



Enabling Digital Forensics Practices in Libraries, Archives and Museums: The BitCurator Experience

By

Christopher Lee and Kam Woods

Presented At

The Digital Forensic Research Conference

DFRWS 2014 USA Denver, CO (Aug 3rd - 6th)

DFRWS is dedicated to the sharing of knowledge and ideas about digital forensics research. Ever since it organized the first open workshop devoted to digital forensics in 2001, DFRWS continues to bring academics and practitioners together in an informal environment. As a non-profit, volunteer organization, DFRWS sponsors technical working groups, annual conferences and challenges to help drive the direction of research and development.

<http://dfrws.org>

Enabling Digital Forensics Practices in Libraries, Archives and Museums: The BitCurator Experience

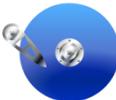
Cal Lee and Kam Woods

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

Digital Forensics Research Workshop

August 3-6, 2014

Denver, CO

BitCurator 



UNC
SCHOOL OF INFORMATION
AND LIBRARY SCIENCE

Goals of Libraries, Archives and Museums (LAMs) When Acquiring Materials

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

Fundamental Archival Principles

Provenance

- Reflect “life history” of records
- Records from a common origin or source should be managed together as an aggregate unit

Original Order

Organize and manage records in ways that reflect their arrangement within the creation/use environment

Chain of Custody

- “Succession of offices or persons who have held materials from the moment they were created”¹
- Ideal recordkeeping system would provide “an unblemished line of responsible custody”²

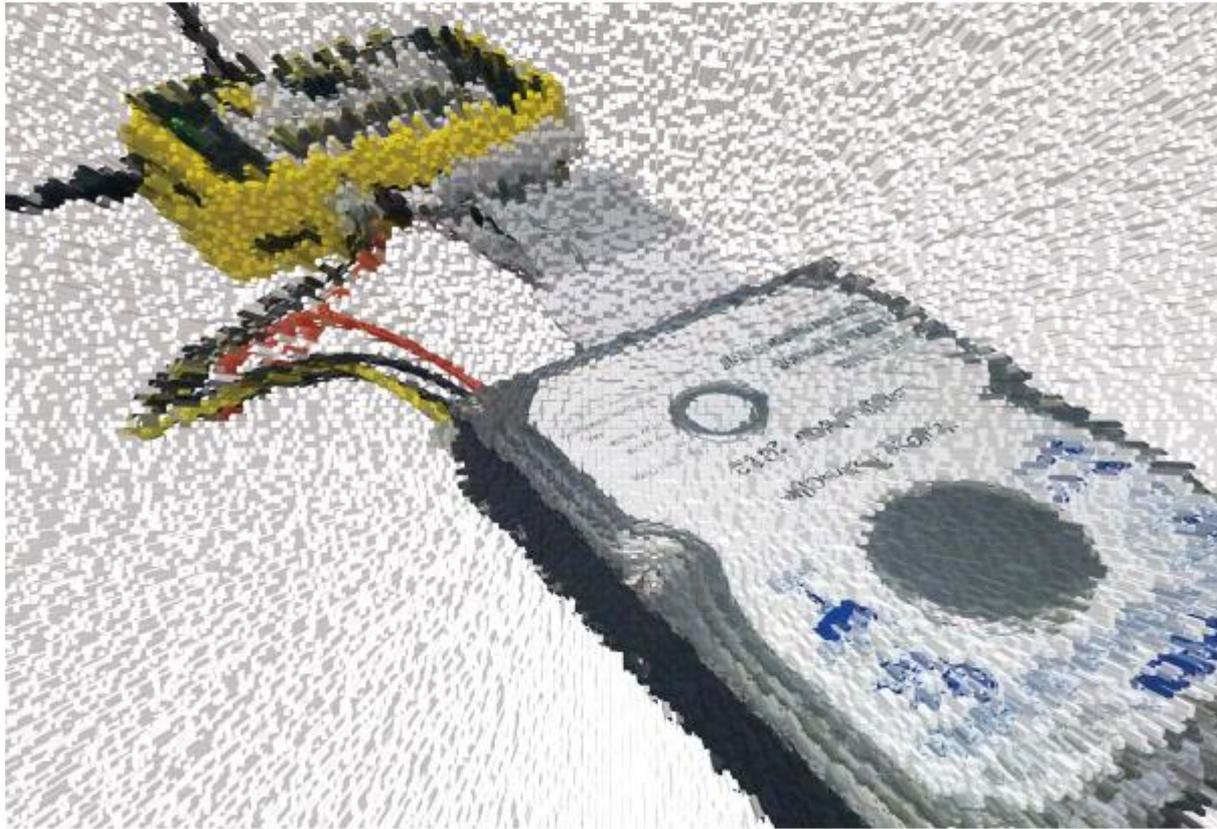
1. Pearce-Moses, Richard. *A Glossary of Archival and Records Terminology*. Chicago, IL: Society of American Archivists, 2005.
2. Hilary Jenkinson, *A Manual of Archive Administration: Including the Problems of War Archives and Archive Making* (Oxford: Clarendon Press, 1922), 11.

Digital Forensics Can Help to Fulfill Archival Principles

- Provenance
 - Identify, extract and save essential information about context of creation
- Original Order
 - Reflect original folder structures, files associations, related applications and user accounts
- Chain of Custody
 - Documentation of how records were acquired and any transformations to them
 - Use well-established hardware and software mechanisms to ensure that data haven't been changed inadvertently
- Identifying Sensitive Information
 - Identify personally identifying information, regardless of where it appears
 - Flag for removal, redaction, closure or restriction

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>

Digital Forensics Lab @ UNC School of Information and Library Science



Digital Forensics in LAMs

- In recent years, LAMs have discovered the value of 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 Tools and Tasks for LAM Users

- While existing digital forensics tools provide valuable functionality, they don't always fit well into primary workflows of LAMs.
- 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)

BitCurator

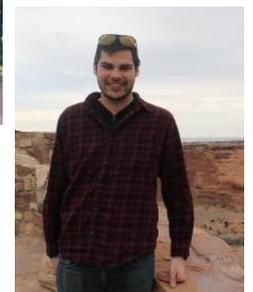
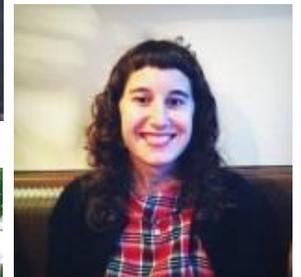
- Funded by Andrew W. Mellon Foundation
 - Phase 1: October 1, 2011 – September 30, 2013
 - Phase 2 – October 1, 2013 – September 30, 2014
- Partners: SILS at UNC and Maryland Institute for Technology in the Humanities (MITH)

BitCurator Goals

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

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)



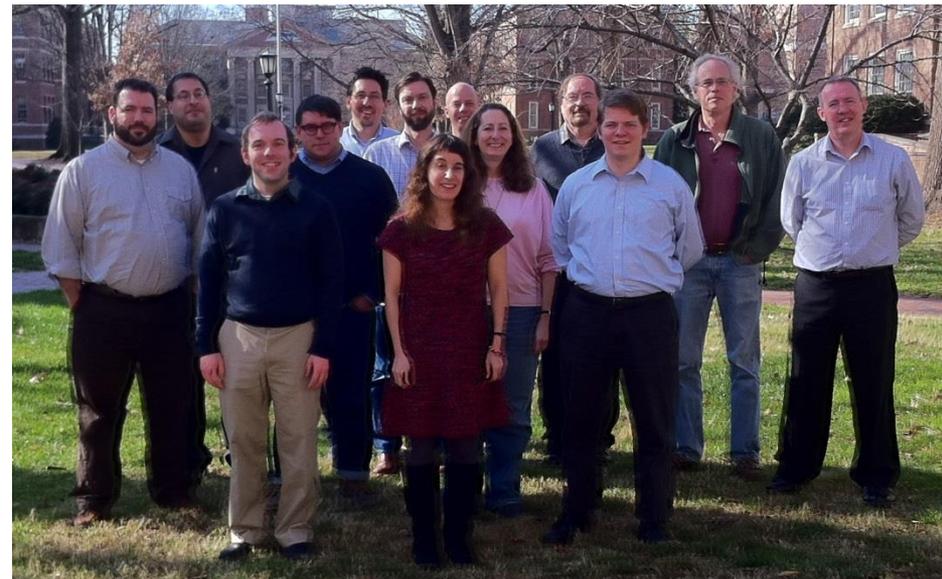
Two Groups of Advisors

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

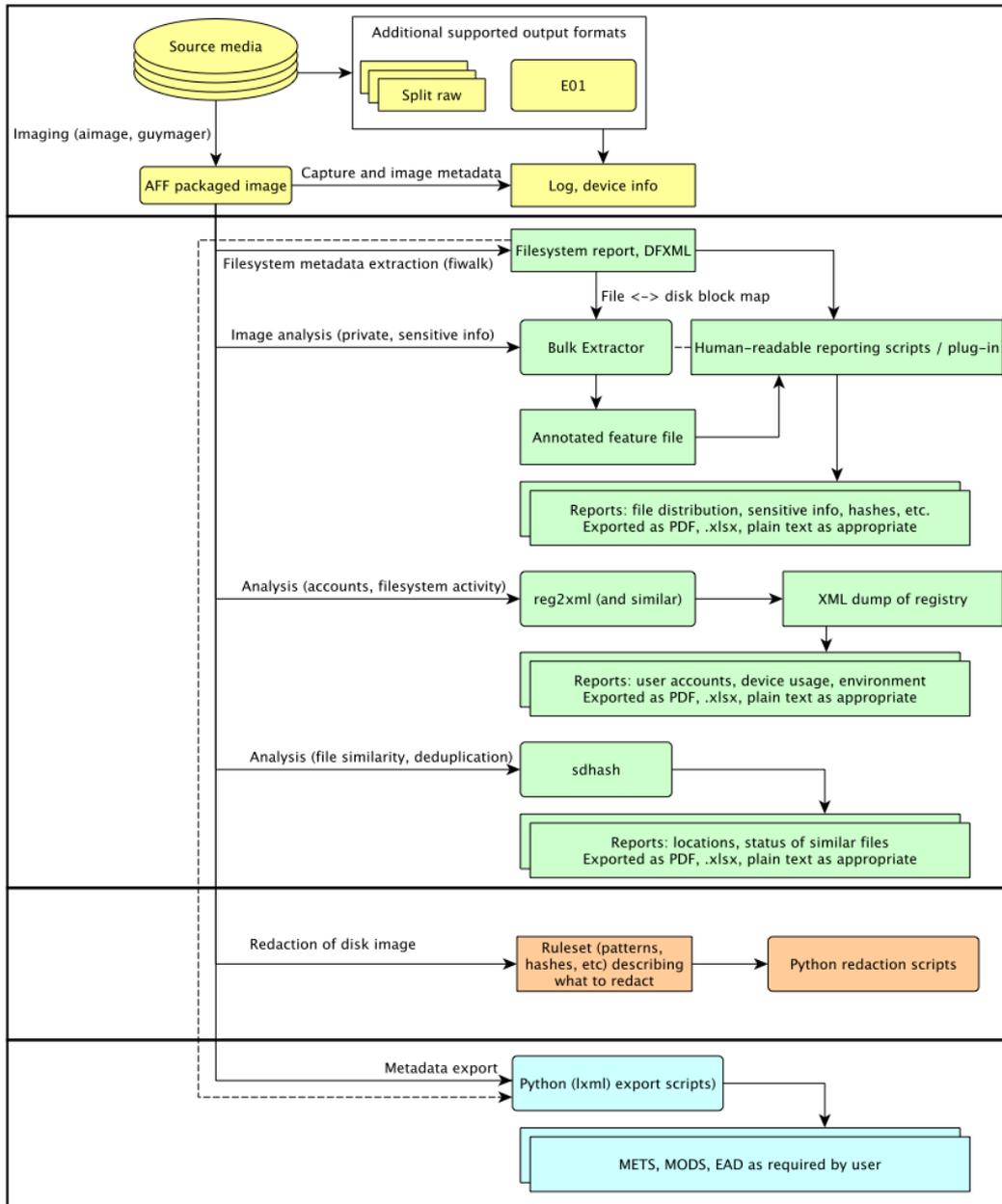


BitCurator Environment*

- Bundles, integrates and extends functionality of open source software: fiwalk, bulk_extractor, Guymager, The Sleuth Kit, sdhash and others
- 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

*To read about and download the environment, see: <http://wiki.bitcurator.net/>

BitCurator-Supported Workflow Elements



Acquisition

Reporting

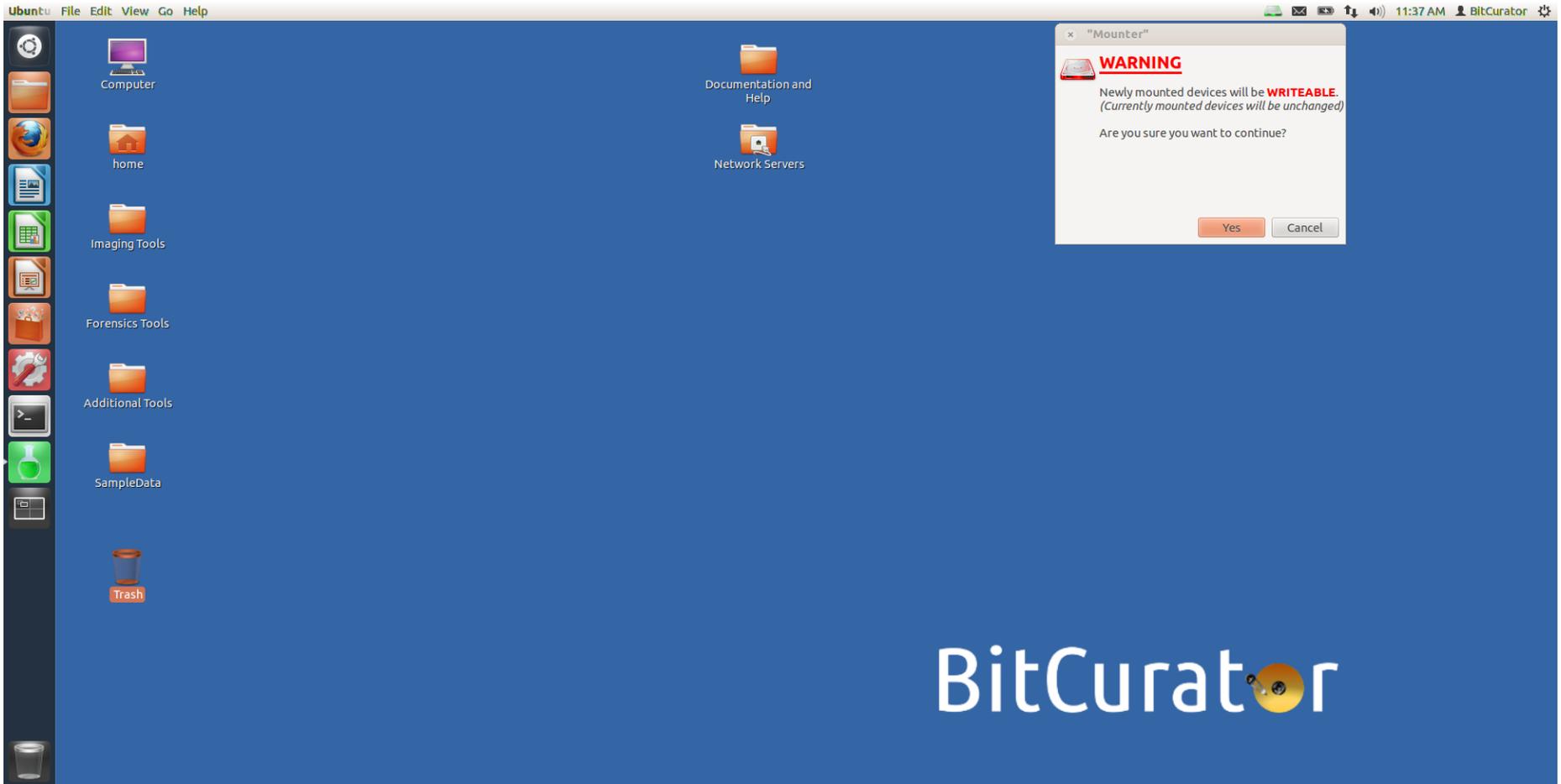
Redaction

Metadata export

- Acquisition
- Reporting
- Redaction
- Metadata Export

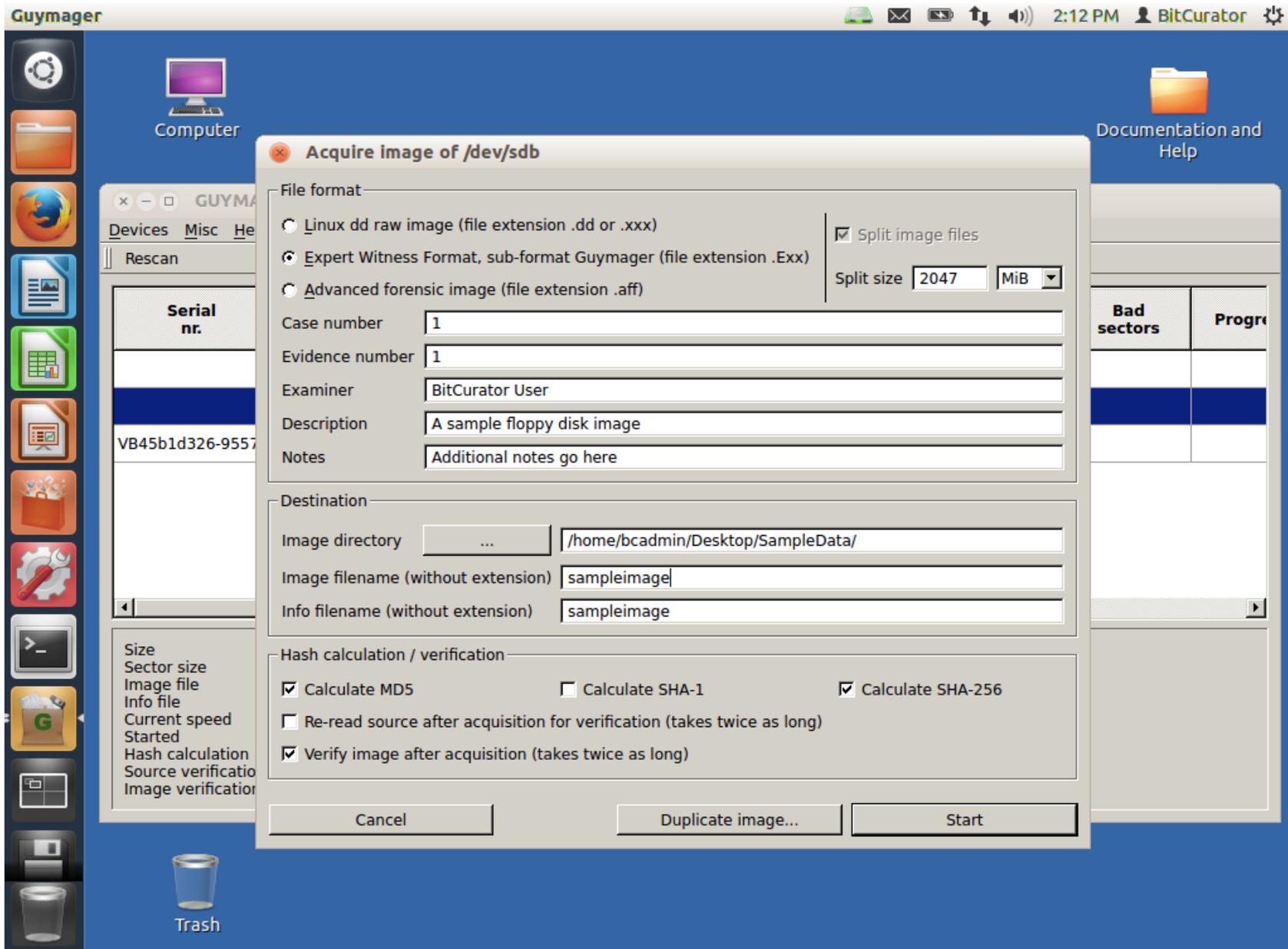
See: <http://bitcurator.net>

Mounted Devices set to Read-Only by Default*



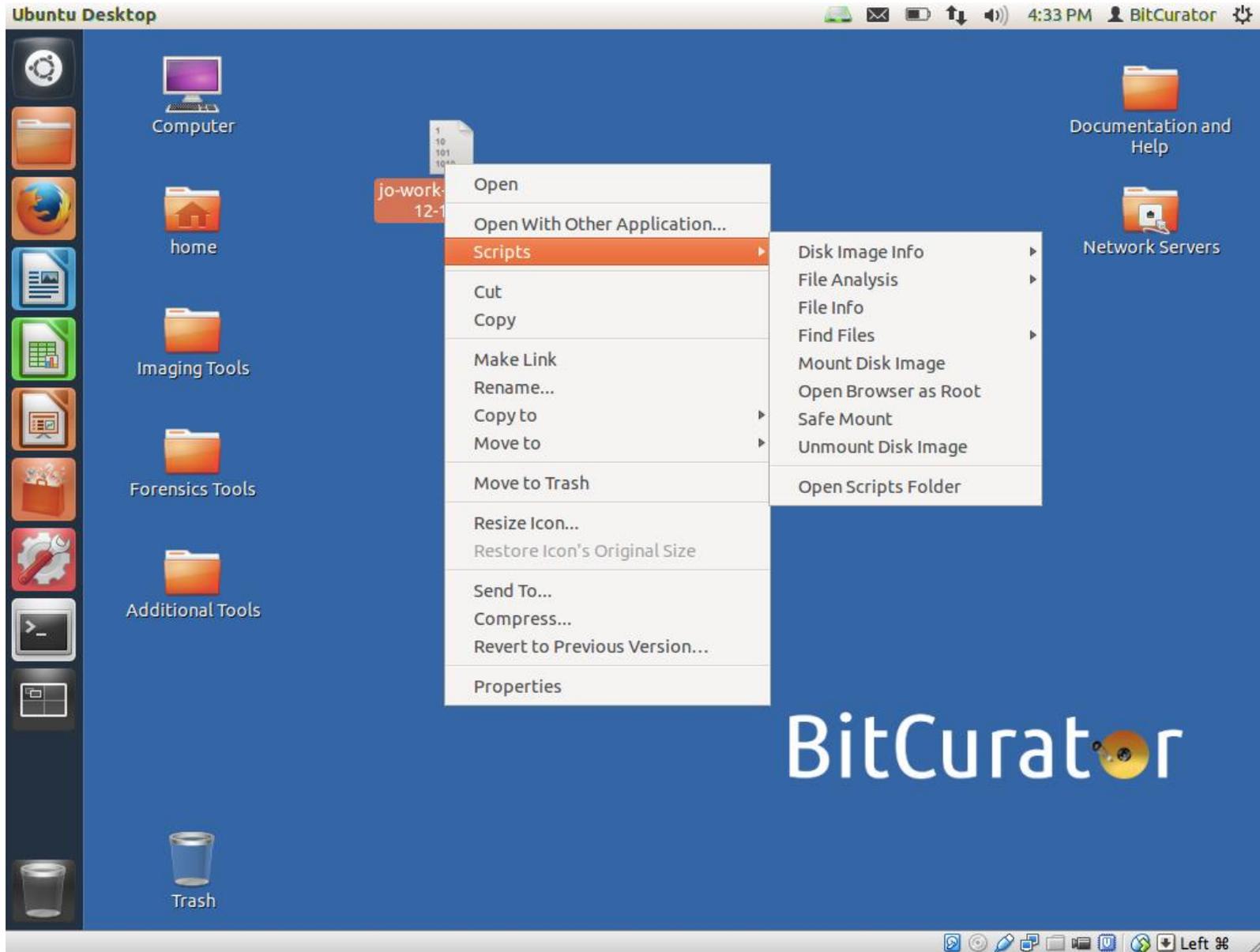
*Not to replace hardware-based write blocking, but useful for various purposes

Creating a Disk Image in Guymager*



*Developed by Guy Voncken

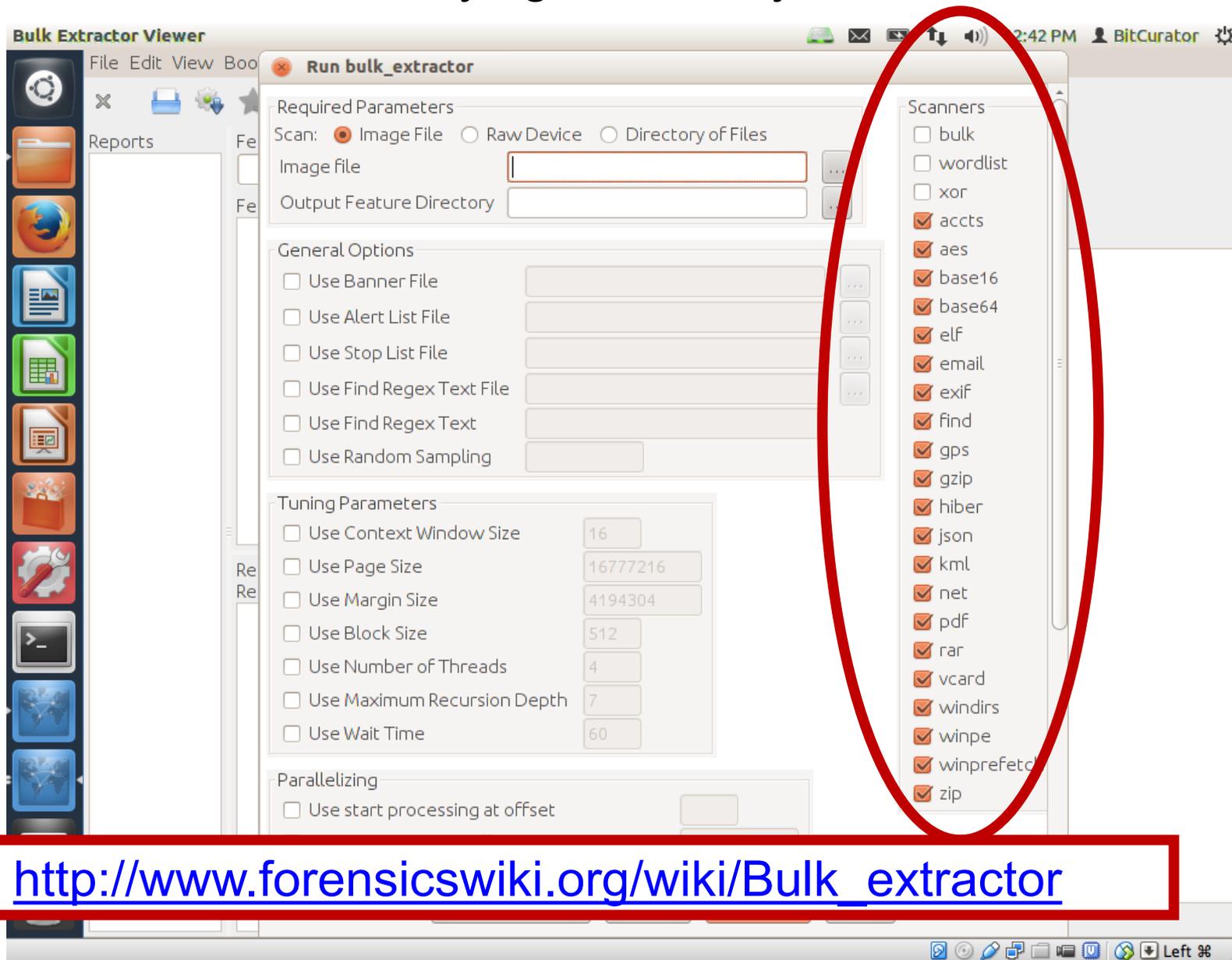
Mounting a Disk Image to Browse the Contents



Mounting a Disk Image to Browse the Contents



Bulk Extractor* – Identifying Potentially Sensitive Information



*Developed by Simson Garfinkel

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

The screenshot shows the BitCurator-0.2.0 Bulk Extractor Viewer interface. The main window displays a histogram of email addresses extracted from various files. The histogram is titled "Histogram File email_histog..." and lists the following data:

Count	Email Address
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@73yjt.lr
n=1	jqnmq@17.pn
n=1	kjph@sj.gr
n=1	nq9@5c7k.sg
n=1	pdcnfb@tft.ao
n=1	gyf@j65.de
n=1	tw+4vsa@xf.ms

Below the histogram is a table of "Referenced Feature File" and "Referenced Feature":

Referenced Feature File	Referenced Feature
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

The right pane shows a specific instance in context, titled "Image File sampleimage.E01". The feature is "privacy@Motorola.com". The image content is a snippet of text from a document, likely a privacy policy, which includes the following text:

your credit card number, so this information can only be viewed by Motorola. Motorola uses Secure Sockets Layer (SSL) encryption technology, the highest level of security on the Internet. The SSL protocol provides server authentication, data integrity, and privacy on the Web. This security measure helps ensure that no impostors, eavesdroppers, or vandals get your personal information. SSL not only encrypts your personal and financial information transmitted, including credit card information, but also verifies the identity of the server and that the original message arrives safely at its destination. However, no data transmission over the Internet can be guaranteed to be 100% secure. As a result, while we strive to protect your personal information, Motorola cannot ensure or warrant the security of any information you transmit to us or from our Web site, and therefore you use our site at your own risk. Once we receive your transmission, we use our best effort to ensure its security on our systems. 0002000007AE000038B6.7A8, As a global company Motorola has international sites and users all over the world. When you give Motorola personal information, that information may be sent electronically to servers outside of the country where you originally entered the information. In addition, that information may be used, stored and processed outside of the country where you entered that information. Whenever Motorola handles personal information, regardless of where this occurs, it takes steps to ensure that your information is treated securely and in accordance with the relevant Terms of Use and this Privacy Policy. How can I correct or change my personal information? If you would like to review, correct or change any personal information you have provided, or remove your name from our mailing list, please e-mail us at privacy@Motorola.com. If you have established a "user profile" on a Motorola website, you may change the information you provided at an

BitCurator Reporting Tool

The screenshot displays the BitCurator Reporting Tool window. The interface includes a sidebar with navigation icons for Computer, home, Imaging Tools, Forensics Tools, and Additional Tools, along with a Trash icon. The main window has a title bar and a menu bar with options: Run All, Fiwalk XML, Annotated Features, Reports, and File Access. The 'Reports' tab is active, showing instructions to run fiwalk, generate Office/PDF reports, and a 'Launch BEViewer' button. Below this, several fields are filled with paths: Image File (/home/bcadmin/Desktop/SampleData/sampleimage.E01), Bulk Extractor Feature Directory (/home/bcadmin/Desktop/SampleData/bulk-extractor-output), Output Directory (/home/bcadmin/Desktop/SampleData/reporting-output), and Config File (Optional) (/Path/To/File). The Command Line Output section shows a list of generated report files and an Excel report, followed by a success message: '>> Success!!! BitCurator Reports generated in the directory: o /home/bcadmin/Desktop/SampleData/reporting-output/reports'. At the bottom, there are buttons for Close, Cancel, and Run.

BitCurator Reporting Tool

Bitcurator Reports

Run All | Fiwalk XML | Annotated Features | Reports | File Access

Run fiwalk, annotate the bulk_extractor output, and generate Office / PDF reports.
If you haven't run bulk_extractor yet, use the button to the right to launch and run it first.

Image File
/home/bcadmin/Desktop/SampleData/sampleimage.E01

Bulk Extractor Feature Directory
/home/bcadmin/Desktop/SampleData/bulk-extractor-output

Output Directory (fiwalk output, annotated features, and reports will appear in here)
/home/bcadmin/Desktop/SampleData/reporting-output

Config File (Optional)
/Path/To/File

Command Line Output

```
1) /home/bcadmin/Desktop/SampleData/reporting-output/reports/bc_format_bargraph.pdf
2) /home/bcadmin/Desktop/SampleData/reporting-output/reports/format_table.pdf
3) /home/bcadmin/Desktop/SampleData/reporting-output/reports/FiwalkReport.pdf
4) /home/bcadmin/Desktop/SampleData/reporting-output/reports/FiwalkDeletedFiles.pdf
5) /home/bcadmin/Desktop/SampleData/reporting-output/reports/BeReport.pdf
Generating Excel report /home/bcadmin/Desktop/SampleData/reporting-output/reports/fiwalk-output.xml.xlsx
```

>> Success!!! BitCurator Reports generated in the directory:
o /home/bcadmin/Desktop/SampleData/reporting-output/reports

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 0x2, O...ctors 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	TIF_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 80386, l..., InnoSetup self-extracting archive	PE3_2e	1
JPEG image data, JFIF standard 1.01, ...25x5C276x5C332hex5C0115jx5C261"	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	40
dat_ata	31
emp_pty	9
PDF_1-4	6
JPE_01	4
PE3_2e	1
TIF_ian	2
emp_ct-	2
ASC_at-	1
Com_ifo	1
PE3_2e	1
ASC_ors	1
Sys_ser	1
x86_x0-	1
MS_ors	1
PCX_ata	1
new_ors	1

Page 1

Look Familiar? Filesystem Metadata - Output from fiwalk*

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

*Developed by Simson Garfinkel

Specialized BitCurator Reports

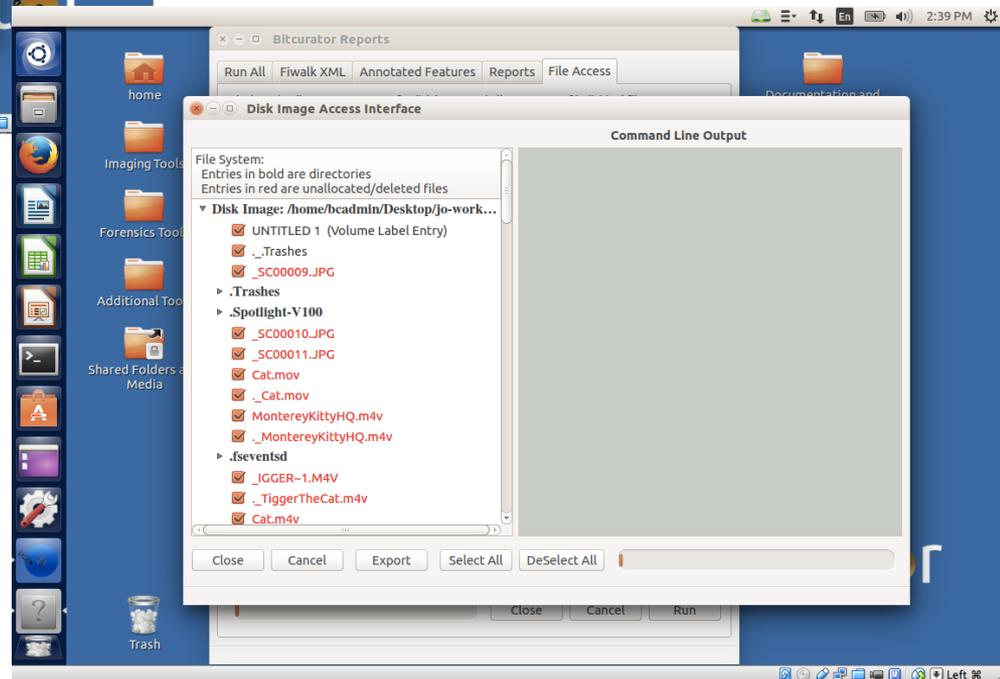
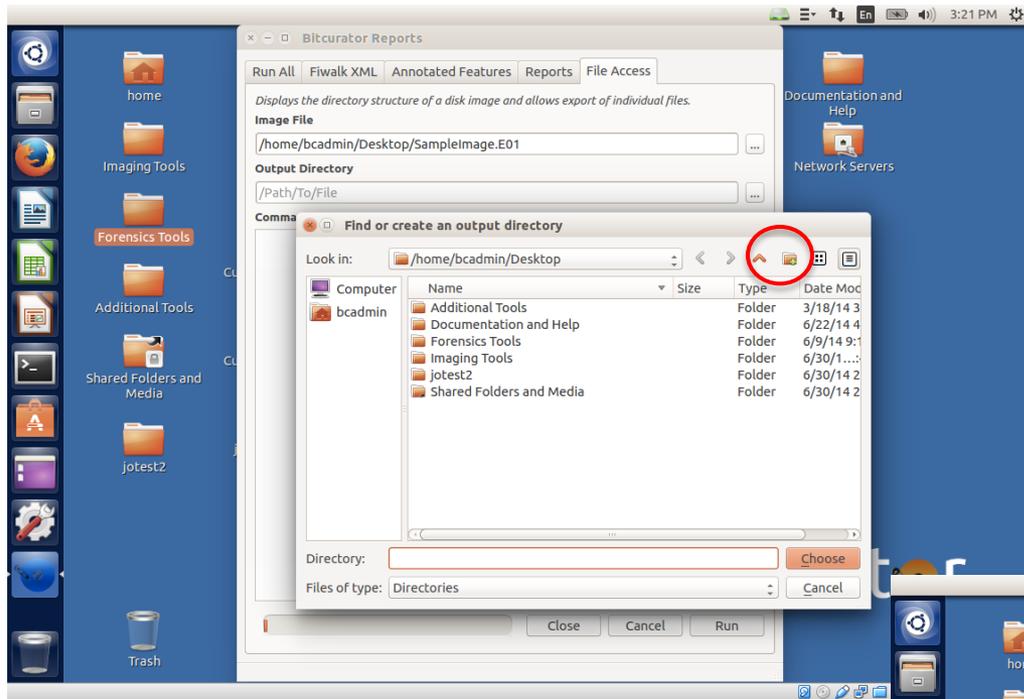
File	Content
bc_format_bargraph.pdf	histogram of file formats found on the volume
bulk_extractor_report.pdf	high-level overview of feature locations on disk
fiwalk_deleted_files.pdf	shows paths to any deleted materials found in a given partition
fiwalk-output.xml.xlsx	Excel converted DFXML output (file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

Maybe Less Familiar? PREMIS (Preservation) Metadata Generated from Running BitCurator Tools – Recorded as PREMIS Events

```
premis.xml (~/Desktop/demo1/demo1reports/reports) - gedit
Open Save Undo 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 |

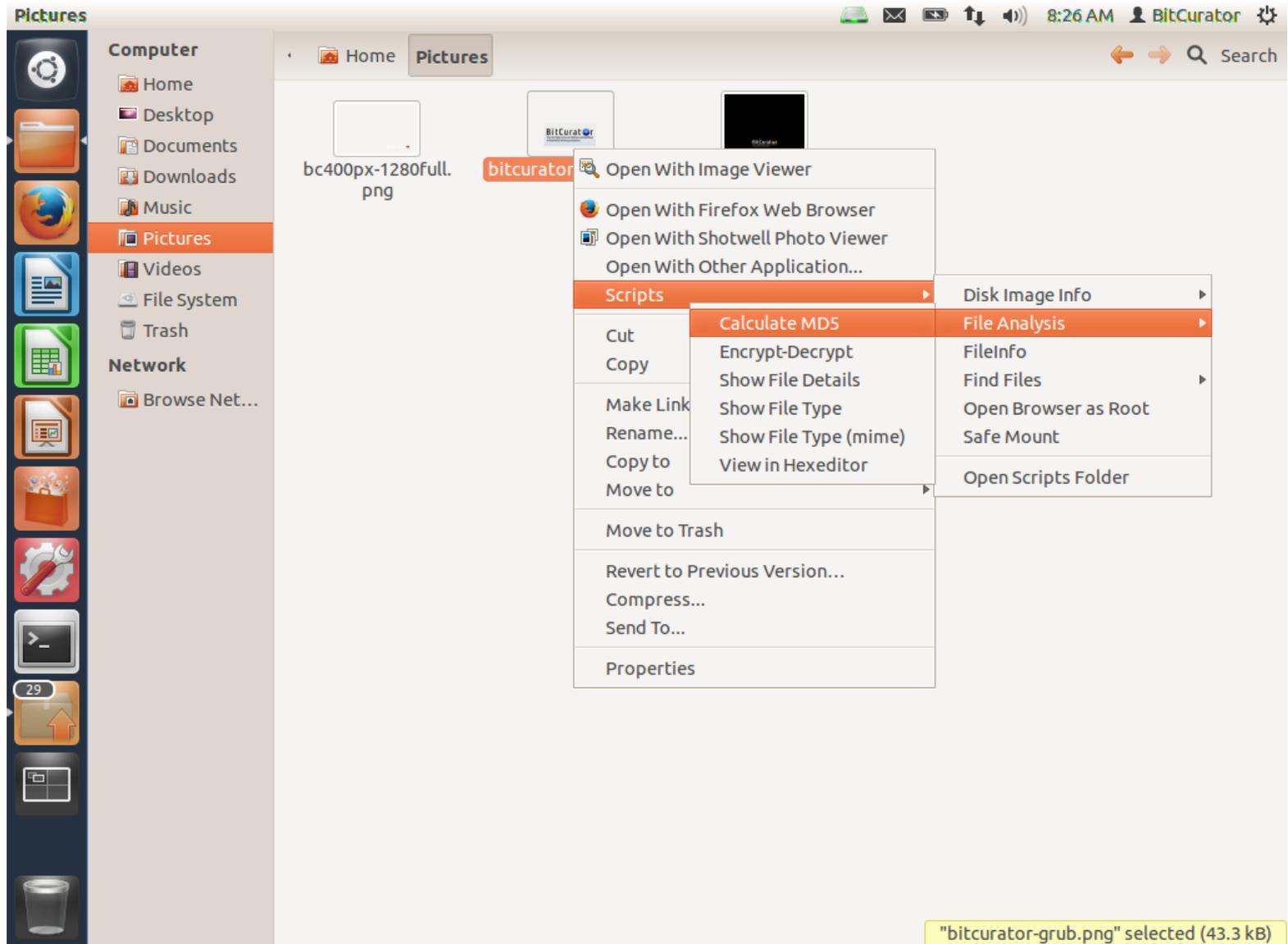
Exporting Files from a Disk Image



Nautilus Scripts

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

Right Click on File or Directory and Calculate MD5



Computer

- Home
- Desktop
- Documents
- Downloads
- Music
- Pictures**
- Videos
- File System
- Trash

Network

- Browse Net...

bc400px-1280full.png

bitcurator-grub.png

bitcurator-grub-new.png

Calculate MD5 (Files and Directories)

Please choose the way you want the MD5 hash to be presented:
(1 file(s) selected)

Handling
<input checked="" type="radio"/> Display on screen
<input type="radio"/> Save to file (the selected filename + .md5 extension)

Cancel OK

Computer

- Home
- Desktop
- Documents
- Downloads
- Music
- Pictures**
- Videos
- File System
- Trash

Network

- Browse Net...

Home Pictures

bc400px-1280full.png bitcurator-grub.png bitcurator-grub-new.png

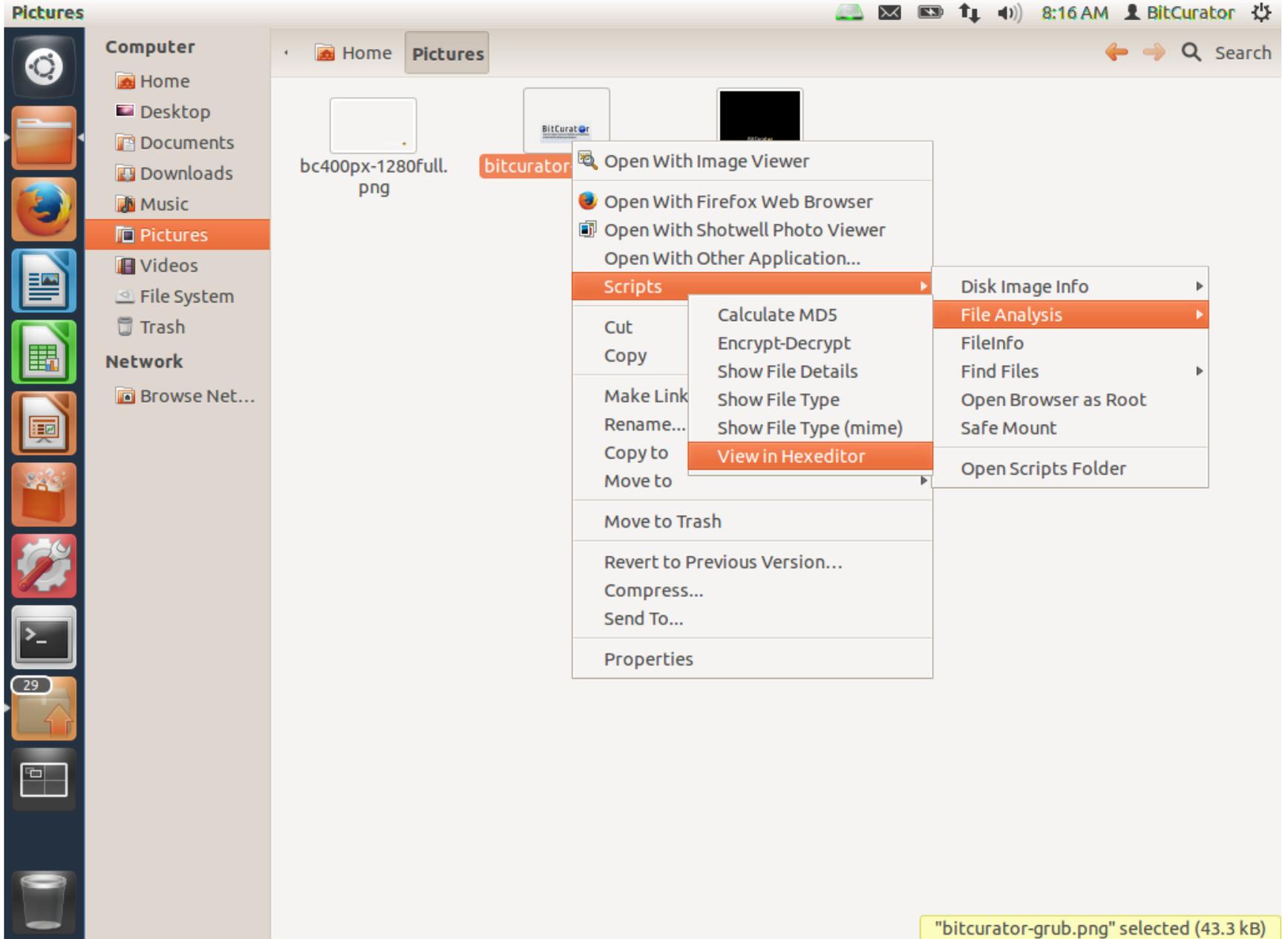
Calculate MD5 (Files and Directories)

The MD5 hash of the selected file:

keb2622125be1231b0fc9babee27942d /home/bcadmin/Pictures/bitcurator-grub.png

Cancel OK

Quick Access to a Hex View:



bitcurator-grub.png - GHex

00000000	89 50 4E 47 0D 0A 1A 0A 00 00 00 0D 49 48 44 52 00 00	.PNG.....IHDR..
00000012	02 80 00 00 01 E0 08 02 00 00 00 BA B3 4B B3 00 00 00K....
00000024	09 70 48 59 73 00 00 0B 13 00 00 0B 13 01 00 9A 9C 18	.pHYs.....
00000036	00 00 0A 4F 69 43 43 50 50 68 6F 74 6F 73 68 6F 70 20	...0iCCPPhotoshop
00000048	49 43 43 20 70 72 6F 66 69 6C 65 00 00 78 DA 9D 53 67	ICC profile..x..Sg
0000005A	54 53 E9 16 3D F7 DE F4 42 4B 88 80 94 4B 6F 52 15 08	TS..=...BK...KoR..
0000006C	20 52 42 8B 80 14 91 26 2A 21 09 10 4A 88 21 A1 D9 15	RB...&*!..J.!...
0000007E	51 C1 11 45 45 04 1B C8 A0 88 03 8E 8E 80 8C 15 51 2C	Q..EE.....Q,
00000090	0C 8A 0A D8 07 E4 21 A2 8E 83 A3 88 8A CA FB E1 7B A3!.....{.
000000A2	6B D6 BC F7 E6 CD FE B5 D7 3E E7 AC F3 9D B3 CF 07 C0	k.....>.....
000000B4	08 0C 96 48 33 51 35 80 0C A9 42 1E 11 E0 83 C7 C4 C6	...H3Q5...B.....
000000C6	E1 E4 2E 40 81 0A 24 70 00 10 08 B3 64 21 73 FD 23 01	...@..\$p....d!s.#.
000000D8	00 F8 7E 3C 3C 2B 22 C0 07 BE 00 01 78 D3 0B 08 00 C0	...~<<+".....x.....
000000EA	4D 9B C0 30 1C 87 FF 0F EA 42 99 5C 01 80 84 01 C0 74	M..0....B.\.....t
000000FC	91 38 4B 08 80 14 00 40 7A 8E 42 A6 00 40 46 01 80 9D	.8K....@z.B..@F...
0000010E	98 26 53 00 A0 04 00 60 CB 63 62 E3 00 50 2D 00 60 27	.&S.....`cb..P-..`
00000120	7F E6 D3 00 80 9D F8 99 7B 01 00 5B 94 21 15 01 A0 91{..[.!.....
00000132	00 20 13 65 88 44 00 68 3B 00 AC CF 56 8A 45 00 58 30	..e.D.h;...V.E.X0
00000144	00 14 66 4B C4 39 00 D8 2D 00 30 49 57 66 48 00 B0 B7	..fK.9.-.0IWfH...
00000156	00 C0 CE 10 0B B2 00 08 0C 00 30 51 88 85 29 00 04 7B0Q...{
00000168	00 60 C8 23 23 78 00 84 99 00 14 46 F2 57 3C F1 2B AE	..`##x.....F.W<.+.
0000017A	10 E7 2A 00 00 78 99 B2 3C B9 24 39 45 81 5B 08 2D 71	..*..x..<.\$9E.[.-q
0000018C	07 57 57 2E 1E 28 CE 49 17 2B 14 36 61 02 61 9A 40 2E	.WW..(.I.+6a.a.@.
0000019E	C2 79 99 19 32 81 34 0F E0 F3 CC 00 00 A0 91 15 11 E0	.y..2.4.....
000001B0	83 E3 ED 78 CE 0E AE CE CE 36 8E B6 0E 5E 2D FA BE 06	x.....6.....

Signed 8 bit:	-119	Signed 32 bit:	1196314761	Hexadecimal:	89
Unsigned 8 bit:	137	Unsigned 32 bit:	1196314761	Octal:	211
Signed 16 bit:	20617	Float 32 bit:	5.281654e+04	Binary:	10001001
Unsigned 16 bit:	20617	Float 64 bit:	5.292398e-260	Stream Length:	8

Documentation and Help

Network Servers

BitCurator



Computer

- Home
- Desktop**
- Documents
- Downloads
- Music
- Pictures
- Videos
- File System
- Trash

Network

- Browse Net...

Additional Tools Documentation and Help Forensics

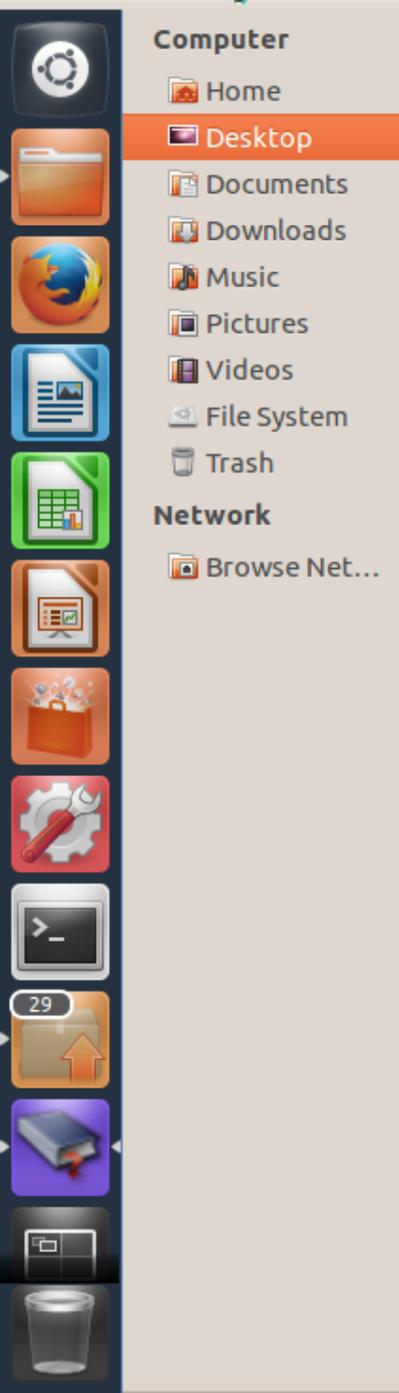
nps-2010-emails.E01

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

- Show AFF Info
- Show E01 Info

- 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



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

Quick Start Guide

Most recent version always available at:

<http://wiki.bitcurator.net/>

BitCurator

Quick Start Guide v0.9.16

Last updated: August 3, 2014



UNC
SCHOOL OF INFORMATION
AND LIBRARY SCIENCE

MITH MARYLAND INSTITUTE FOR
TECHNOLOGY IN THE HUMANITIES

Other Functionality to Meet Identified LAM Needs:

Function	Tool(s)
Identify duplicate files	FSLint
Characterize files	FITS
Examine, copy and extract information from old Mac disks	HFSExplorer
Package files for storage and/or transfer	BagIt (Java) library
Scan for viruses	ClamTK
Read contents of Microsoft Outlook PST files	readpst
Examine embedded header information in images	pyExifToolGUI
Generate images of problematic disks or particular disk types	dd, dcfldd, ddrescue, cdrdao (in addition to Guymager)
Identify files that are partially similar but not identical	sdhash, ssdeep

Considerable Uptake

- 118 members of the BitCurator Users mailing list
- 640 individuals who have participated in BitCurator-related events (workshops, tutorials, hackathons)
- Numerous publications and reports by library/archives students and professionals about their testing and use of the software
- 934 Twitter followers (whatever that means)

Open Source Development Strategy

- Rapid development with numerous iterations based on several channels of user feedback
- Code released under GPL, v3 (perhaps moving to Apache) – available through GitHub
- Existing code incorporated is generally GPL or public domain (government products)
- Packaging elements of the code to be integrated into other environments (e.g. Archivematica)
- Significant community engagement, including work of Community Lead (Porter Olsen)

BitCurator Consortium

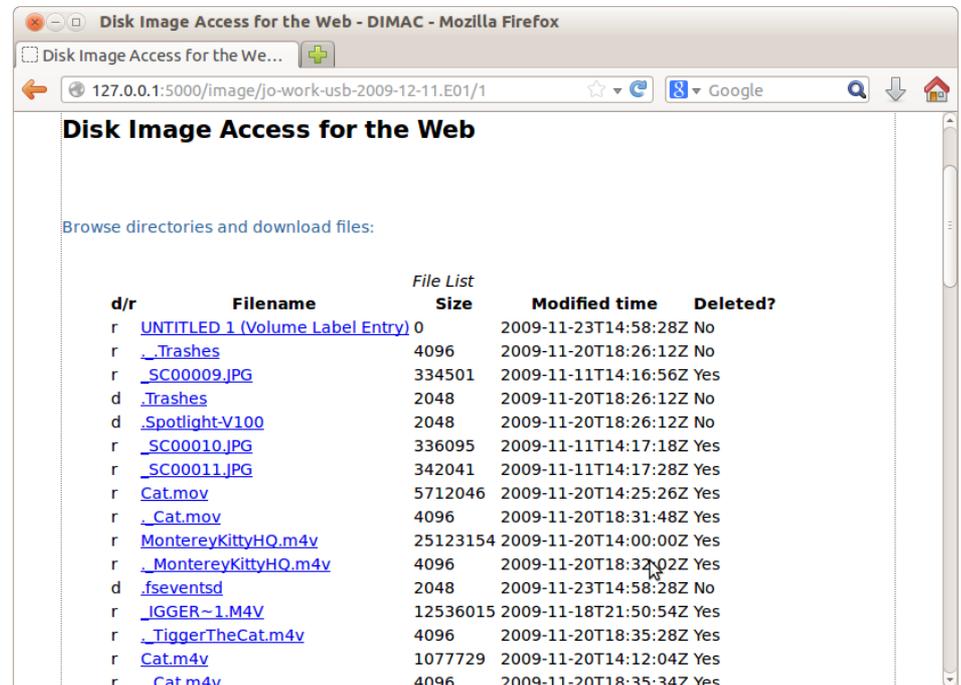
- Continuing home for hosting, stewardship and support of BitCurator tools and associated user engagement
- Administrative home: Educopia Institute
- Funding based on membership dues
- Institutions as members, with two categories of membership: Charter and General
- Software and documentation will continue to be free and open source, but membership provides further benefits (e.g. support, training, development priority)

DIMAC (Disk Image Access for the Web)

- Developed by Sunitha Misra and Kam Woods
- To dynamically navigate and download contents of a disk image, without having to download or mount the full image
- See: <https://github.com/kamwoods/dimac>
- Demo at:

<http://www.youtube.com/watch?v=BwiWFqxYzQ8>

See: Sunitha Misra, Christopher A. Lee, and Kam Woods, “A Web Service for File-Level Access to Disk Images,” *Code4Lib Journal*, <http://journal.code4lib.org/articles/9773>.



Disk Image Access for the Web

Browse directories and download files:

d/r	Filename	File List Size	Modified time	Deleted?
r	UNTITLED 1 (Volume Label Entry)	0	2009-11-23T14:58:28Z	No
r	_Trashes	4096	2009-11-20T18:26:12Z	No
r	_SC00009.JPG	334501	2009-11-11T14:16:56Z	Yes
d	_Trashes	2048	2009-11-20T18:26:12Z	No
d	_Spotlight-V100	2048	2009-11-20T18:26:12Z	No
r	_SC00010.JPG	336095	2009-11-11T14:17:18Z	Yes
r	_SC00011.JPG	342041	2009-11-11T14:17:28Z	Yes
r	_Cat.mov	5712046	2009-11-20T14:25:26Z	Yes
r	_Cat.mov	4096	2009-11-20T18:31:48Z	Yes
r	MontereyKittyHQ.m4v	25123154	2009-11-20T14:00:00Z	Yes
r	_MontereyKittyHQ.m4v	4096	2009-11-20T18:32:02Z	Yes
d	_fsevents	2048	2009-11-23T14:58:28Z	No
r	_JGGER~1.M4V	12536015	2009-11-18T21:50:54Z	Yes
r	_TiggerTheCat.m4v	4096	2009-11-20T18:35:28Z	Yes
r	_Cat.m4v	1077729	2009-11-20T14:12:04Z	Yes
r	_Cat.m4v	4096	2009-11-20T18:35:34Z	Yes

APPLYING FORENSICS TO PRESERVING THE PAST: CURRENT ACTIVITIES AND FUTURE POSSIBILITIES

- Organizers: Cal Lee, Jeremy Leighton John, Susan Thomas
- To be held at Digital Libraries 2014, London, September, 8-12, 2014
- One-day event, split across an afternoon and following morning (Sept 11-12)
- Short papers and talks, group discussion and formation of steps for further action

Thank You!

BitCurator is a suite of open source digital forensics and data analysis tools to help collecting institutions (libraries, archives, and museums) process born-digital materials. BitCurator supports positive digital preservation outcomes using software and practices adopted from the digital forensics community.

- **Create forensic disk images:** Disk images packaged with metadata about devices, file systems, and the creation process.
- **Analyze files and file systems:** View details on file system contents from a wide variety of file systems.
- **Extract file system metadata:** File system metadata is a critical link in the chain of custody and in records of provenance.
- **Identify and redact sensitive information:** Locate private and sensitive information on digital media and prepare materials for public access.
- **Identify and remove duplicate files:** Know what files to keep and what can be discarded.

Find more details about specific programs available in BitCurator on our software page.

Documentation

1. Add BitCurator to Your Archival Workflow
2. Install BitCurator
3. Prepare Your Media
4. Data Triage
5. Using Digital Forensics Tools

The following descriptions include text-and-screenshot tutorials. You can also visit our video gallery to watch related screencasts.

Get BitCurator

- BitCurator Virtual Machine (v0.9.16) [Download Mirror] [MD5 checksum]
- BitCurator Installation ISO (v0.9.16) [Download Mirror] [MD5 checksum]
- Quickstart Guide v0.9.16 Walk through installation and BitCurator use.

The BitCurator Community Get support and speak with members of the team.

BitCurator

Install Using BitCurator Community

About Blog Research People Support Community

News

BitCurator Users Gathering @ Archives*Records - August 12, 2014

Posted July 23, 2014 by Porter Olsen

Please join us for a gathering of BitCurator users on the evening of August 12th after the SAA Research Forum from 5:30PM to 7:30PM. Come meet other BitCurator users and share your experiences working with digital forensics and BitCurator. In addition to collaborating with other BitCurator users, meet with BitCurator Co-PI's Cal Lee and Matt ... [READ MORE](#)

BitCurator Consortium — Membership Now Open

Posted July 8, 2014 by Porter Olsen

The BitCurator Consortium (BCC) is an independent, community-led membership association that will serve as the host and center of administrative, user and community support for the BitCurator environment. Its purpose is to support curation of born-digital materials through the application of open-source digital forensics tools by institutions responsible for such

Video Spotlight

BitCurator

Tools for Digital Forensics Methods and Workflows in Real-World Collecting Institutions

Screencast Tutorial: Installing BitCurator as a virtual machine in VirtualBox

UNC MITH HARVARD INSTITUTE FOR TECHNOLOGY IN THE HUMANITIES

Events

BitCurator Webinar on Digital Forensics Metadata - Session

Get the software
Documentation and technical specifications
Screencasts
Google Group
<http://wiki.bitcurator.net/>

People
Project overview
Publications
News
<http://www.bitcurator.net/>

Twitter: @bitcurator