

Digital Forensics with BitCurator

Based on the July 29, 2014
BitCurator workshop
at METRO

Kevin Schlottmann
August 26, 2014

What is digital forensics?

"Process of identifying, preserving, analyzing, and presenting digital evidence in a manner that is legally acceptable."



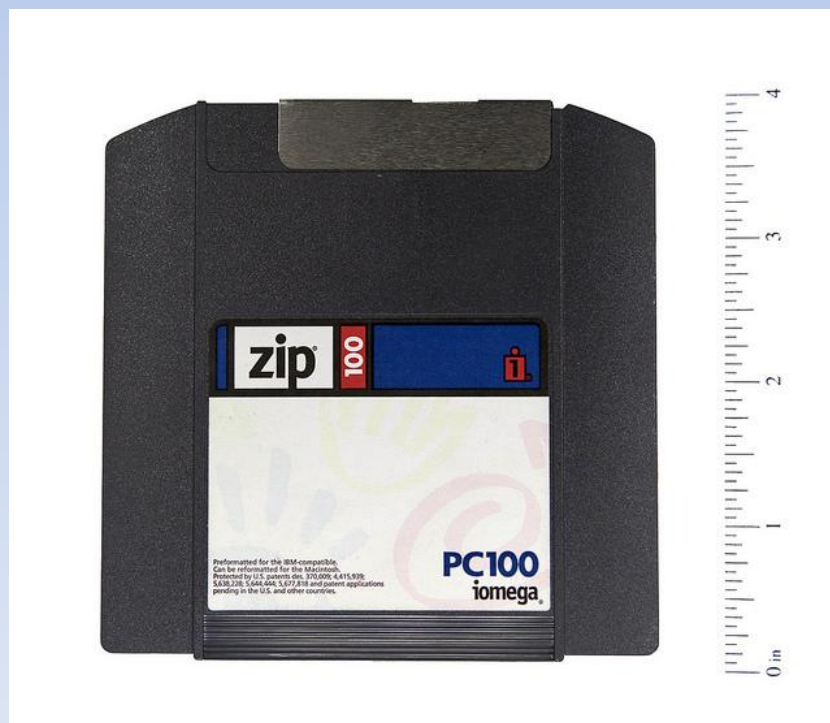
Why apply digital forensics?

- *Ensure data integrity
- *Ease of automation and processing



Why apply digital forensics?

*In other words: to ensure provenance, original order, chain of custody of digital objects



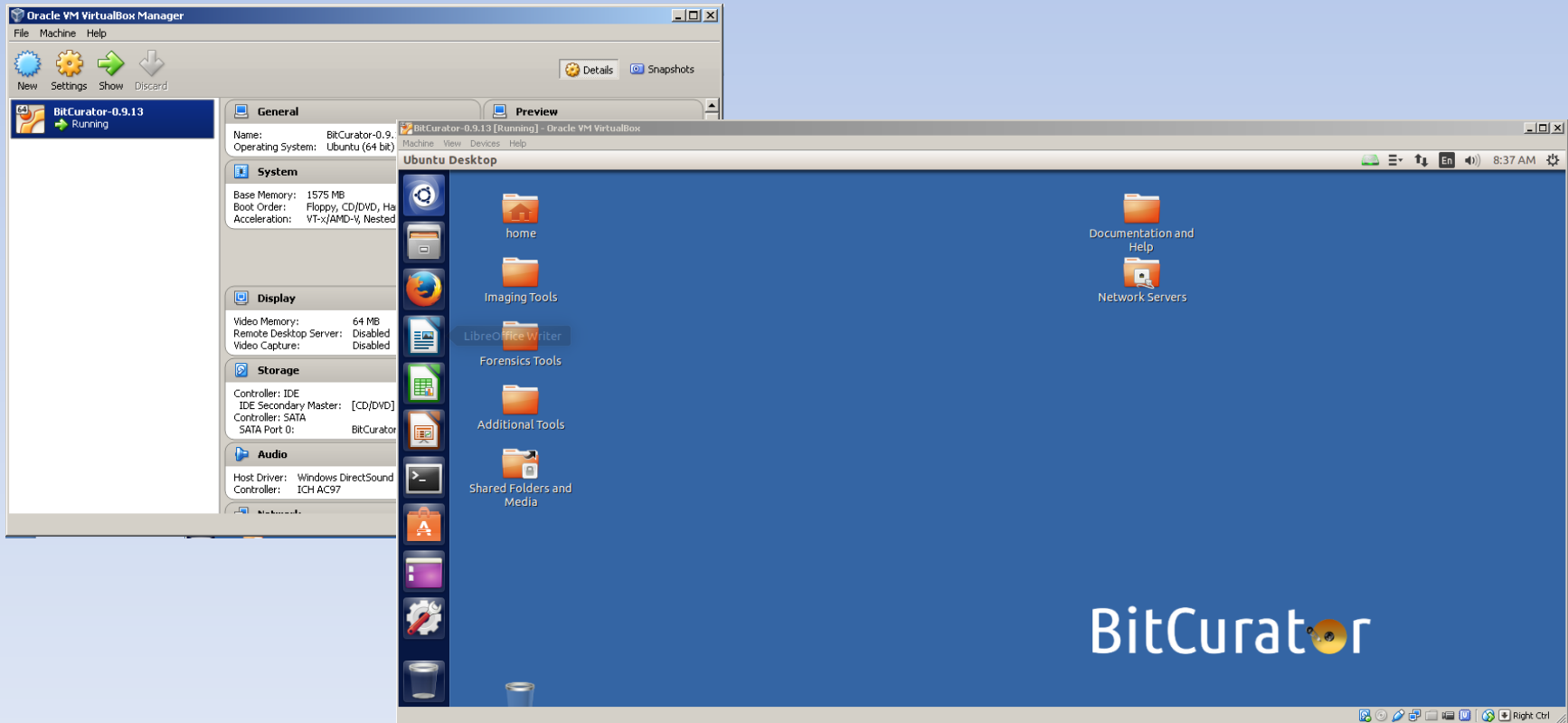
What is BitCurator?

*Customized Linux OS running in virtual machine with a tightly integrated, well-documented suite of open-source digital forensics tools



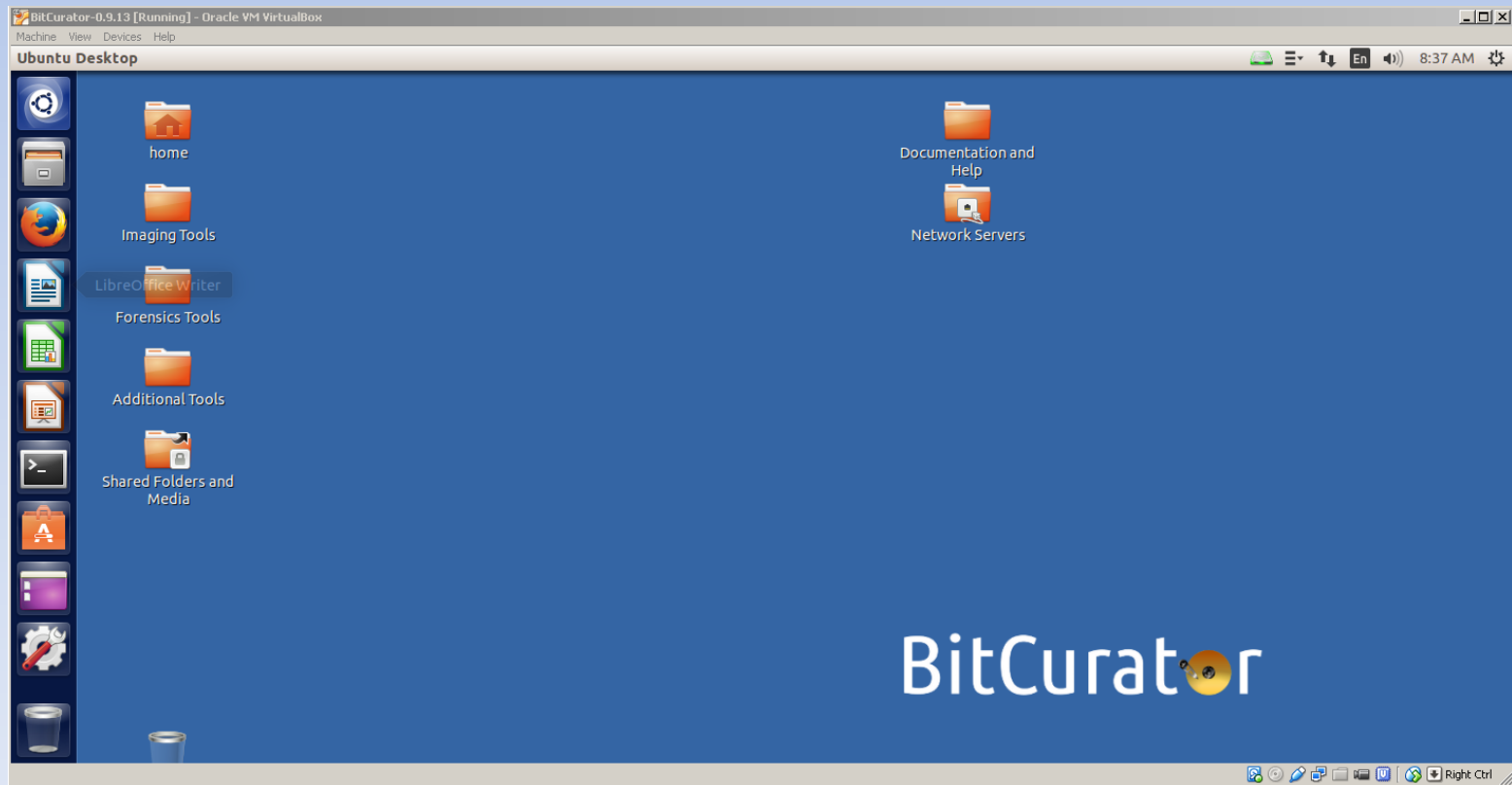
What is BitCurator?

***Customized Linux OS running in virtual machine...**



What is BitCurator?

***Customized Linux OS running in virtual machine...**



What is BitCurator?

***Customized Linux OS running in virtual machine...**

```
bcadmin@ubuntu:~$ cd /home/bcadmin/Tools/fits
bcadmin@ubuntu:~/Tools/fits$ ./fits.sh -h

usage: fits
-h          print this message
-i <arg>   input file or directory
-o <arg>   output file
-r          process directories recursively when -i is a directory
-v          print version information
-x          convert FITS output to a standard metadata schema
-xc        output using a standard metadata schema and include FITS xml
bcadmin@ubuntu:~/Tools/fits$
bcadmin@ubuntu:~/Tools/fits$ █
```


What is BitCurator?

*** ...a tightly integrated, well-documented suite of open-source digital forensics tools**

Guymager homepage

Introduction

Guymager is a free forensic imager for media acquisition

Home

Projects

Blog

The Sleuth Kit

[File Systems](#)
[Plug-in Framework](#)
[Download](#)
[Documents](#)
[History](#)
[Licenses](#)

[Autopsy](#)

[Sleuth Kit Hadoop](#)

[mac-robber](#)

Overview

The Sleuth Kit® (TSK) is a library for analyzing disk images. The core functionality framework allows you to incorporate various file systems. The library can be incorporated directly into your application or used to find evidence.

- [Volume and File System](#)
- [Plug-in Framework](#)
- [Download](#)
- [Documents](#)
- [History](#)
- [Licenses](#)

 [simsong / bulk_extractor](#)

Home

BruceMty edited this page on Jun 19 · 16 revisions

Welcome to the bulk_extractor wiki!

bulk_

extra

are s

tools

more

proc

pytho

Bulk

and f

Fiwalk

fiwalk is a batch forensics analysis program written in C that uses SleuthKit. The program can output in XML or ARFF format

Contents [hide]

- 1 Temporary Distribution Point
- 2 Legacy Distribution
- 3 XML Example
- 4 Availability
- 5 See Also

1. Creating a disk image

The screenshot shows the GUYMAGER application window. The title bar reads "GUYMAGER". Below the title bar are menu items: "Devices", "Misc", and "Help". A "Rescan" button is visible. The main area contains a table with the following data:

Serial nr.	Linux device	Model	State	Size	Hidden areas	Bad sectors	Progress
VB27f4cd1c-fc439adf	/dev/sda	ATA VBOX HARDDISK	● Idle	274.9GB	unknown		

Below the table is a progress bar. At the bottom of the window, the following information is displayed:

Size 274,877,906,944 bytes (256GiB / 275GB)
Sector size 512
Image file
Info file
Current speed
Started
Hash calculation
Source verification
Image verification

2. Analyzing the disk image

Bulk Extractor Viewer

Highlight: Match case

Reports

- be_data
 - ccn.txt
 - ccn_histogram.txt
 - domain.txt
 - domain_histogram.txt**
 - email.txt
 - email_histogram.txt
 - exif.txt
 - gps.txt
 - rfc822.txt
 - tcp.txt
 - tcp_histogram.txt
 - telephone.txt
 - telephone_histogram.txt
 - url.txt
 - url_histogram.txt
 - url_services.txt
 - windirs.txt
 - zip.txt

Feature Filter Match case

Histogram File: domain_histogram.txt

n=12	www.zotero.org
n=11	www.paradigm.ac.uk
n=10	www.oclc.org
n=8	bitcurator.set
n=8	graphml-graphdrawing.org
n=8	nokia.com
n=8	thanete.com
n=8	www.cldr.org
n=7	mail.ipa.es
n=7	www.archivists.org
n=7	www.bl.uk
n=7	www.loc.gov
n=7	www.soh.gov
n=7	www.slideshare.net
n=6	britishlibrary.typepad.co.uk
n=6	document.com
n=6	expo.lcs.mit.edu
n=6	gnoma.org
n=6	matienzo.org
n=6	mitb.umd.edu
n=6	netscape.com

Referenced Feature File: domain.txt

14273536	www.archivists.org
14273651	www.archivists.org
14285577	www.archivists.org
1111 (10593435-PDF)	www.archivists.org
1173 (10593435-PDF)	www.archivists.org
1251 (10593435-PDF)	www.archivists.org

Navigation

backupcd E01, 10503435-PDF-1049, www.archivists.org

Image File backupcd E01

Feature File domain.txt

Feature Path 1049(10503435-PDF)

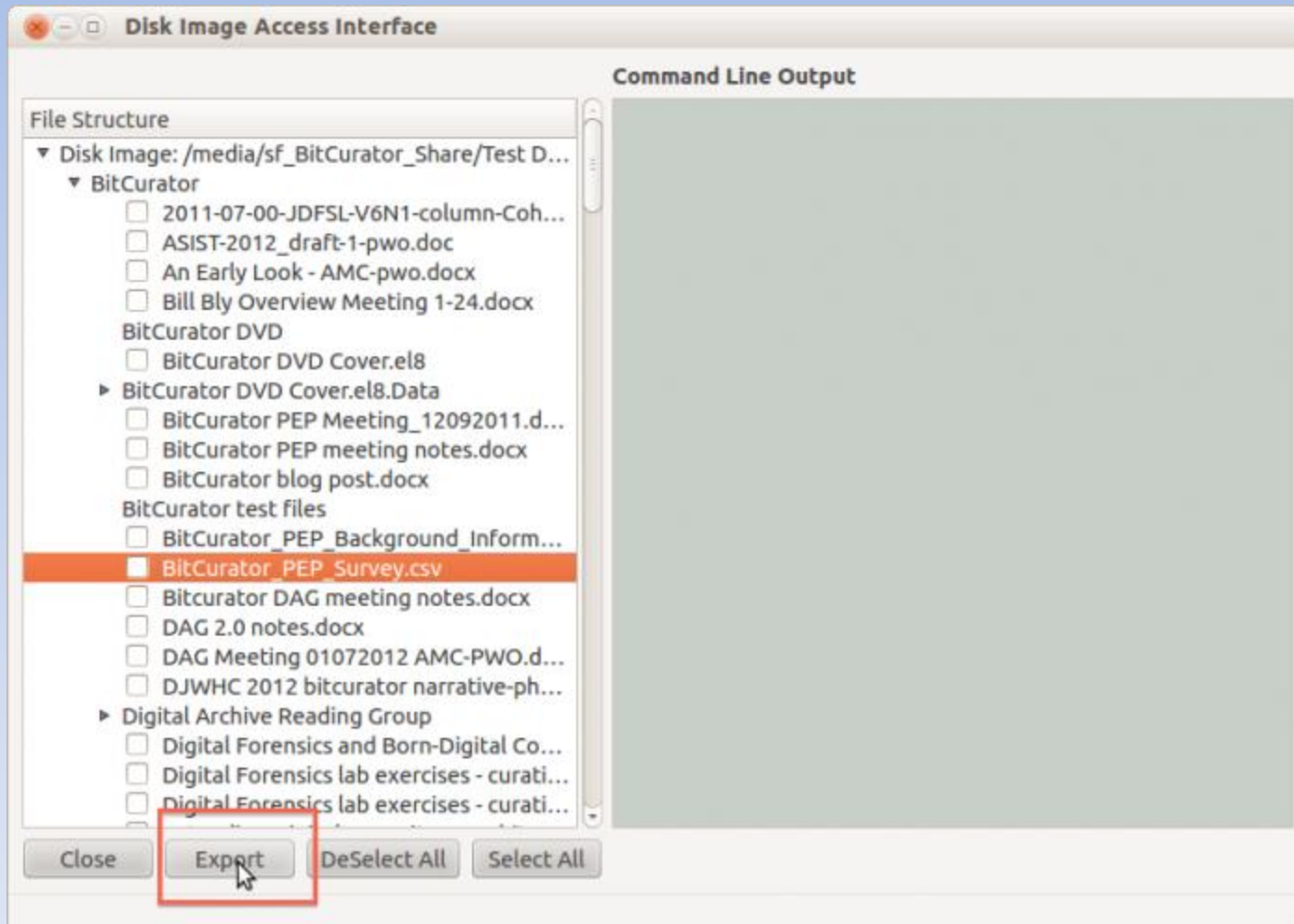
Feature www.archivists.org

Image

```
B Times New Roman Times New Roman Arial Identity Adobe Symbol Time
64 s New Roman Identity Adobe Identity Adobe Identity Adobe http://
120 hdl.handle.net/2027/spo.3310410.0010.105 http://pyrag.phparch.co
192 e/ http://matienzo.org/files/2010/11/UsingTheOCLCWorldCstAPIs.pd
256 f http://archives.org.uk/publications/archaeassociationmonthlye
320 agazine.html http://matienzo.org/files/2010/11/ARCArticleJune201
384 0.pdf http://www.oclc.org/research/publications/library/2010/201
448 0-04.pdf http://www.oclc.org/research/publications/library/2010/
512 2010-04.pdf http://www.oclc.org/research/publications/library/20
576 10/2010-04.pdf http://matienzo.org/publications http://matienzo
640 org/mailto:mark@matienzo.org http://www.nycarchivists.org/?p=67
704 2 http://www.slideshare.net/anarchivist/learning-to-take-learnin
768 g-to-give-linking-as-repurposing-metadata http://www.ibiblio.org
832 /saswiki/2007/index.php/Session:_Rethinking_Access_and_Descripti
896 ve_Practice_%26Session_503%29 http://www.ibiblio.org/saswiki/200
960 7/index.php/Session:_Rethinking_Access_and_Descriptive_Practice
1024 %26Session_503%29 http://www.archivists.org/publications/proceed
1088 ings/EADq10.asp http://www.archivists.org/publications/proceedin
1152 gs/EADq10.asp http://www.archivists.org/publications/proceedings
1216 /EADq10/Matienzo-EADq10.pdf http://www.archivists.org/publicatio
1280 ns/proceedings/EADq10/Matienzo-EADq10.pdf http://hdl.handle.net/
1344 10150/106496
```

Text Hex

3. Create access copy



Potential workflows



CJH, METRO, vendor, Partner creates disk image and content reports

CJH or Partner uses BitCurator to analyze data and prepare access copies

CJH, vendor, Partner stores disk image as digital preservation, makes access copies accessible in reading room or remotely

Additional Reading

*BitCurator wiki

[http://wiki.bitcurator.net/index.php?title=Main_Page]

*From Bitstreams to Heritage report

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

*You've Got to Walk Before You Can Run: First Steps for Managing Born-Digital Content Received on Physical Media

[<http://www.oclc.org/content/dam/research/publications/library/2012/2012-06.pdf?urlm=168601>]

Thank you!