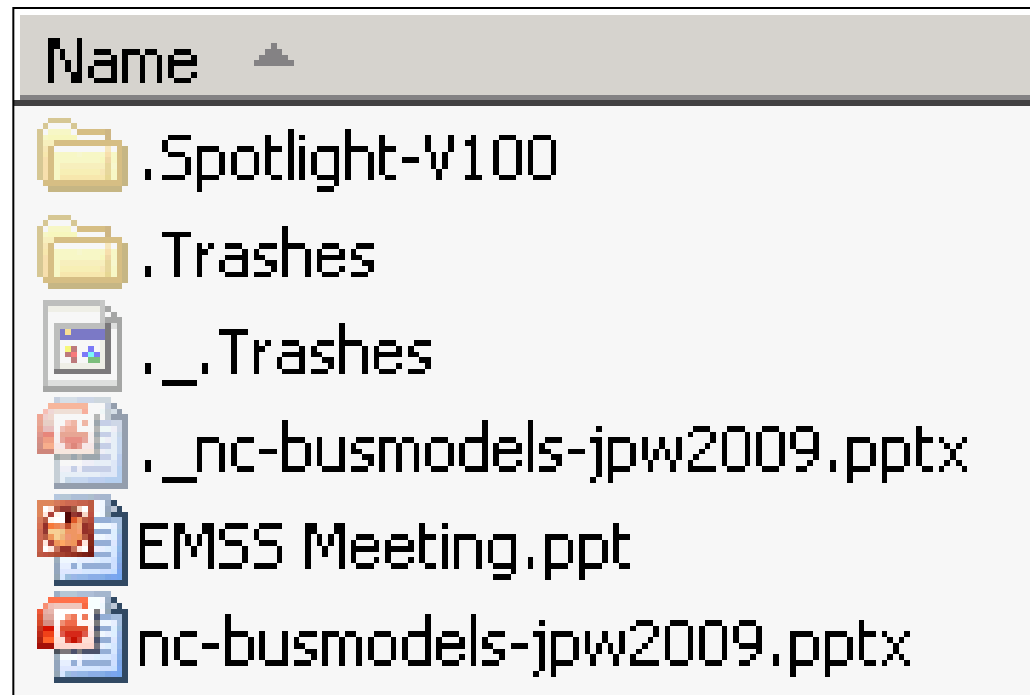
```
C:\WINDOWS\system32\cmd.exe
Microsoft Windows XP [Version 5.1.2600]
(C) Copyright 1985-2001 Microsoft Corp.

G:\>dir /a
Volume in drive G is KINGSTON
Volume Serial Number is 17E9-242F

Directory of G:\

03/12/2009  08:54 AM                4,096  ._.Trashes
03/12/2009  08:54 AM                <DIR>  .Trashes
03/12/2009  08:54 AM                <DIR>  .Spotlight-V100
03/11/2009  07:07 PM          1,023,213  nc-busmodels-jpw2009.pptx
03/12/2009  08:55 AM                4,096  .nc-busmodels-jpw2009.pptx
03/31/2009  01:23 PM          6,442,496  EMSS Meeting.ppt
                                4 File(s)          7,473,901 bytes
                                2 Dir(s)          120,145,920 bytes free

G:\>
```



# Interaction Examples

## Level

Aggregation of objects

Object or package

In-application rendering

File through filesystem

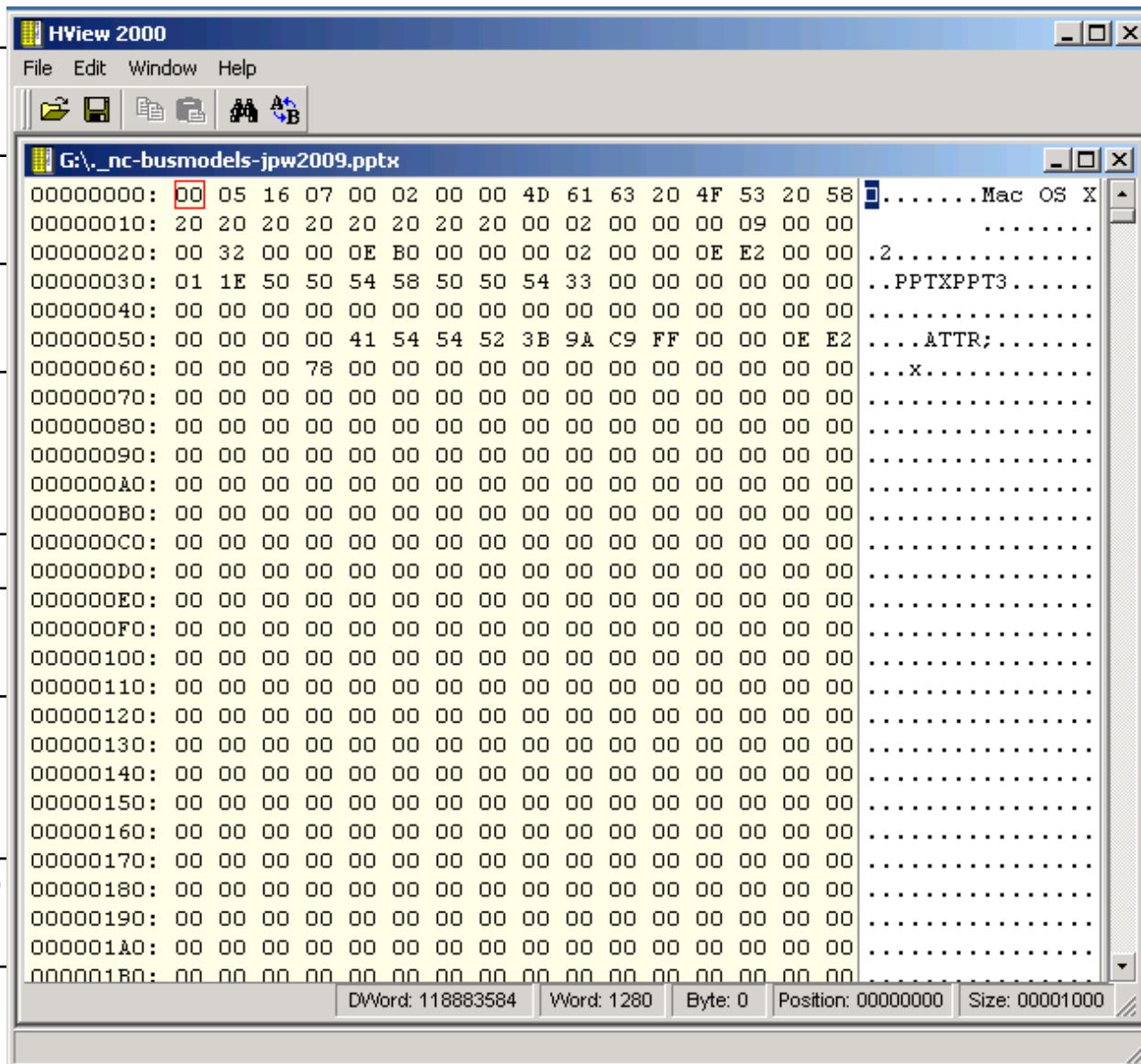
File as “raw” bitstream

Sub-file data structure

Bitstream through I/O  
equipment

Raw signal stream through I/O  
equipment

Bitstream on physical medium



# Interaction Examples

## Level

Aggregation of objects

Object or package

In-application rendering

File through filesystem

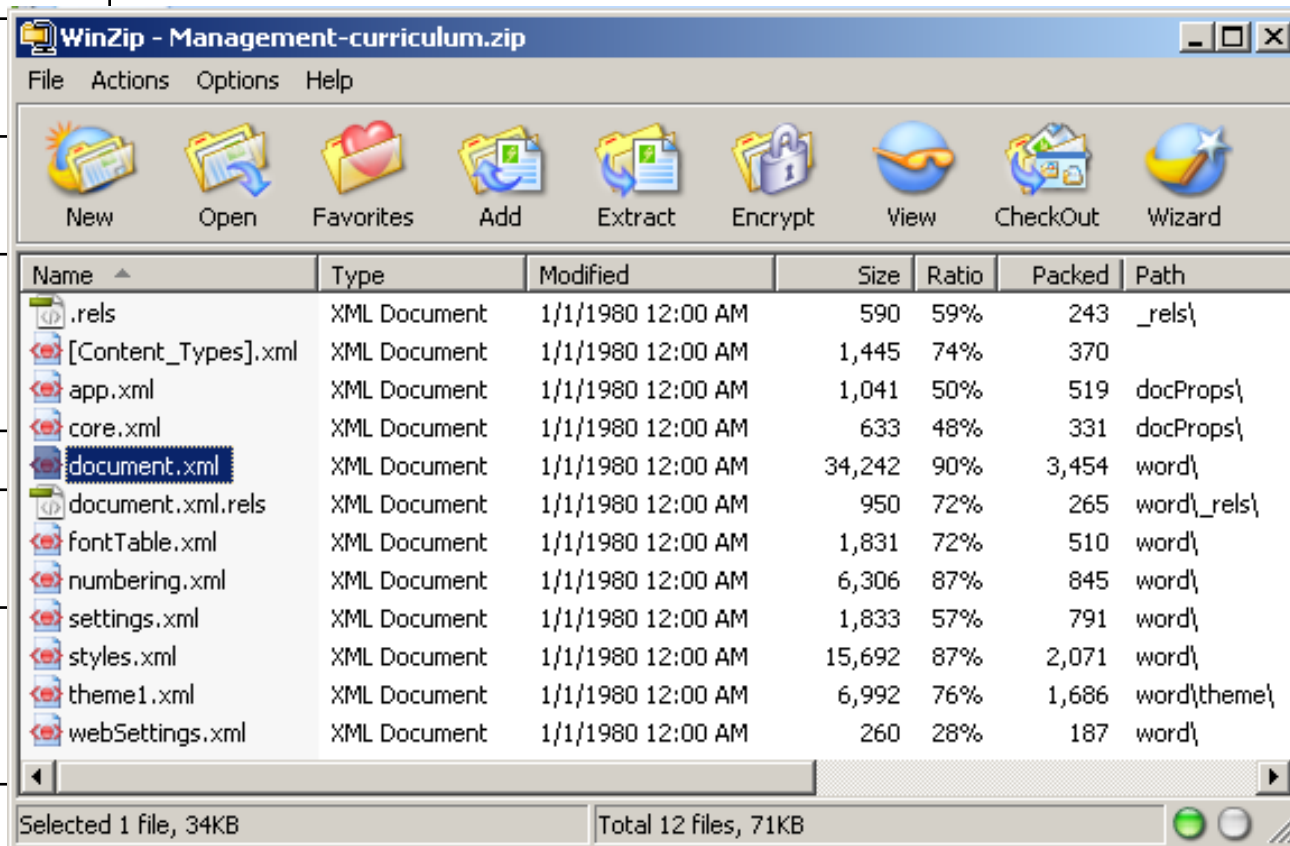
File as “raw” bitstream

Sub-file data structure

Bitstream through I/O  
equipment

Raw signal stream through  
equipment

Bitstream on physical medium





# Interaction Examples

## Level

Aggregation of object

Object or package

In-application render

File through filesystem

File as “raw” bitstream

Sub-file data structure

Bitstream through  
I/O equipment

Raw signal stream through  
I/O equipment

Bitstream on physical medium

The screenshot shows the Guymager application window. The top menu bar includes 'Devices', 'Misc', and 'Help'. Below the menu is a 'Rescan' button. The main window displays a table of devices with the following columns: Serial nr., Linux device, Model, State, Size, Hidden Areas, Bad sectors, and Progress.

Serial nr.	Linux device	Model	State	Size	Hidden Areas	Bad sectors	Progress
100726PBN303GTHXUWUS	/dev/sda	ATA HITACHI HTS545032B9A300	Idle	320.1GB	unknown		
20071114173400000	/dev/sdb	Generic- Multi-Card	Acquisition running	2.0GB	unknown	0	8%

Below the table, a detailed view of the selected device (20071114173400000) is shown, including the following information:

- Size: 2,032,664,576 bytes (1.89GiB / 2.03GB)
- Sector size: 512
- Image file: /home/kam/Desktop/Datasets/SDCardImageMay2012.E??
- Info file: /home/kam/Desktop/Datasets/SDCardImageMay2012.info
- Current speed: 8.32 MB/s
- Started: 26. May 11:18:23 (00:00:37)
- Hash calculation: MD5 and SHA-256
- Source verification: on
- Image verification: on

The interface also includes a sidebar with various icons for different functions and a bottom status bar.



# Interaction Examples

## Examples



In-application rendering

Browsing the contents of an archival collection using a finding aid

Viewing a web page that contains several files, including HTML, a style sheet and several images

Using Microsoft Excel to view an .xls file, watching an online

File through filesystem



Using Windows Explorer, typing to show the contents of a

File as "raw" bitstream

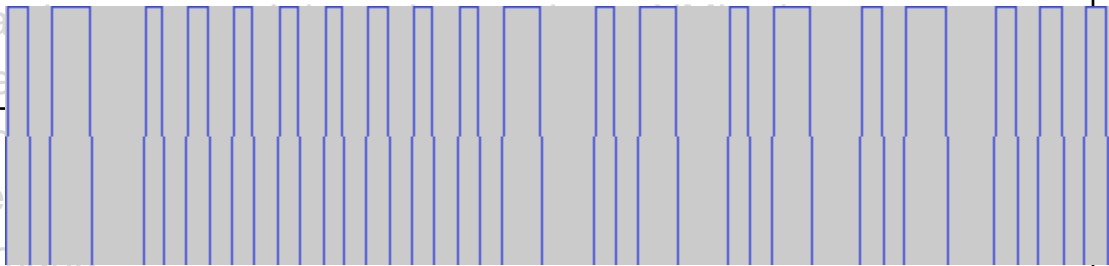
Opening an individual file in a hex editor

Sub-file data structure

Extra value

Bitstream through I/O equipment

Conn gene command



**Raw signal stream through I/O equipment**

Connecting a floppy drive to a host computer and then generating a magnetic flux transition image of the disk

Bitstream on physical medium

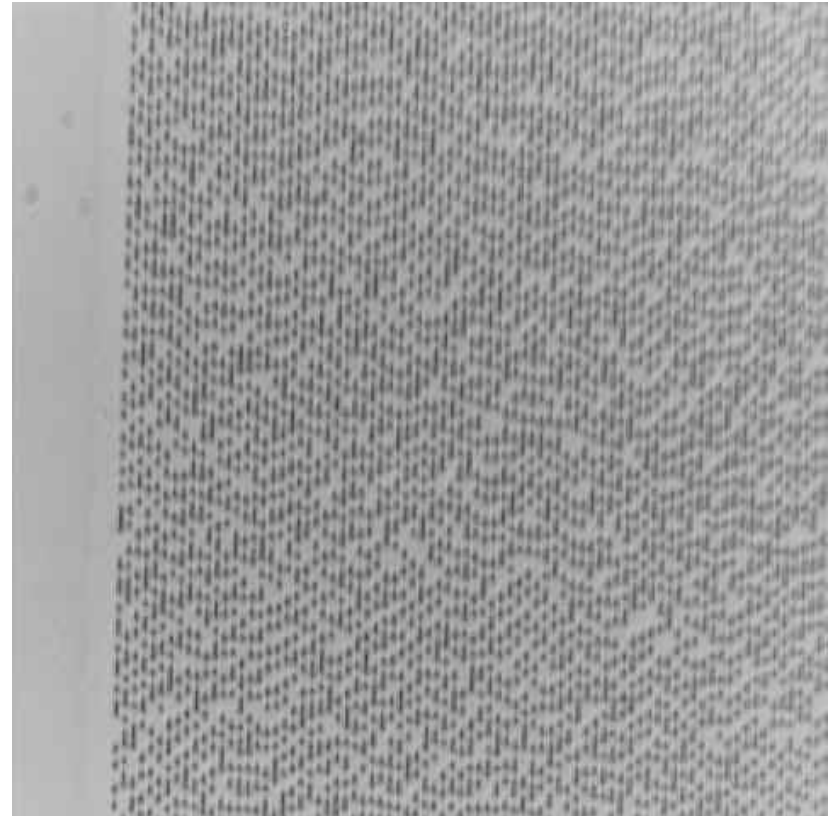
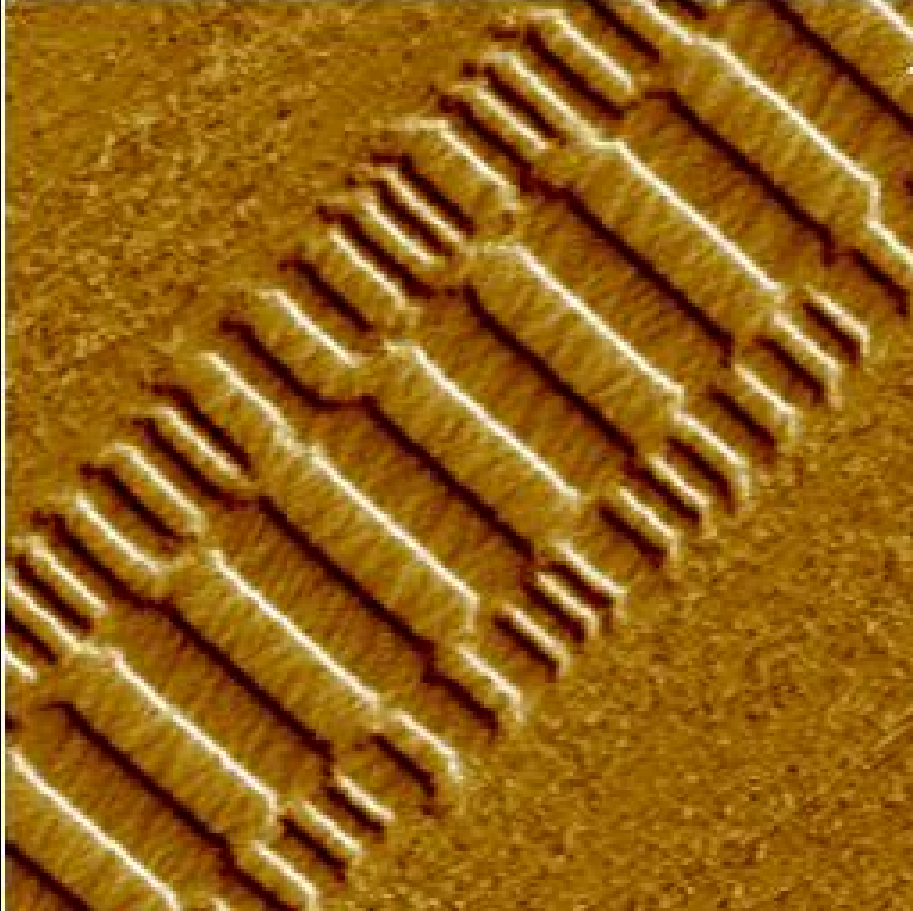
Using a high-power microscope and camera to take a picture of the surface of a hard drive or pits and lands on an optical disk

<http://www.pagetable.com/?p=32>

# Interaction Examples

## Level

### Aggregation of objects



### Bitstream on physical medium

Veeco Instruments.  
[http://www.veeco.com/library/nanotheater\\_detail.php?type=application&id=78&app\\_id=34](http://www.veeco.com/library/nanotheater_detail.php?type=application&id=78&app_id=34)

Compact Disk (CD). USByte. [http://www.usbyte.com/common/compact\\_disk\\_3.htm](http://www.usbyte.com/common/compact_disk_3.htm)



# Digital Forensics in Libraries, Archives and Museums (LAMs)

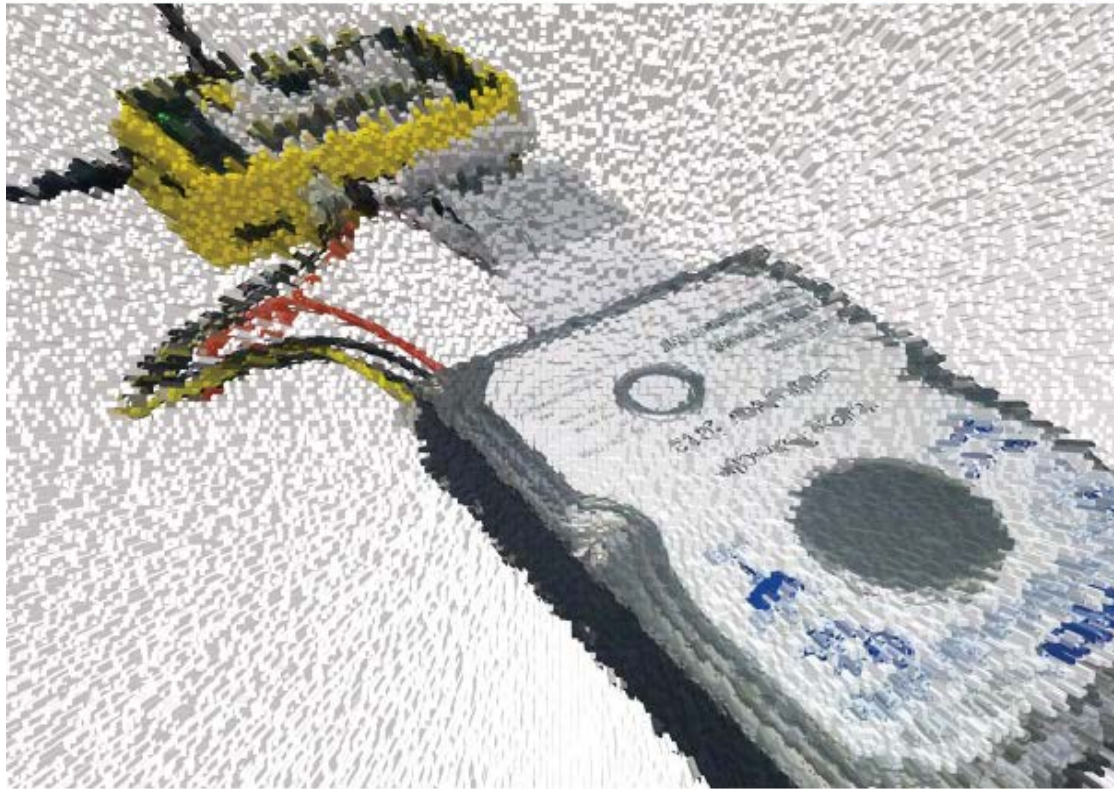
- In recent years, LAMs have been applying various digital forensics methods, for example:
  - use of write blockers
  - generation of disk images
  - applying cryptographic hashes to files
  - capture of Digital Forensics XML (DFXML)
  - scanning bitstreams for personally identifying information

# Need for Adaptation of Digital Forensics Tools and Tasks for LAMs

- Existing digital forensics tools provide valuable functionality, but they don't always fit well into primary LAM workflows.
- For example, LAMs 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

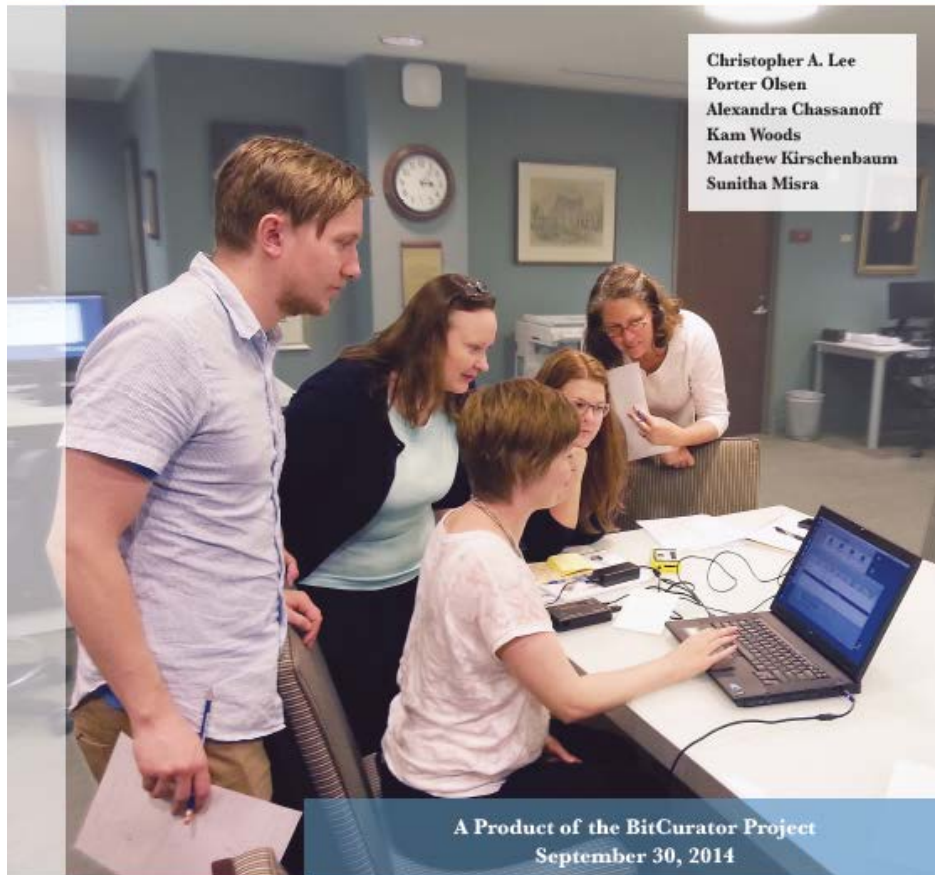
<http://www.bitcurator.net/docs/bitstreams-to-heritage.pdf>



# From Code to Community:

## Building and Sustaining BitCurator through Community Engagement

Christopher A. Lee  
Porter Olsen  
Alexandra Chassanoff  
Kam Woods  
Matthew Kirschenbaum  
Sunitha Misra



A Product of the BitCurator Project  
September 30, 2014

<http://www.bitcurator.net/wp-content/uploads/2014/11/code-to-community.pdf>



# 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)**

# Core BitCurator Team

**Cal Lee, PI**

**Matt Kirschenbaum, Co-PI**

**Kam Woods, Technical Lead**

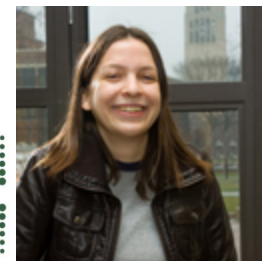
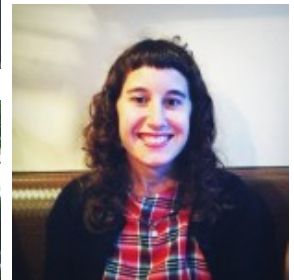
**Porter Olsen, Community Lead**

**Alex Chassanoff, Project  
Manager**

**Sunitha Misra, Software  
Developer (UNC)**

**Kyle Bickoff, GA (MITH)**

**Amanda Visconti, GA (MITH)**



**CoSA**

**State Electronic  
Records Initiative**

COUNCIL OF STATE ARCHIVISTS



# Two Groups of Advisors

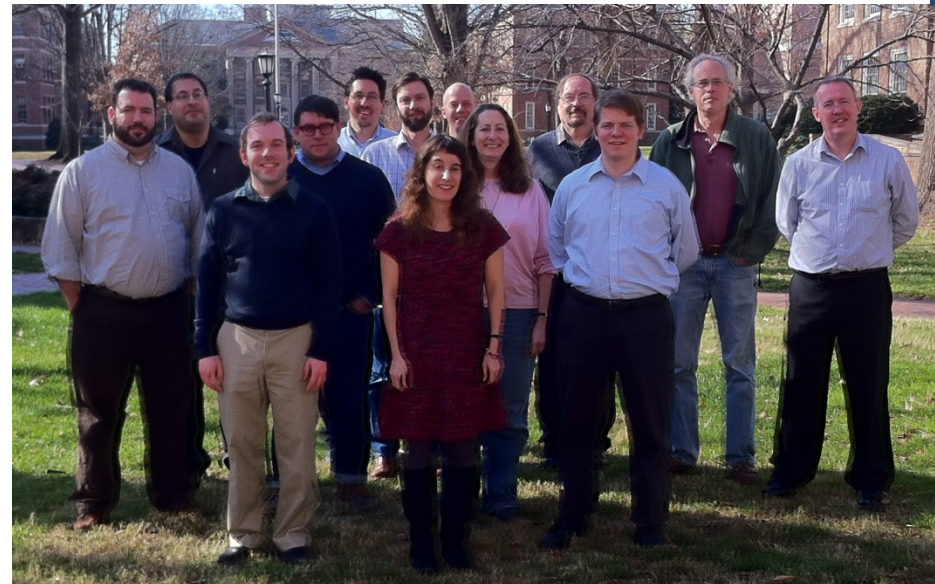
## Professional Experts Panel

- Bradley Daigle, University of Virginia Library
- Erika Farr, Emory University
- Jennie Levine Knies, University of Maryland
- Jeremy Leighton John, British Library
- Leslie Johnston, US National Archives and Records Administration
- Naomi Nelson, Duke University
- Erin O'Meara, Gates Archive
- Michael Olson, Stanford University Libraries
- Gabriela Redwine, Beinecke, Yale University
- Susan Thomas, Bodleian Library, University of Oxford



## Development Advisory Group

- Barbara Guttman, National Institute of Standards and Technology
- Jerome McDonough, University of Illinois
- Mark Matienzo, Digital Public Library of America
- Courtney Mumma, Artefactual Systems
- David Pearson, National Library of Australia
- Doug Reside, New York Public Library
- Seth Shaw, University Archives, Duke University
- William Underwood, Georgia Tech



# 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 LAM ingest and collection management environments
- Provision of public access to the data

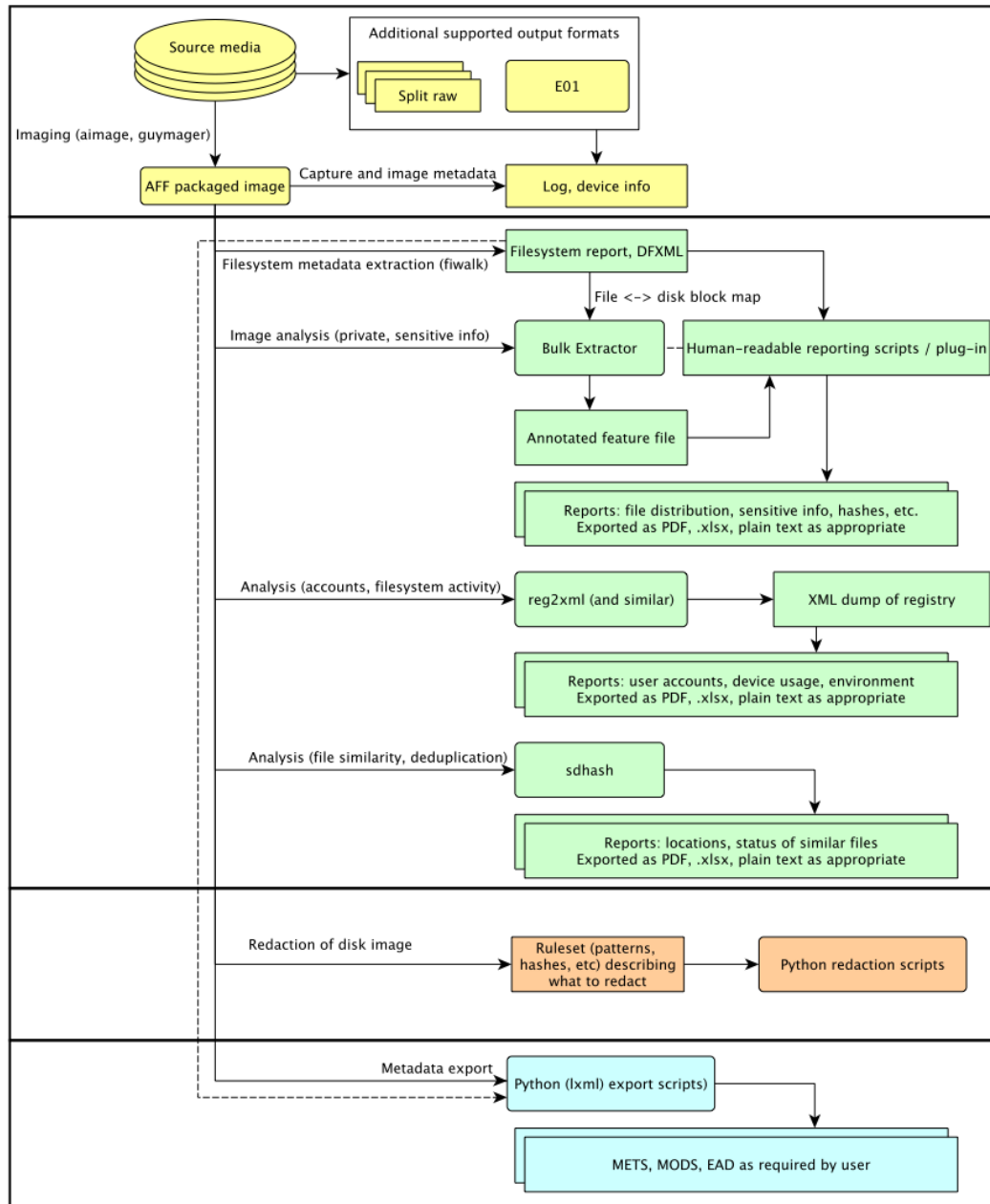
# BitCurator Environment

**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)
- Self-contained Linux environment in a virtual machine using e.g. Virtual Box or VMWare
- As individual components run directly in your own Linux environment or (whenever possible) Windows environment

# BitCurator-Supported Workflow



See: <http://bitcurator.net>

- **Acquisition**
- **Reporting**
- **Redaction**
- **Metadata Export**



INSTITUTE of  
Museum and Library  
SERVICES



**SERI**  
State Electronic  
Records Initiative  
COUNCIL OF STATE ARCHIVISTS



Computer



home



Imaging Tools



Forensics Tools



Additional Tools



Shared Folders and  
Media



Trash



Documentation and  
Help



Network Servers

# BitCurator

# BitCurator Consortium

Continuing home for hosting, stewardship and support of BitCurator (and BitCurator Access) tools and associated user engagement

Administrative home: Educopia Institute

Funding based on membership dues

Institutions as members, with two categories of membership: Charter and General

The most important member benefit is assurance that the BitCurator software will persist in future years

*<https://bitcuratorconsortium.org/>*







## A Growing Community

The BitCurator Consortium provides spaces for members to share documentation, develop their skills, and improve the BitCurator environment.

[Membership is open >](#)

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](#).

## 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](#) 

## How our members are using BitCurator

## 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  
University of

---

**So let's now look at specific curation actions and processes...**



# Write Blocking – One-Way Streets 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



# Dedicated Hardware Write Blockers



UltraBay II (legacy ports)

UltraBay 4D (contemporary ports)

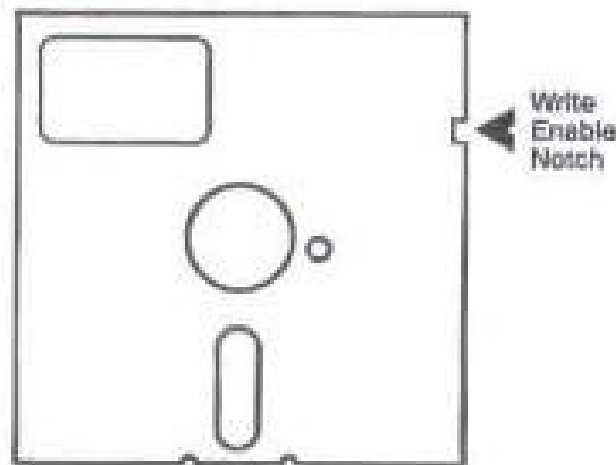


## 5.25 Inch Floppy – If light can get through, it's not write protected



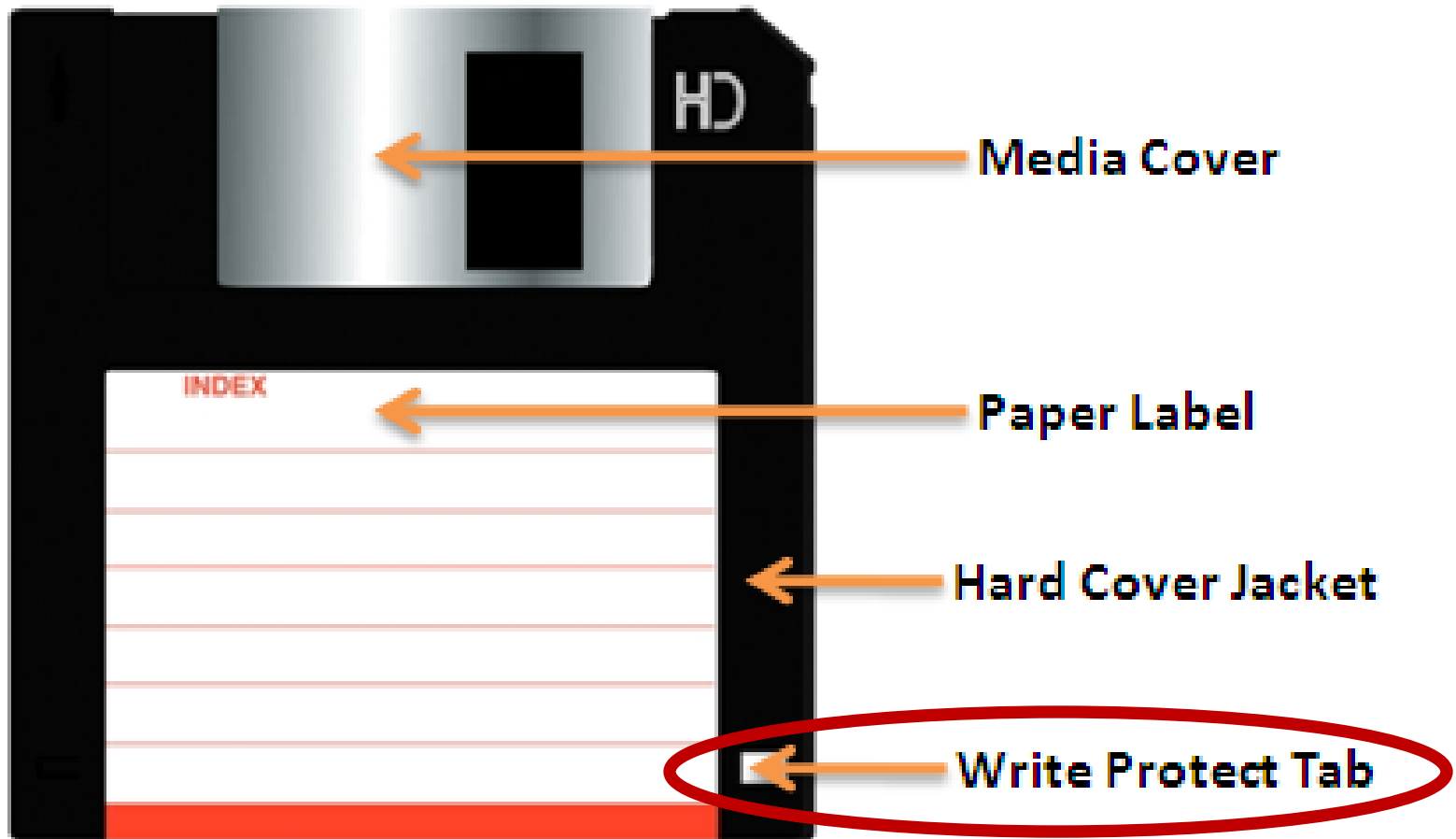
These black Write Protect Tabs should be used when you do not want stored information on your disk to be changed or lost accidentally.

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



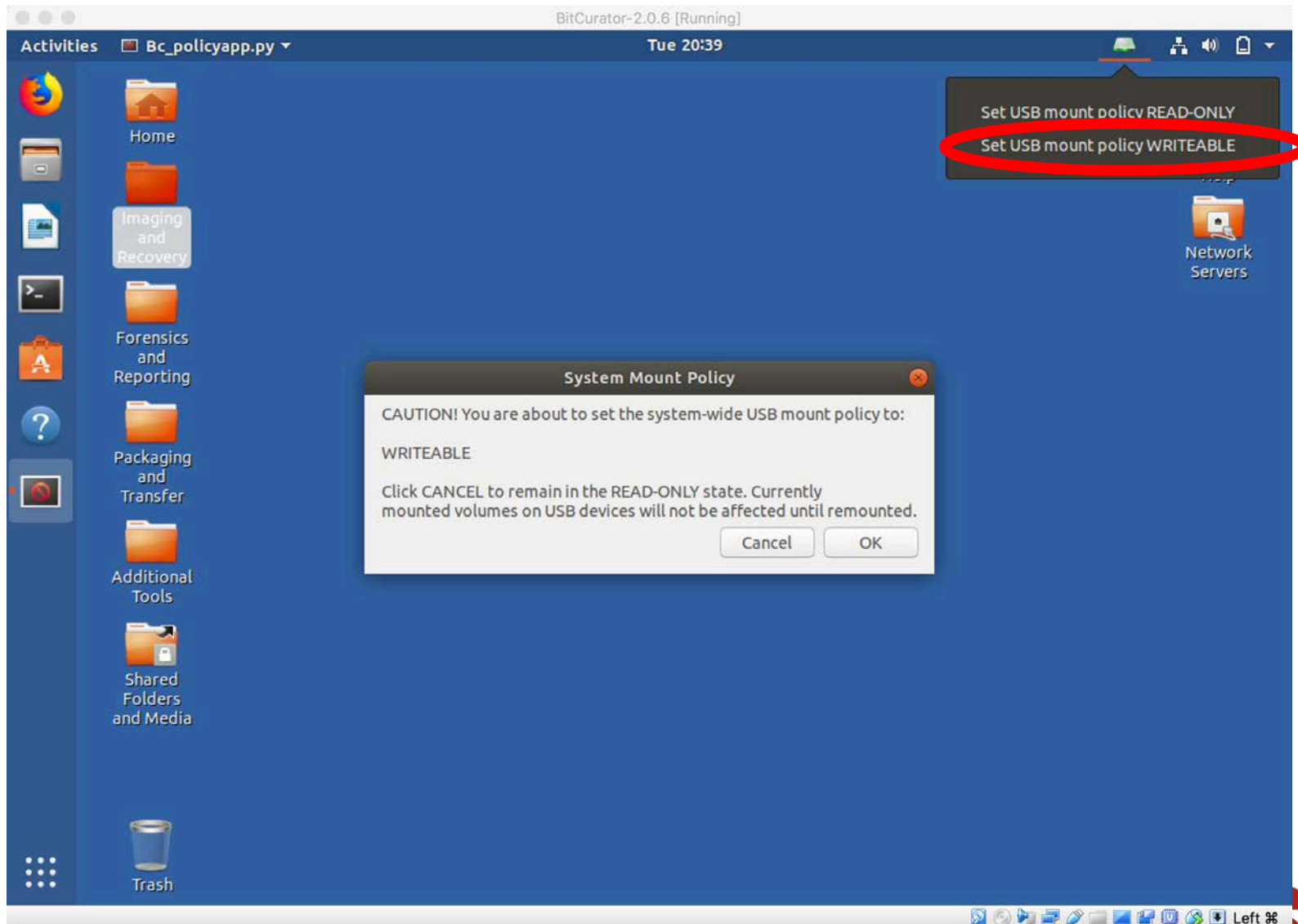
[http://en.wikipedia.org/wiki/File:Floppy\\_tabs\\_3x2.jpg](http://en.wikipedia.org/wiki/File:Floppy_tabs_3x2.jpg)

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



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

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



# Getting below the File System – Low-Level Copying

**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

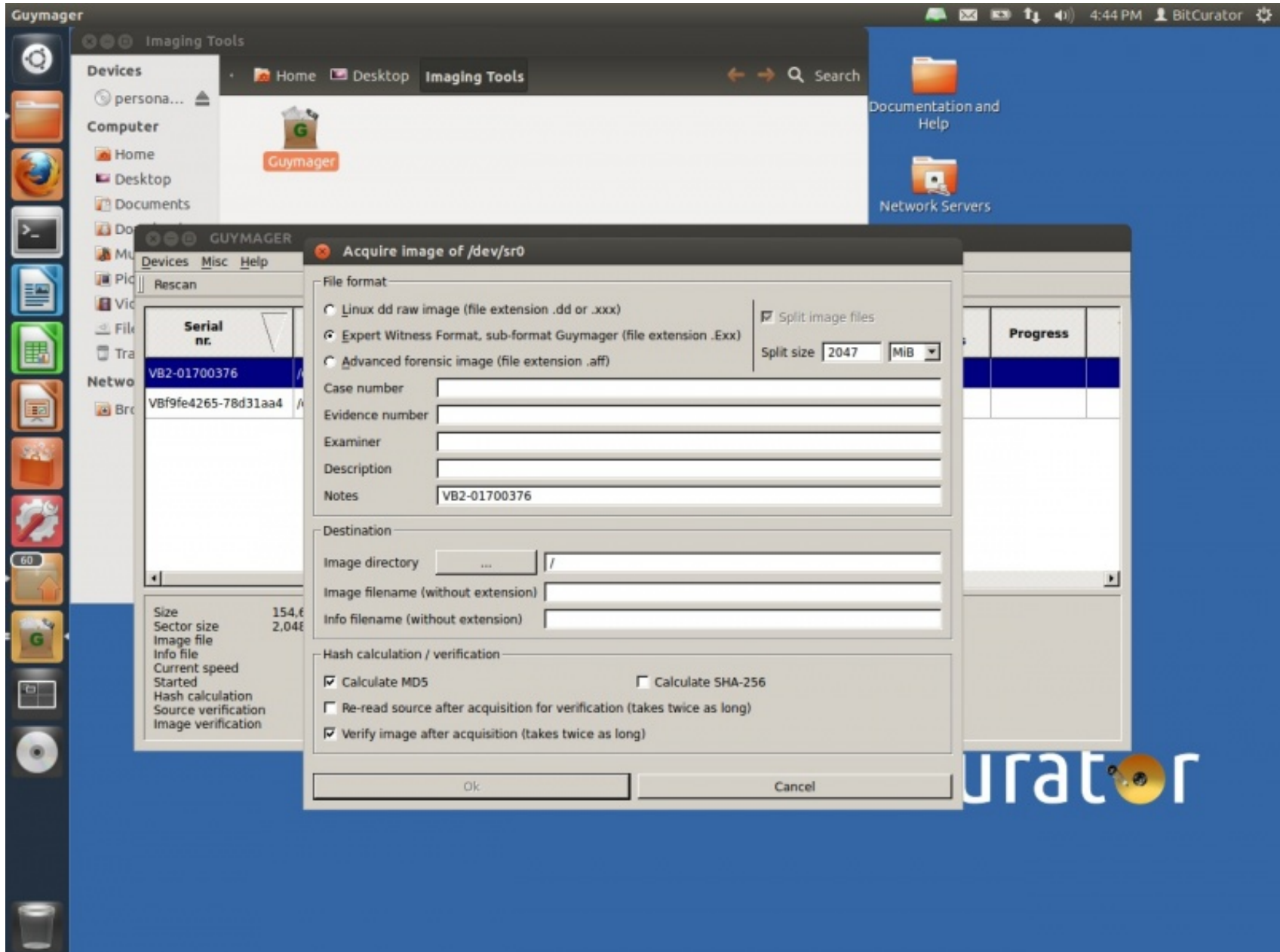
**Most commonly used tool is dd (or variant) - UNIX program for low-level copying and conversion of data from a storage device**

**More specialized tools for creating forensic images include:**

- FTK Imager
- Guymager
- Imaging utilities in commercial applications (e.g. EnCase)




# Main Acquisition Interface for Guymager



 Desktop


Documents

 Downloads Music Pictures Videos

 File System

 Trash

## Network



## Additional Tools

Documentation and Help

## Forensics

[Show AFF Info](#)[Show E01 Info](#)

nps-2010-emails.  
E01

### Disk Image Info

## File Analysis

FileInfo

## Find Files

## Open Browser as Root

### Safe Mount

### Open Scripts Folder

Open

Open With Other Application...

## Scripts

Cut

Copy

Make Link

Rename...

Copy to

Move to

Move to Trash

[Revert to Previous Version...](#)

Compress...

Send To...

## Properties





## Computer

Home

Desktop

Documents

Downloads

Music

Pictures

Videos

File System

Trash

## Network

Browse Net...

Home

Desktop

Search

## EnCase Disk Image Info

ewfinfo 20130416

## Acquiry information

Acquisition date: Wed Jan 19 12:09:18 2011

System date: Wed Jan 19 12:09:18 2011

Operating system used: Linux

Software version used: 20100226

Password: N/A

## EWF information

File format: EnCase 6

Sectors per chunk: 64

Error granularity: 64

Compression method: deflate

Compression level: best compression

Set identifier: 4eb6701d-6cf0-2f4a-a0c6-0cb5d5e20959

## Media information

Media type: fixed disk

Is physical: yes

Bytes per sector: 512

Number of sectors: 2068480

Media size: 1010 MiB (1059061760 bytes)

## Digest hash information

MD5: 9c0de6c8532d7a66ddcf01861dfb6535

Cancel

OK

charlie-work-usb-  
2009-12-11.E01

# Four Ways to Interact with Disk Images

## Emulation

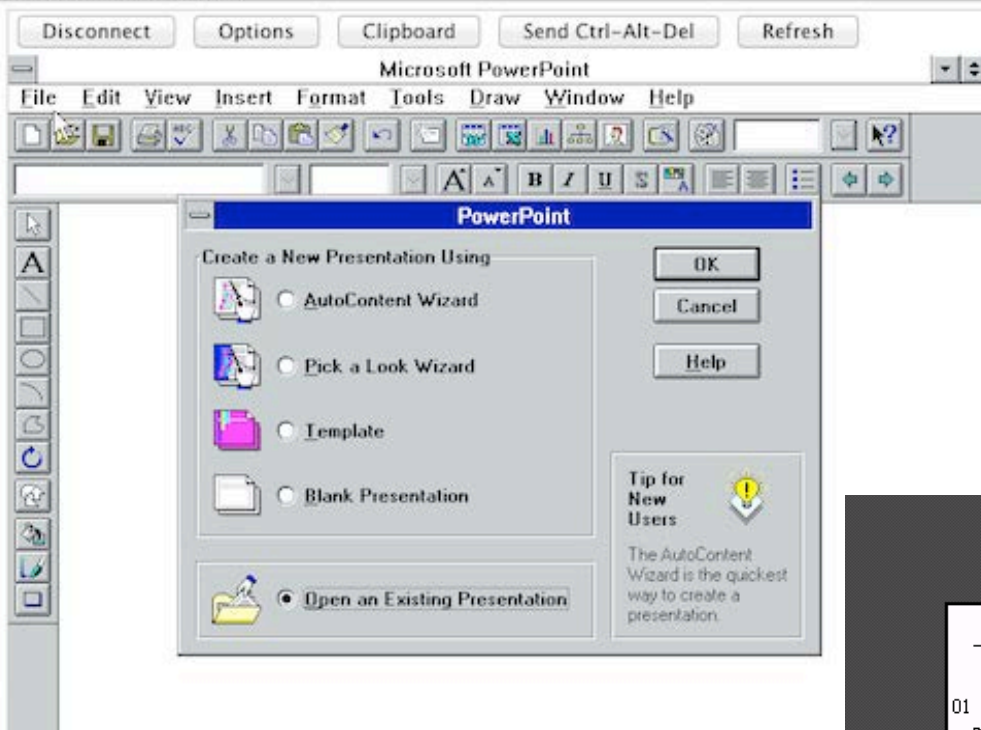
### Mount them like regular drives:

- For ISO images - disk utilities in Mac OS or Windows 8/10
- For forensically packaged disk images: ewfmount, OSFMount, BitCurator (mounting scripts built into the environment)

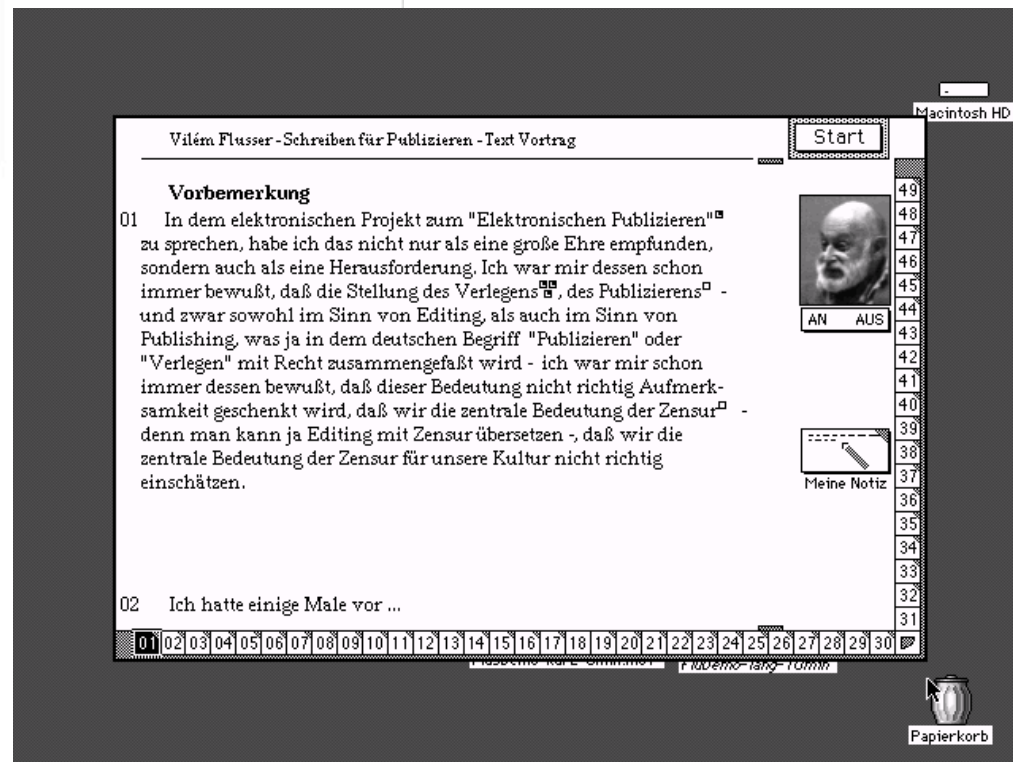
### Inspect them as forensic objects

- FTK Imager
- Autopsy
- BitCurator (Disk Image Access tool)

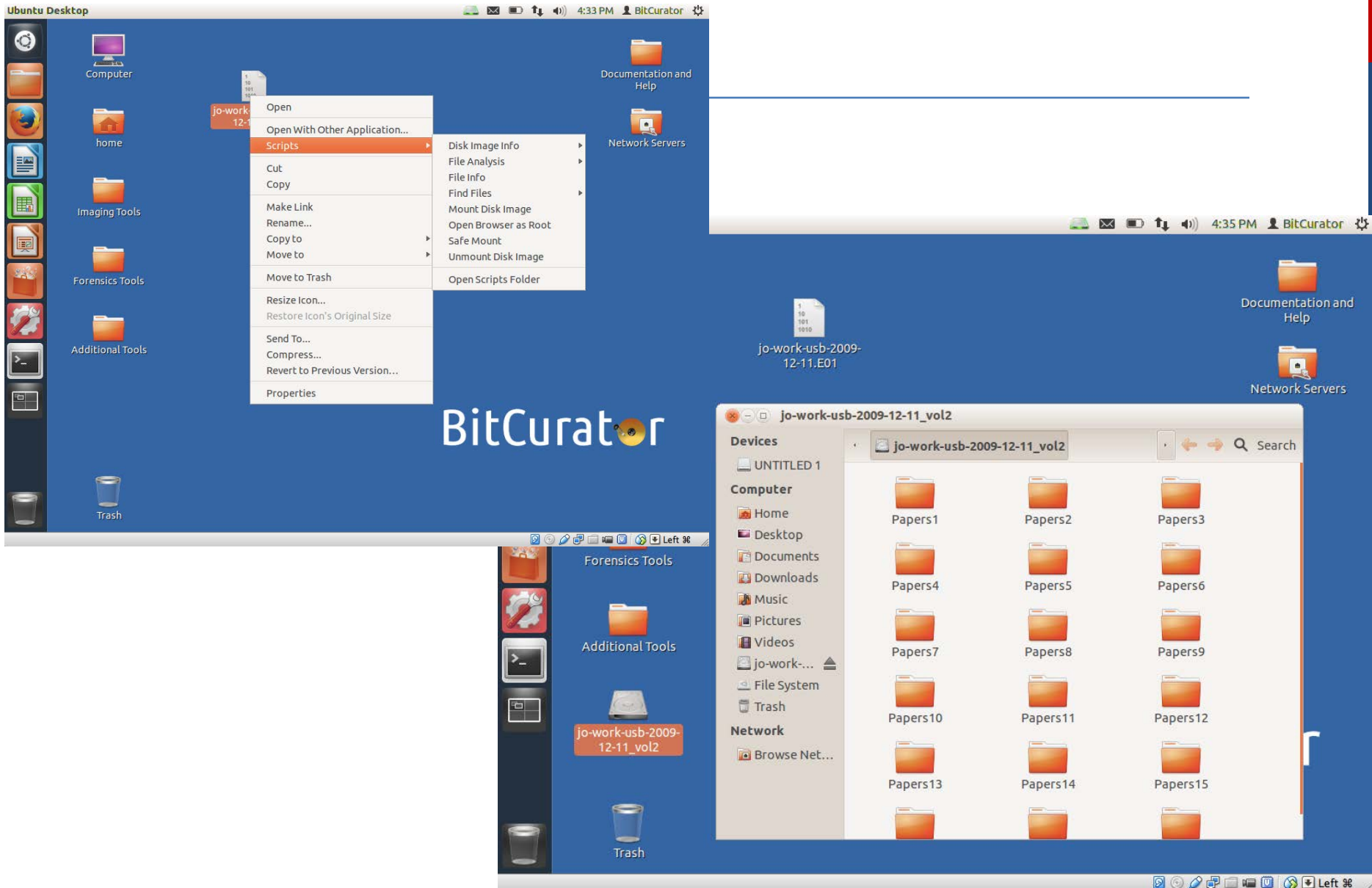
### Dynamically navigate them from within a web browser (BitCurator Access Webtools)



# Emulation as a Service



# Mounting a Forensically Packaged Disk Image in the BitCurator Environment



# Exporting Files from a Disk Image

The screenshot displays a Linux desktop environment with a blue background. A window titled "Forensics Tools" is open, showing a list of applications. The "BitCurator Disk Image Access" icon is circled in red. Other icons include "BitCurator Mounter", "BitCurator Reporting Tool", "Bulk Extractor Viewer", "FSint", "pyExifToolGUI", "SDHash (file similarity)", and "ssdeep (fuzzy hashes)".

In the foreground, the "Disk Image Access Interface" window is open. It features a toolbar with buttons: "Open disk image", "Close disk image", "Select All", "DeSelect All", "Export selections", and "Cancel export".

The "File System:" section contains a list of files and directories. Entries in bold are directories, and entries in red are unallocated/deleted files. The list includes:

- ☐ .urlspersona.txt
- ☐ urlspatents.txt
- ☐ .urlspatents.txt
- ☐ urlscryptography.txt
- ☐ .urlscryptography.txt
- ☐ urlscopyright.txt
- ☐ .urlscopyright.txt
- ☐ patentauto.py
- ☐ .patentauto.py
- ☐ patentterms.txt
- ☐ .patentterms.txt
- ☐ **\_54402.EXE**
- ☐ R54402.EXE
- ☐ **\_461531\_**
- ☐ **\_604468\_**
- ☐ **\_078421\_**
- ☐ **\_452781\_**
- ☐ **\_189812\_**
- ☐ **xpadvancedkeylogger.exe**
- ☒ **Log**
- ☐ vnc-4\_1\_3-x86\_win32.exe
- ☐ \$MBR
- ☐ \$FAT1

The "Image Info" section displays the following information:

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

The "Messages" section shows a log of actions:

```
/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 ...
```

BitCurator Access Webtools | x

dogwood.ils.unc.edu:8080

Home Images Status

Search text...

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

<b>All Images</b> All images included recursively.	Images: 12
<b>ISO test</b> Set of ISO test disk images.	Images: 2
<b>Mixed test</b> Set of mixed-format test disk images.	Images: 10

<https://github.com/BitCurator/bitcurator-access-webtools>



BitCurator Access Webtools | f x

Kam

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

Kevin Fairbanks, Johns Hopkins Applied Physics Lab

http://simson.net/

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

Kevin Fairbanks, Johns Hopkins Applied Physics Lab

http://simson.net/



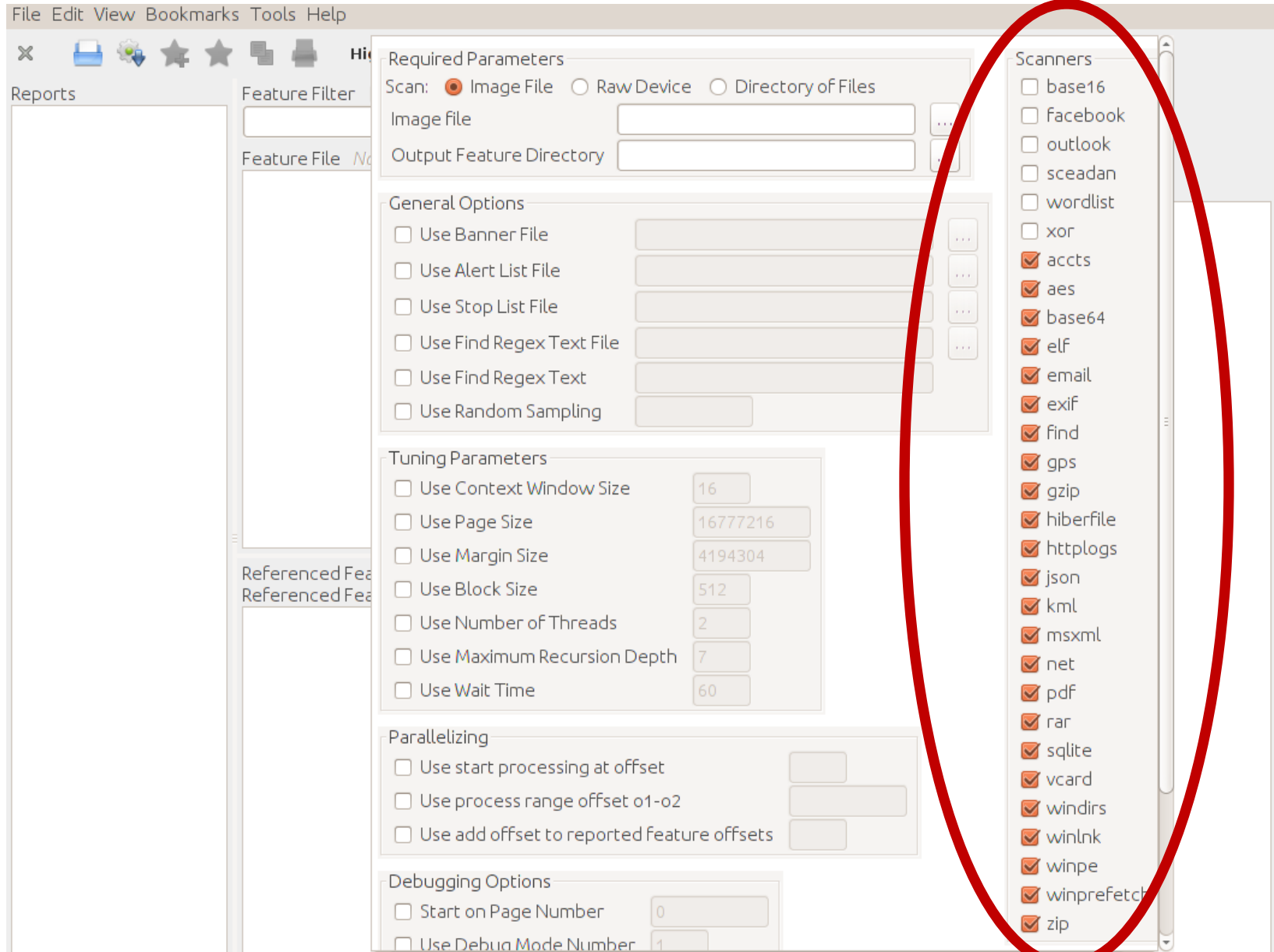
INSTITUTE of  
Museum and Library  
SERVICES



CoSA

**SERI**  
State Electronic  
Records Initiative  
COUNCIL OF STATE ARCHIVISTS

# Identifying Potentially Sensitive Data using Bulk Extractor - Scanning Options



The screenshot shows the Bulk Extractor application window. The 'Scanners' list on the right is circled in red. The 'Required Parameters' section shows 'Scan' set to 'Image File'. The 'General Options' section has several checkboxes for file types. The 'Tuning Parameters' section has input fields for various settings. The 'Parallelizing' section has checkboxes for processing options. The 'Debugging Options' section has checkboxes for debugging settings.

File Edit View Bookmarks Tools Help

Reports

Feature Filter

Feature File No

Referenced Feature

Referenced Feature

Required Parameters

Scan: ☒ Image File ☐ Raw Device ☐ Directory of Files

Image file

Output Feature Directory

General Options

☐ Use Banner File

☐ Use Alert List File

☐ Use Stop List File

☐ Use Find Regex Text File

☐ Use Find Regex Text

☐ Use Random Sampling

Tuning Parameters

☐ Use Context Window Size 16

☐ Use Page Size 16777216

☐ Use Margin Size 4194304

☐ Use Block Size 512

☐ Use Number of Threads 2

☐ Use Maximum Recursion Depth 7

☐ Use Wait Time 60

Parallelizing

☐ Use start processing at offset

☐ Use process range offset o1-o2

☐ Use add offset to reported feature offsets

Debugging Options

☐ Start on Page Number 0

☐ Use Debug Mode Number 1

Scanners

- ☐ base16
- ☐ facebook
- ☐ outlook
- ☐ sceanan
- ☐ wordlist
- ☐ xor
- ☒ accts
- ☒ aes
- ☒ base64
- ☒ elf
- ☒ email
- ☒ exif
- ☒ find
- ☒ gps
- ☒ gzip
- ☒ hiberfile
- ☒ httplogs
- ☒ json
- ☒ kml
- ☒ msxml
- ☒ net
- ☒ pdf
- ☒ rar
- ☒ sqlite
- ☒ vcard
- ☒ windirs
- ☒ winlnk
- ☒ winpe
- ☒ winprefetch
- ☒ zip

See: [http://www.forensicswiki.org/wiki/Bulk\\_extractor](http://www.forensicswiki.org/wiki/Bulk_extractor)

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

BitCurator-0.2.0 [Running]

Bulk Extractor Viewer

File Edit View Tools Help

Highlight:  ☒ Match case

Reports

- beoutput
- domain.txt
- domain\_histogram.txt
- email.txt
- email\_histogram.txt**
- ether.txt
- ether\_histogram.txt
- json.txt
- packets.pcap
- rfc822.txt
- tcp.txt
- tcp\_histogram.txt
- url.txt
- url\_histogram.txt
- url\_services.txt
- windirs.txt
- winpe.txt

Feature Filter ☐ Match case

Navigation

sampleimage.E01, 42273785, privacy@Motorola.com

Image File sampleimage.E01

Feature File email.txt

Feature Path 42273785

Feature privacy@Motorola.com

Image

42271936 your credit card number, so this information can only be viewed  
42272000 by Motorola. .Motorola uses Secure Sockets Layer (SSL) encrypti  
42272064 on technology, the highest level of security on the Internet. Th  
42272128 e SSL protocol provides server authentication, data integrity, a  
42272192 nd privacy on the Web. This security measure helps ensure that n  
42272256 o impostors, eavesdroppers, or vandals get your personal informa  
42272320 tion. SSL not only encrypts your personal and financial informat  
42272384 ion transmitted, including credit card information, but also ver  
42272448 ifies the identity of the server and that the original message a  
42272512 rrives safely at its destination. .However, no data transmission  
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  
42272704 ola cannot ensure or warrant the security of any information you  
42272768 transmit to us or from our Web site, and therefore you use our  
42272832 site at your own risk. Once we receive your transmission, we use  
42272896 our best effort to ensure its security on our systems. .000200  
42272960 0007AE000038B6.7A8.As a global company Motorola has internation  
42273024 l sites and users all over the world. When you give Motorola per  
42273088 sonal information, that information may be sent electronically t  
42273152 o servers outside of the country where you originally entered the  
42273216 e information. In addition, that information may be used, stored  
42273280 and processed outside of the country where you entered that inf  
42273344 ormation. Whenever Motorola handles personal information, regard  
42273408 less of where this occurs, it takes steps to ensure that your in  
42273472 formation is treated securely and in accordance with the relevan  
42273536 t Terms of Use and this Privacy Policy. .How can I correct or ch  
42273600 ange my personal information? .If you would like to review, corr  
42273664 ect or change any personal information you have provided, or rem  
42273728 ove your name from our mailing list, please e-mail us at **privacy**  
42273792 **@Motorola.com**. If you have established a "user profile" on a Mot  
42273856 orola website, you may change the information you provided at an

Histogram File email\_histog...

n=12	privacy@motorola.com
n=3	0mj5nj@0itgx.ib.dj
n=3	73t@fo.pa
n=3	john@humaniz.com
n=3	newton@planetb.fr
n=3	sales@integrationnew
n=1	5kda_c@kqahw.sl
n=1	dqf@40mt.ro
n=1	fodfv@nwa4.ck
n=1	imki@73yt.lr
n=1	jcnmq@17.pn
n=1	kjph@sj.gr
n=1	nq9@5c7k.sg
n=1	pdcnfb@tft.ao
n=1	qyf@j65.de
n=1	tw+4vsa@xf.ms

Referenced Feature File e...

Referenced Feature pri...

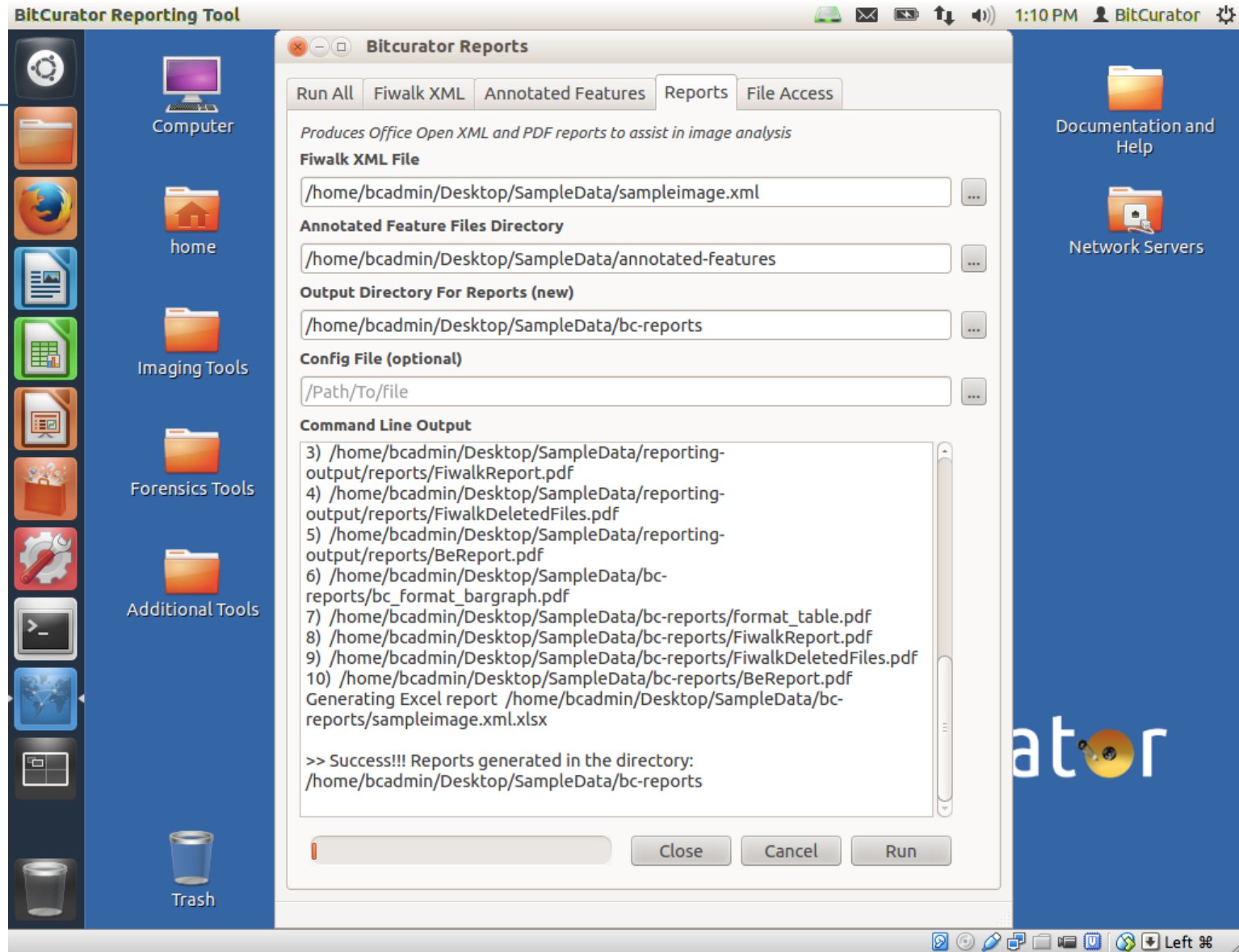
34804080	privacy@Motor
34807246	privacy@Motor
34808676	privacy@Motor
42271602	privacy@Motor
42273785	privacy@Motor
42274743	privacy@Motor
42347307	privacy@Motor
42349490	privacy@Motor
42350448	privacy@Motor
74735841	privacy@Motor
74738019	privacy@Motor
74738989	privacy@Motor

Text Hex

Left 98

Electronic Initiative

# Generating BitCurator Reports



- Provenance metadata - about the disk capture process
- Technical metadata - about the specific storage partition(s) on the disk

BitCurator-0.2.0 [Running]

Mozilla Firefox

file:///home/b...mpleimage.xml

file:///home/bcadmin/Desktop/SampleData/sampleimage.xml

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>
- <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>
- <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>
  <block_count>36839</block_count>

```

Left

ERI  
Electronic  
ds Initiative  
COUNCIL OF STATE ARCHIVISTS

# Operationalizing Original Order - Filesystem Metadata Output from fiwalk\*

```
-<fileobject>
  -<parent_object>
    <inode>102</inode>
  </parent_object>
  <filename>Papers8/37638.BrannyPhyle.Joseph+Moore.pdf</filename>
  <partition>1</partition>
  <id>901</id>
  <name_type>r</name_type>
  <filesize>100857</filesize>
  <alloc>1</alloc>
  <used>1</used>
  <inode>6783</inode>
  <meta_type>1</meta_type>
  <mode>511</mode>
  <nlink>1</nlink>
  <uid>0</uid>
  <gid>0</gid>
  <mtime prec="2">2009-11-17T19:35:10</mtime>
  <atime prec="86400">2009-12-10T05:00:00</atime>
  <ctime prec="2">2009-12-10T19:34:11</ctime>
  <libmagic>PDF document, version 1.4 </libmagic>
  -<byte_runs>
    <byte_run file_offset="0" fs_offset="56621568" img_offset="56653824" len="100857"/>
  </byte_runs>
  <hashdigest type="md5">eb60256dabffa67cef7211bcba659815</hashdigest>
  <hashdigest type="sha1">e56f606877f10daf91dc0304ea120b35452bd36e</hashdigest>
</fileobject>
```

\*Developed by Simson Garfinkel



## XML Schema for Digital Forensics XML

43 commits

1 branch

9 releases

1 contributor

 branch: master **dfxml\_schema** / +

Document an XML validation step ...

**ajnelson** authored on Dec 4, 2014 latest commit 4c8aab566e **README.md**[Code](#)[Issues](#)

8

[Pull requests](#)

0

[Pulse](#)[Graphs](#)

HTTPS clone URL

<https://github.com/>You can clone with [HTTPS](#) or [Subversion](#).[Clone in Desktop](#)[Download ZIP](#)

[https://github.com/dfxml-working-group/dfxml\\_schema](https://github.com/dfxml-working-group/dfxml_schema)



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

```
premis.xml (~/.Desktop/demo1/demo1reports/reports) - gedit
Open Save Undo Redo Find
premis.xml x
<?xml version="1.0" encoding="UTF-8"?>
<premis xmlns="info:lc/xmlns/premis-v2" version="2.0" xsi="http://www.w3c.org/2001/XMLSchema-instance">
  <object>
    <objectIdentifier>
      <objectIdentifierType>0d4e30d6-b8dc-11e3-a80f-080027f8dfea</objectIdentifierType>
      <objectIdentifierValue>/home/bcadmin/Desktop/terry-work-usb-2009-12-11.E01</objectIdentifierValue>
    </objectIdentifier>
  </object>
  <event>
    <eventIdentifier>
      <eventIdentifierType>0d4ea1ce-b8dc-11e3-a80f-080027f8dfea</eventIdentifierType>
      <eventIdentifierValue>E01/home/bcadmin/Desktop/terry-work-usb-2009-12-11.E01</eventIdentifierValue>
    </eventIdentifier>
    <eventType>Capture</eventType>
    <eventDateTime>Wed Jan 19 12</eventDateTime>
    <eventOutcomeInformation>
      <eventOutcome>E01</eventOutcome>
      <eventOutcomeDetail>Version: 20100226
, Image size: 512</eventOutcomeDetail>
    </eventOutcomeInformation>
  </event>
  <event>
    <eventIdentifier>
      <eventIdentifierType>19882604-b8dc-11e3-93f0-080027f8dfea</eventIdentifierType>
      <eventIdentifierValue>bulk_extractor -o /home/bcadmin/Desktop/demo1 /home/bcadmin/Desktop/terry-
work-usb-2009-12-11.E01</eventIdentifierValue>
    </eventIdentifier>
    <eventType>Feature Stream Analysis</eventType>
    <eventDateTime>2014-03-31T13:49:59Z</eventDateTime>
    <eventOutcomeInformation>
      <eventOutcome>Bulk Extractor Output</eventOutcome>
      <eventOutcomeDetail>version: 1.4.4</eventOutcomeDetail>
    </eventOutcomeInformation>
  </event>
</premis>
```

XML Tab Width: 8 Ln 1, Col 1 INS ↑

# Various Specialized BitCurator Reports

BitCurator-Demo-0.3.4 [Running]

Document Viewer

format\_table.pdf

Previous Next 1 (1 of 1) Fit Page Width

Report: File System Statistics and Files BitCurator

File Format Table

Disk Image: sampleimage.E01

Format	Short Form	Files
data	dat_ata	31
news or mail, ASCII text, with CR/LF line terminators	new_ors	1
PCX ver. 2.5 image data	PCX_ata	1
PDF document, version 1.4	PDF_1-4	6
MS Windows icon resource - 21 icons, 3x, 4-colors	MS_ors	1
x86 boot sector, code offset 0x52, O...doors 1, dos < 4.0 BootSector (0x0)	x86_x0-	1
SysEx File - GreyMatter	Sys_ser	1
empty (Zip archive data, at least v1.0 to extract)	emp_ct-	2
TIFF image data, little-endian	TIFF_ian	2
ASCII text, with no line terminators (OpenDocument Text)	ASC_at-	1
JPEG image data, JFIF standard 1.01	JPE_01	4
PE32 executable (GUI) Intel x86_64, f..._InnoSetup self-extracting archive	PE3_je	1
JPEG image data, JFIF standard 1.01, ...25(x5C276(x5C332(x5C0115(x5C261"	JPE_61-	2
...	ASC_ors	40
summary info	Com_ifo	1
...	emp_pty	9
ata, at least v2.0 to extract)	ASC_ct-	1

bc\_format\_bargraph.pdf

Previous Next 1 (1 of 1) Fit Page Width

Thumbnails

Disk Image: sampleimage.E01 File counts (by format)

Format	Count
ASC_ors	31
emp_pty	9
JPE_01	6
JPE_61	4
TIFF_ian	2
emp_ct-	2
ASC_at-	2
Com_ifo	1
PE3_je	1
ASC_ata	1
Sys_ser	1
x86_x0-	1
MS_ors	1
PCX_ata	1
new_ors	1

Page 1

Electronic Records Initiative

## Other Functionality to Meet Identified 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 Microsoft Outlook PST files	readpst
Examine embedded header information in images	pyExifToolGUI
Generate images of problematic disks or particular disk types (in 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	BagIt (Java) library, Bagger
File preview (left-click on file then hit space bar)	gnome-sushi

# Other Functionality to Meet Identified User Needs (Continued):

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	nsrlookup
View and edit bytestreams (hex editor)	Bless Hex Editor, GHex

# Incorporating digital forensics into LAM workflows



# 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.

<http://digitalcurationexchange.org/system/files/gengenbach-forensic-workflows-2012.pdf>

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

[http://www2.lib.virginia.edu/aims/whitepaper/AIMS\\_final.pdf](http://www2.lib.virginia.edu/aims/whitepaper/AIMS_final.pdf)

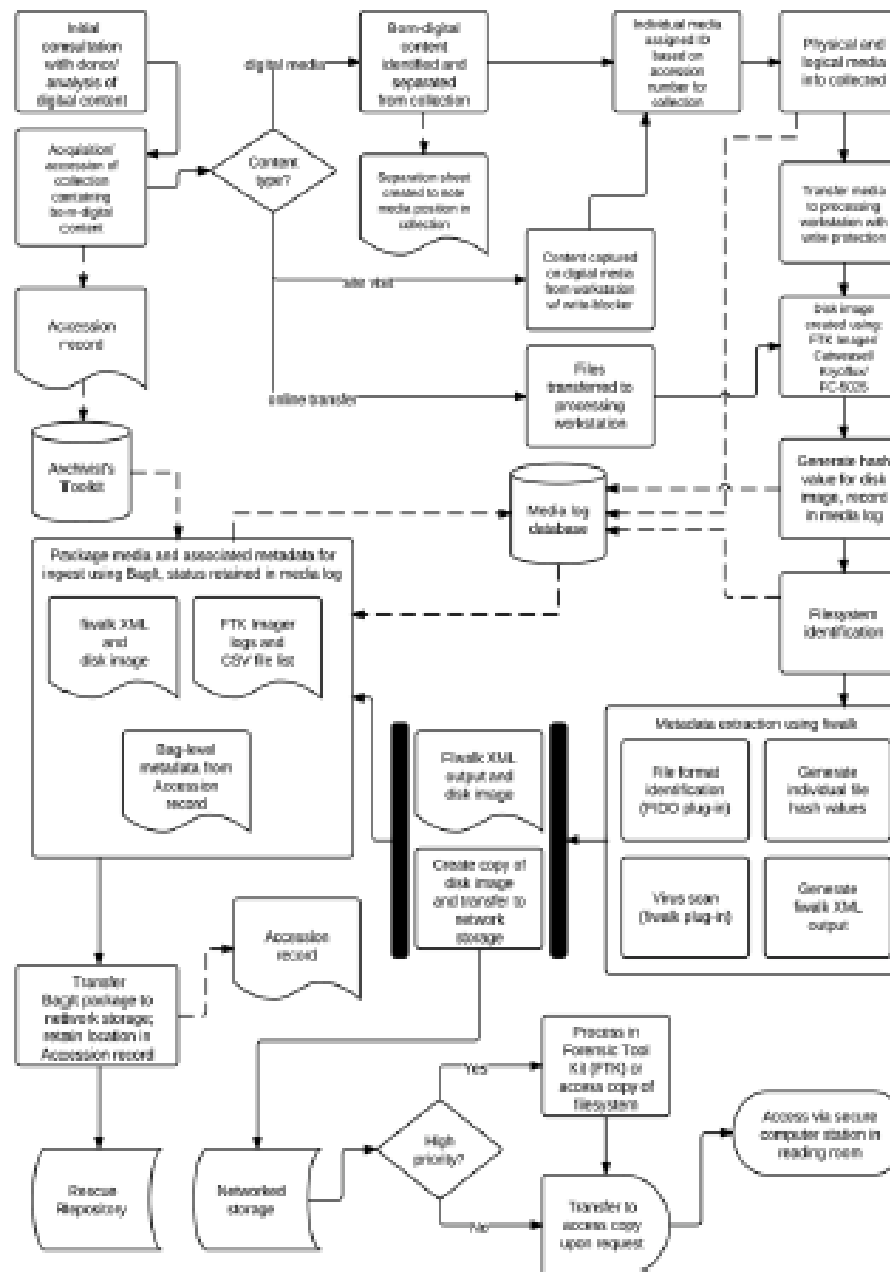
Digital Sustainability Lab – Massachusetts Institute of Technology

[http://www.dpworkshop.org/sites/default/files/DCM-Pipeline\\_28Apr2015.pdf](http://www.dpworkshop.org/sites/default/files/DCM-Pipeline_28Apr2015.pdf)

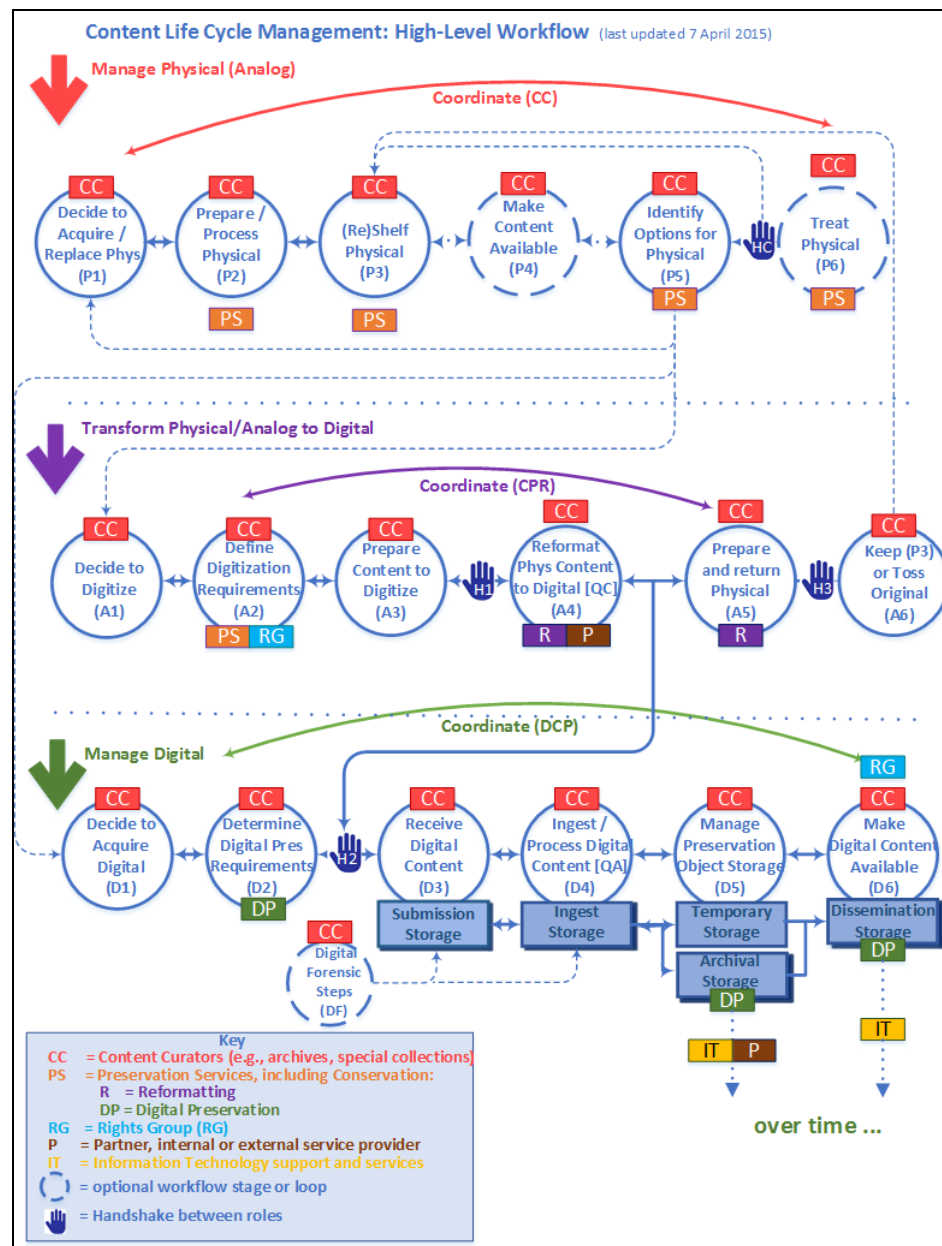
Workflows, BitCurator Consortium <https://bitcuratorconsortium.org/workflows>

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

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







Kari Smith, Massachusetts Institute of Technology.

[http://www.dpworkshop.org/sites/default/files/DCM-Pipeline\\_28Apr2015.pdf](http://www.dpworkshop.org/sites/default/files/DCM-Pipeline_28Apr2015.pdf)

## 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.

## Workflow

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.](#)

If you are interested in adding a workflow to our listing, please [contact us](#).

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

<https://bitcuratorconsortium.org/workflows>

## 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:

[OSSArcFlow proposal](#)

## 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>



INSTITUTE of  
Museum and Library  
SERVICES



**SERI**  
State Electronic  
Records Initiative  
COUNCIL OF STATE ARCHIVISTS



## As-Is Workflows (June 2018)

In the fall of 2017, the project team worked with partners at each institution to mockup a visual representation of their current workflow activities. Representing a “snapshot in time,” these documents show how a diverse group of institutions are using OSS tools in their workflows to curate born-digital content. They also provide an essential starting point for synthesizing and comparing both the gaps and overlaps that currently exist between common OSS tools and environments.

1. [Atlanta University Center, Robert W. Woodruff Library](#)
2. [District of Columbia Public Library](#)
3. [Duke University](#)
4. [Emory University](#)
5. [Kansas Historical Society](#)
6. [Massachusetts Institute of Technology](#)
7. [Mount Holyoke College](#)
8. [New York Public Library](#)
9. [Rice University](#)
10. [Stanford University](#)
11. [New York University](#)
12. [Odum Institute](#)

<https://educopia.org/research/ossarcflow>



INSTITUTE of  
Museum and Library  
SERVICES



**SERI**  
State Electronic  
Records Initiative  
COUNCIL OF STATE ARCHIVISTS

# As-Is Digital Curation Workflow, August 2017

Duke University

OssArcFlow Project

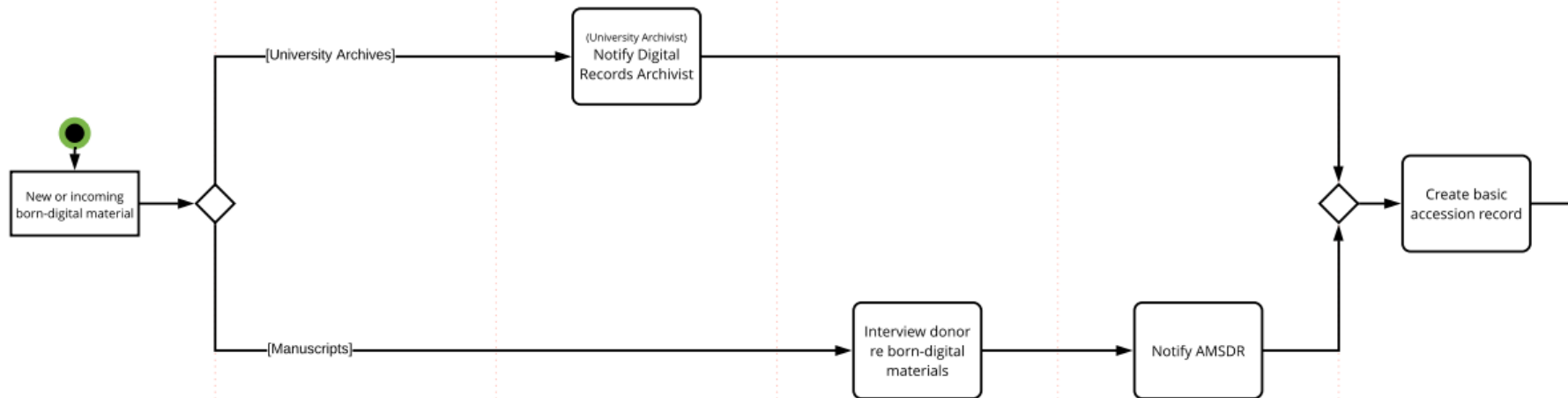
<https://educopia.org/research/ossarcflow>

Born-digital content

ArchivesSpace

archivematica.

BitCurator



Acquisition

Digital Records Archivist (DRA)

Archivist for Metadata, Systems, and Digital Records (AMSDR)

Processing archivist or curator

Other (role noted in parentheses)

Primary / preferred / required

Secondary / optional

Concurrency

Start / end point

Split / Merge

Action

Connection

Object / Data

Delay

Pain point

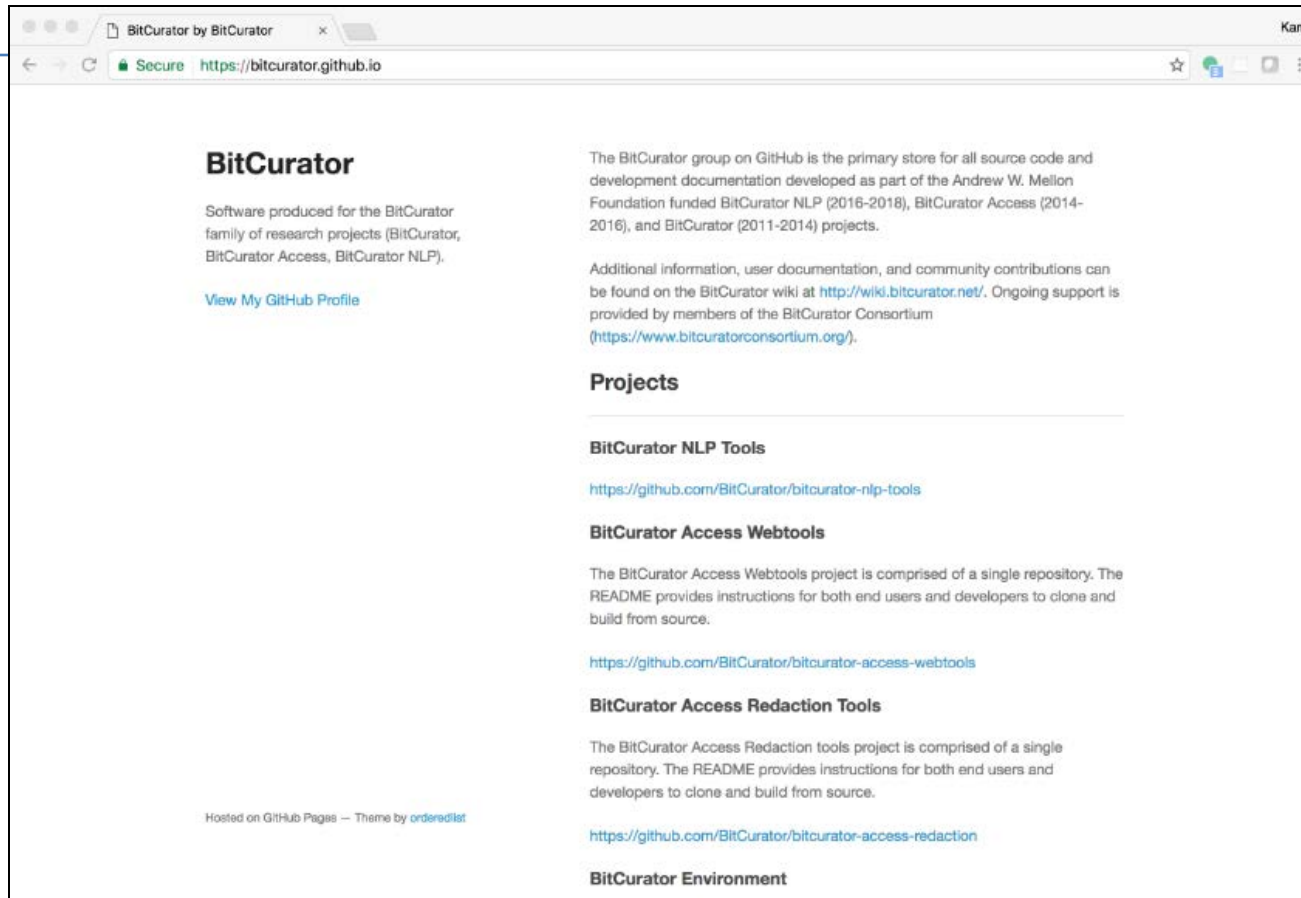


INSTITUTE of  
Museum and Library  
SERVICES



SERI  
State Electronic  
Records Initiative  
COUNCIL OF STATE ARCHIVISTS

# For Further Information



<https://bitcurator.github.io/>

# BitCurator

## Quick Start Guide

Last updated: August 1, 2018

Release(s): 2.0.4 and later



UNC  
SCHOOL OF INFORMATION  
AND LIBRARY SCIENCE

BitCurator  
CONSORTIUM

<https://github.com/BitCurator/bitcurator-distro/wiki/Releases>



INSTITUTE of  
Museum and Library  
SERVICES



**SERI**  
State Electronic  
Records Initiative  
COUNCIL OF STATE ARCHIVISTS

# WEBINAR EVALUATION

---

**We appreciate your feedback!**

**After you exit the webinar, you will automatically be taken to an online webinar evaluation.**

**Please take a couple of minutes to complete the survey and help us plan future webinars.**

# STAY CONNECTED & INFORMED

---

CoSA Website

<http://www.statearchivists.org>

CoSA Twitter Handle

@StateArchivists

CoSA Facebook Page

[www.facebook.com/CouncilOfStateArchivists](http://www.facebook.com/CouncilOfStateArchivists)

