

BitCurator: An Open-Source Project for Libraries and Archives that Takes Bitstreams Seriously

Cal Lee

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

September 17, 2013

**National Digital Stewardship Alliance
Infrastructure Working Group**

BitCurator 

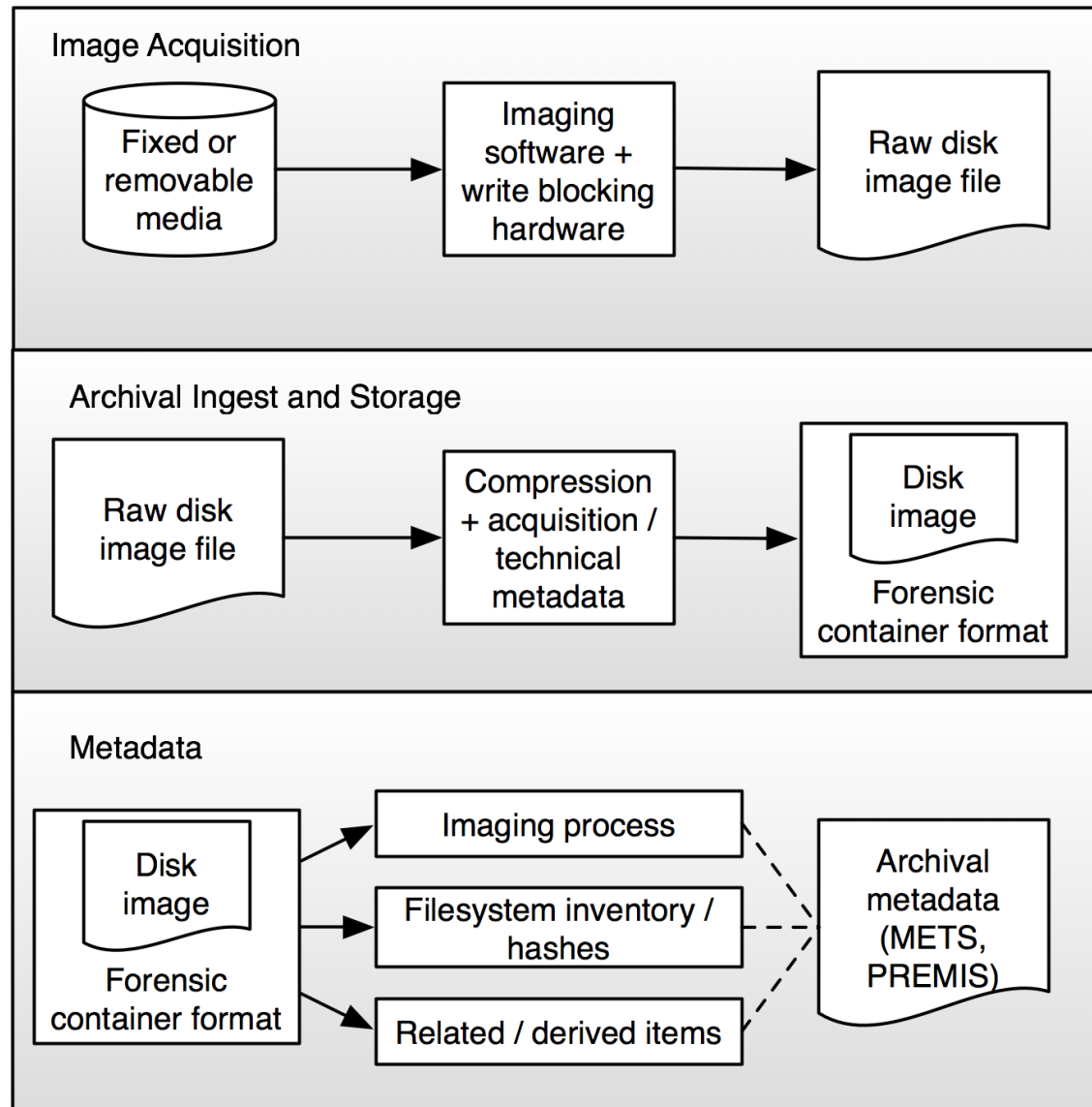


UNC
SCHOOL OF INFORMATION
AND LIBRARY SCIENCE

Why should we care about digital forensics?

- **Not** to solve crimes or catch malicious users
- Recognition of how data can be recovered when layers of technology fail or are no longer available
- Capturing information from places that are not immediately visible
- Ensuring that actions taken on files don't make irreversible changes to essential characteristics (e.g. MAC values)
- Attending to order of volatility – some types of data change more quickly and often than others
- Learning about available tools and techniques to deal with files
- Established practices for documenting acquisition and processing, so others will know what might have changed
- Overlap between technical knowledge required to do digital forensics and ad hoc acquisition of digital materials by libraries/archives

Storage Media Acquisition and Handling Profile for Digital Repositories*



*Woods, Kam, Christopher A. Lee, and Simson Garfinkel. "Extending Digital Repository Architectures to Support Disk Image Preservation and Access." In *JCDL '11: Proceeding of the 11th Annual International ACM/IEEE Joint Conference on Digital Libraries*, 57-66. New York, NY: ACM Press, 2011.

BitCurator

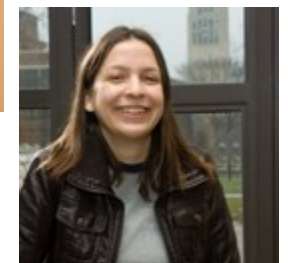
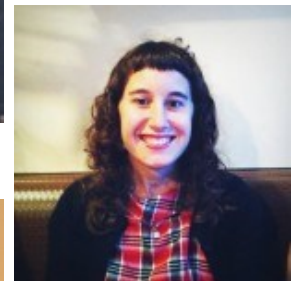
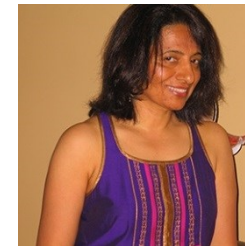
- Funded by Andrew W. Mellon Foundation
 - Phase 1: October 1, 2011 – September 30, 2013
 - Phase 2 – October 1, 2013 – September 30, 2014
- Partners: SILS at UNC and Maryland Institute for Technology in the Humanities (MITH)

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 Chassonoff, Project Manager
- Sunitha Misra, GA (UNC)
- Amanda Visconti, GA (MITH)



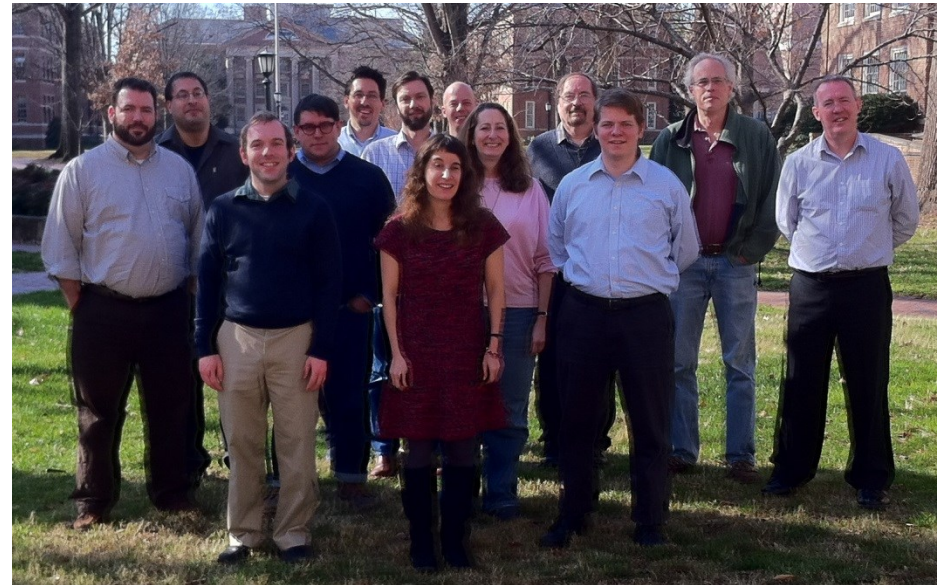
Two Groups of Advisors

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, Library of Congress
- Naomi Nelson, Duke University
- Erin O'Meara, Gates Archive
- Michael Olson, Stanford University Libraries
- Gabriela Redwine, Harry Ransom Center, University of Texas
- Susan Thomas, Bodleian Library, University of Oxford

Development Advisory Group

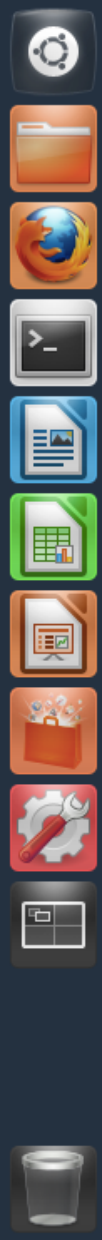
- Barbara Guttman, National Institute of Standards and Technology
- Jerome McDonough, University of Illinois
- Mark Matienzo, Yale University
- Courtney Mumma, Artefactual Systems
- David Pearson, National Library of Australia
- Doug Reside, New York Public Library
- Seth Shaw, University Archives, Duke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



BitCurator Environment*

- Bundles, integrates and extends functionality (primarily data capture and reporting) 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/>



Computer



home



Imaging Tools



Forensics Tools



Reporting and
Metadata Tools



Trash



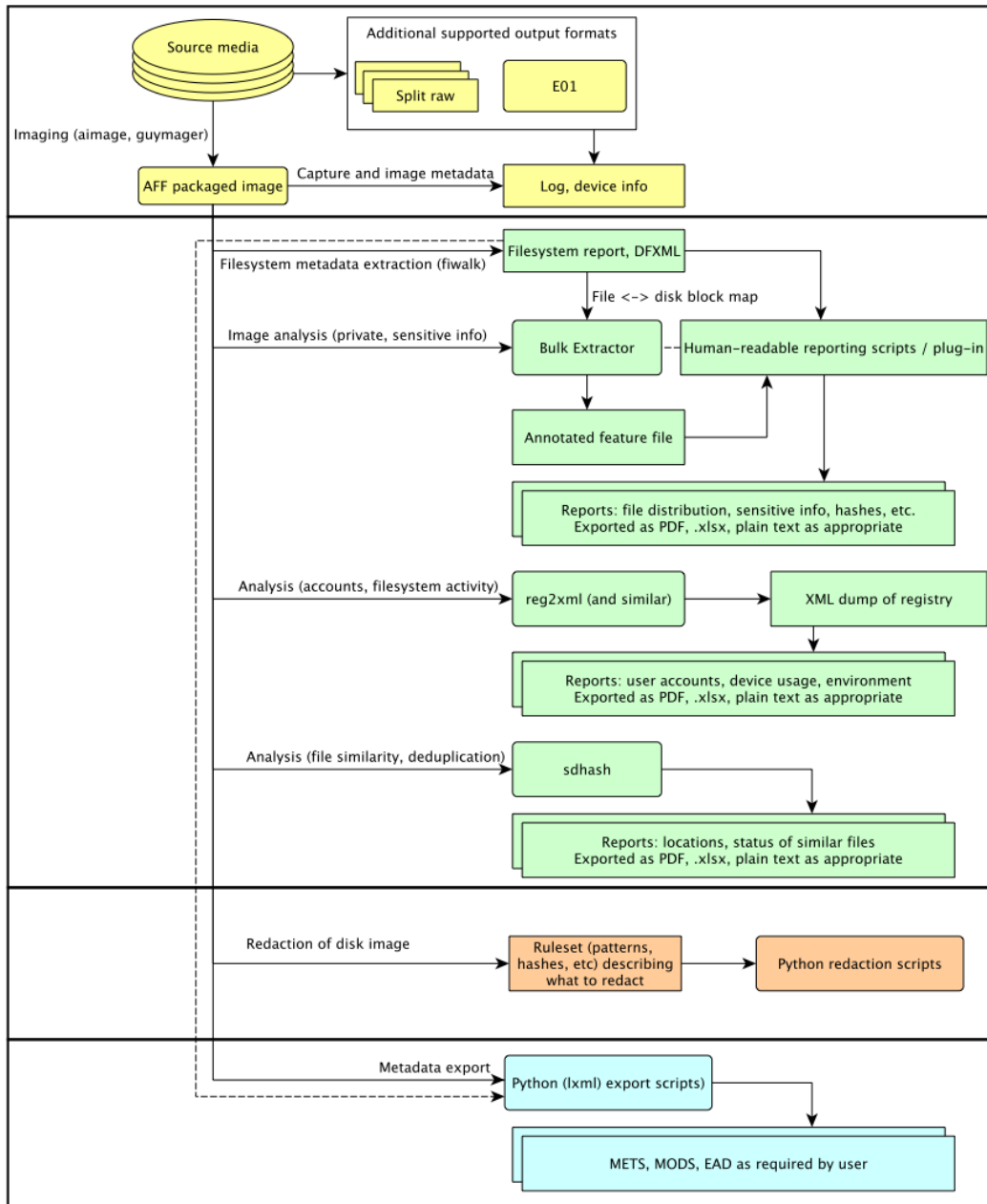
Documentation and
Help



Network Servers

BitCurator

BitCurator-Supported Workflow



Acquisition

Reporting

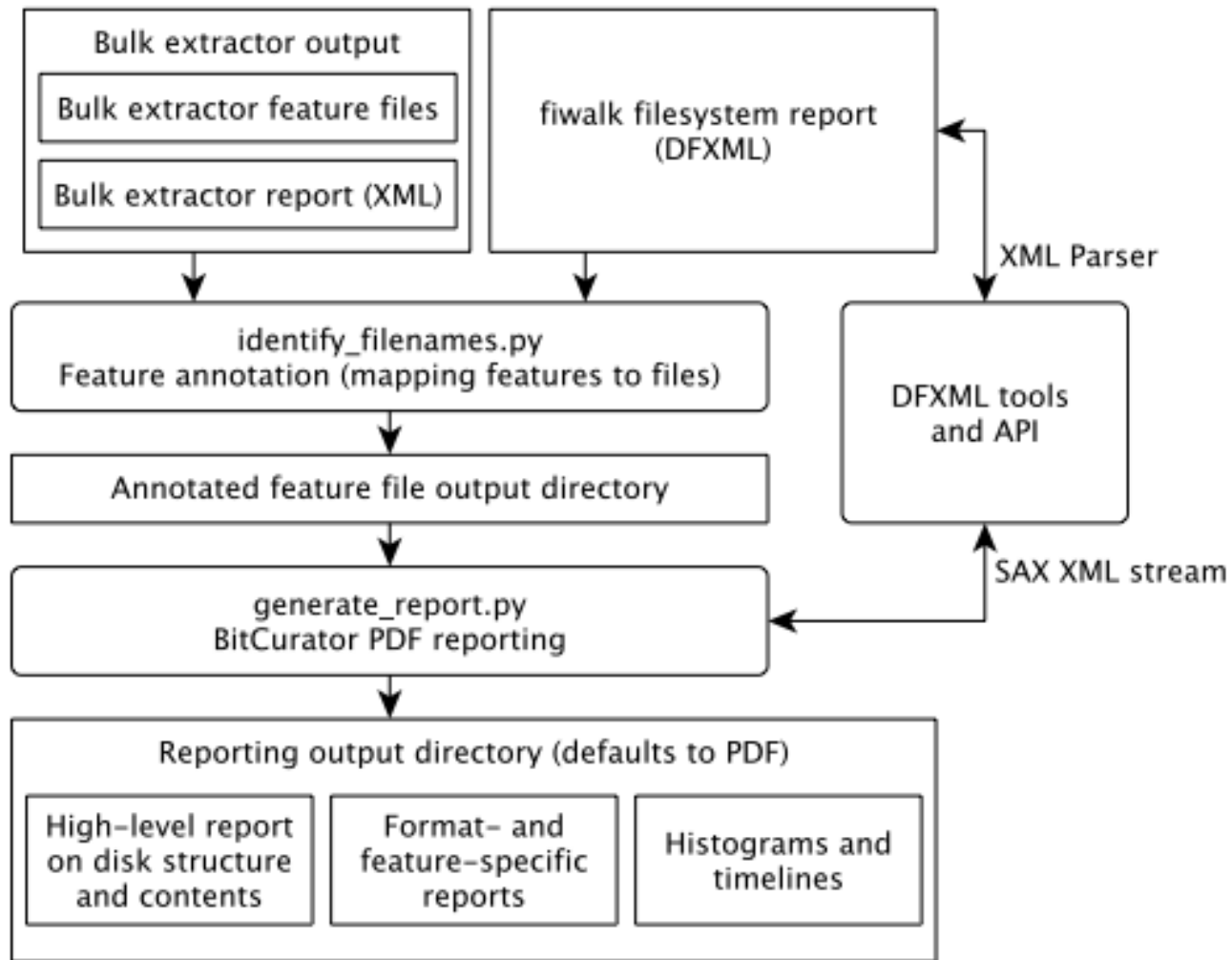
Redaction

Metadata export

- Acquisition
- Reporting
- Redaction
- Metadata Export

See: <http://bitcurator.net>

High-Level view of Metadata Generation and Reporting



See: Woods, Kam, Christopher Lee, and Sunitha Misra. "Automated Analysis and Visualization of Disk Images and File Systems for Preservation." In *Proceedings of Archiving 2013* (Springfield, VA: Society for Imaging Science and Technology, 2013), 239-244.

Documentation of Digital Forensics XML (DFXML) Elements

BitCurator-Test-0.1.7 [Running]

LibreOffice Calc

Documentation and Help

DFXML tag library v3.xlsx - LibreOffice Calc

File Edit View Insert Format Tools Data Window Help

Arial 11

A62 f(x) Σ = <compiler>

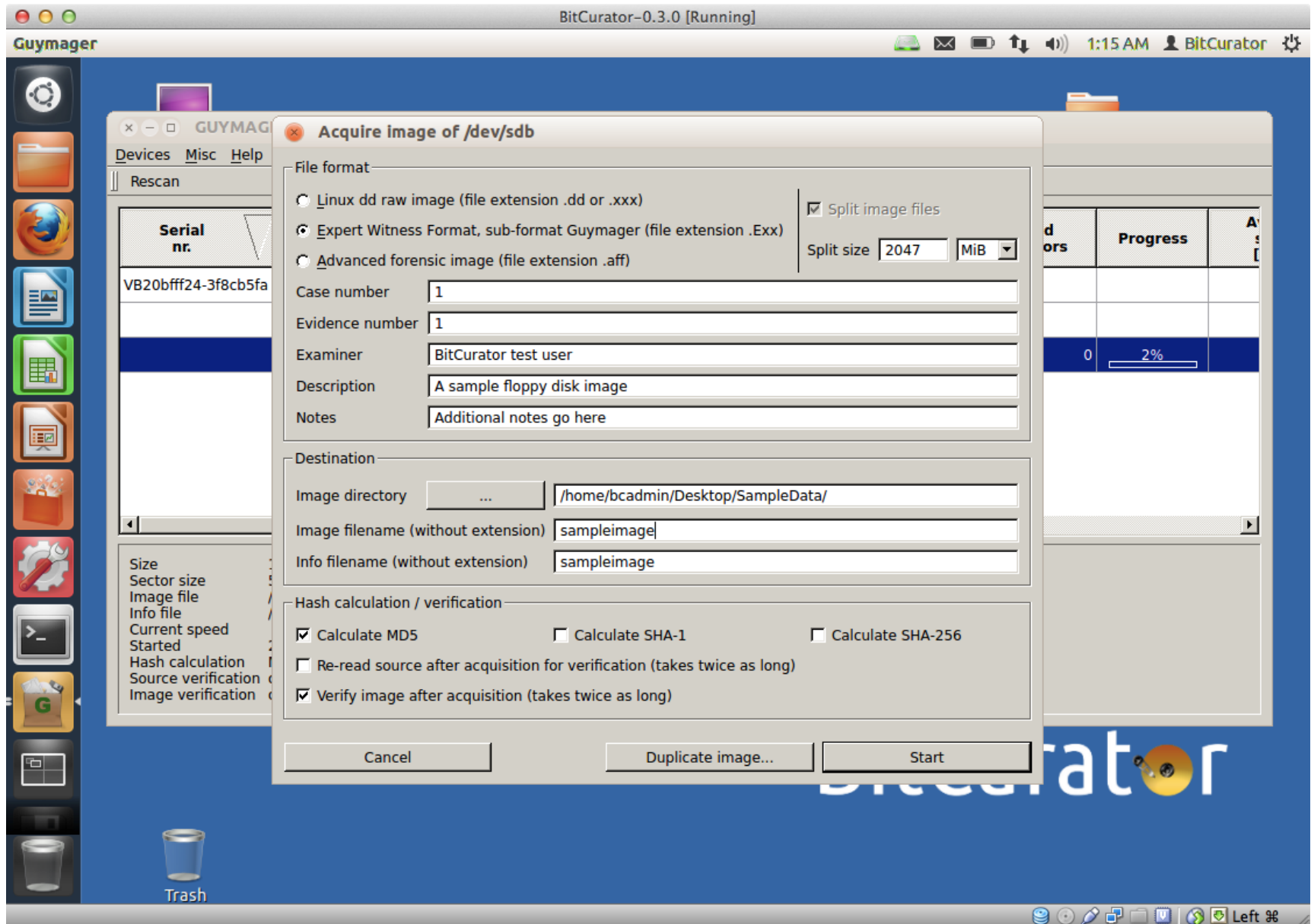
	A	B	C	D
1	Tag name	Element name	Description	May contain
2	<dfxml>	DFXML	Root element, marks the beginning and end of the DFXML metadata file. The <dfxml> element contains the primary elements reported in <i>fiwalk's xml</i> structure: <metadata>, <creator>, <source>, <volume>, and <runstats>.	<metadata>, <creator>, <source>, <volume>, <runstats>, <sectorsize>, <pagesize>, <acquisition seconds>
3	<metadata>	Metadata	The <metadata> tag provides header information that defines the metadata in the DFXML document. Includes namespace declaration, namespace schema location, and other information that is used to define the elements used in the XML file. These declarations provide information on the types of standardization schemes used to convey information in the DFXML document. The <metadata> tag may also contain high level descriptive information about the DFXML document rendered in Dublin Core (dc), in order to increase interoperability.	<dc:type>, <dc:creator>, <dc:title>, <dc:description>; for more information on Dublin Core element set, see (21).
4	<creator>	Creator	The Creator element provides documentation about the program and computing environment in which the disk analysis (or capture) take place. <Creator> includes tags documenting the program that initiated the capture creating the DFXML file, and other contextual information about the system on which	<program>, <version>, <build_environment>, <execution_environment>

fiwalk / bulkextractor / attributes / color code / Sheet2

Sheet 1 / 5 PageStyle_fiwalk STD Sum=0 100%

<http://www.bitcurator.net/2013/02/06/dfxml-tag-library/>

Acquiring Disk Images with Guymager



Exporting Filesystem Content Using fiwalk

The screenshot shows the BitCurator Reports dialog box in an Ubuntu Desktop environment. The dialog box has three tabs: "Fiwalk XML", "Annotated Features", and "Reports". The "Reports" tab is selected, displaying the following text: "Fiwalk produces a DFXML file showing the volumes, directories, and files contained within a disk image." Below this text are two input fields: "Image File" with the path "/home/bcadmin/Desktop/SampleData/sampleimage.E01" and "Output XML File" with the path "/home/bcadmin/Desktop/SampleData/sampleimage.xml". At the bottom of the dialog box are three buttons: "Cancel", "Close", and "OK". The background shows the Ubuntu Desktop with a blue background, a sidebar on the left with icons for Computer, home, Imaging Tools, Forensics Tools, and Additional Tools, and a desktop with icons for Documentation and Help, Network Servers, and Trash. The BitCurator logo is visible in the bottom right corner of the desktop.

BitCurator-0.3.0 [Running]

Ubuntu Desktop

1:32 AM BitCurator

Bitcurator Reports

Fiwalk XML Annotated Features Reports

Fiwalk produces a DFXML file showing the volumes, directories, and files contained within a disk image.

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

Output XML File
/home/bcadmin/Desktop/SampleData/sampleimage.xml

Command Line Output:

Cancel Close OK

tCurator

Viewing the Command Line Output

The screenshot displays the BitCurator Forensics GUI with a 'Bitcurator Reports' dialog box open. The dialog has three tabs: 'Fiwalk XML', 'Annotated Features', and 'Reports'. The 'Reports' tab is active, showing a text area with the following content:

Fiwalk produces a DFXML file showing the volumes, directories, and files contained within a disk image.

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

Output XML File
/home/bcadmin/Desktop/SampleData/sampleimage.xml

Command Line Output:

```
>> Command Executed for Fiwalk = ['fiwalk', '-f', '-X',  
'/home/bcadmin/Desktop/SampleData/sampleimage.xml',  
'/home/bcadmin/Desktop/SampleData/sampleimage.E01']  
  
>> Success!!! Fiwalk crated the following file(s):  
  
o /home/bcadmin/Desktop/SampleData/sampleimage.xml
```

The 'Command Line Output' section is circled in red. At the bottom of the dialog are 'Cancel', 'Close', and 'OK' buttons. The background shows a desktop environment with various icons and a taskbar at the bottom.

DFXML Output from fiwalk

BitCurator-0.3.0 [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>4.1.0</version>
    -<build_environment>
      <compiler>GCC 4.6</compiler>
      <library name="afflib" version="3.7.1"/>
      <library name="libewf" version="20130416"/>
    </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>2013-07-20T05:34:37Z</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>
    <sector_size>512</sector_size>
    <block_size>512</block_size>
    <ftype>2</ftype>
  </volume>
</dfxml>
```

Left 36

DFXML for a Specific File

```
<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>
  <ctime>1257729636</ctime>
  <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'>
    6fb2a38dc107eachb41cf1656e899cf70
  </hashdigest>
  <hashdigest type='SHA1'>
    4eee44b18576e84de7b163142b537d2fe6231845
  </hashdigest>
</fileobject>
```

Identifying “Features” of Interest in Disk Images or Directories

Bulk Extractor

Bulk Extractor Viewer

File Edit View Tools Help



Highlight:

Reports Feature Filter

Feature File None

Referenced Feature
Referenced Feature

Run bulk_extractor

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

Tuning Parameters

- Use Context Window Size
- Use Page Size
- Use Margin Size
- Use Min Word Size
- Use Max Word Size
- Use Block Size
- Use Number of Threads

Scanner Controls

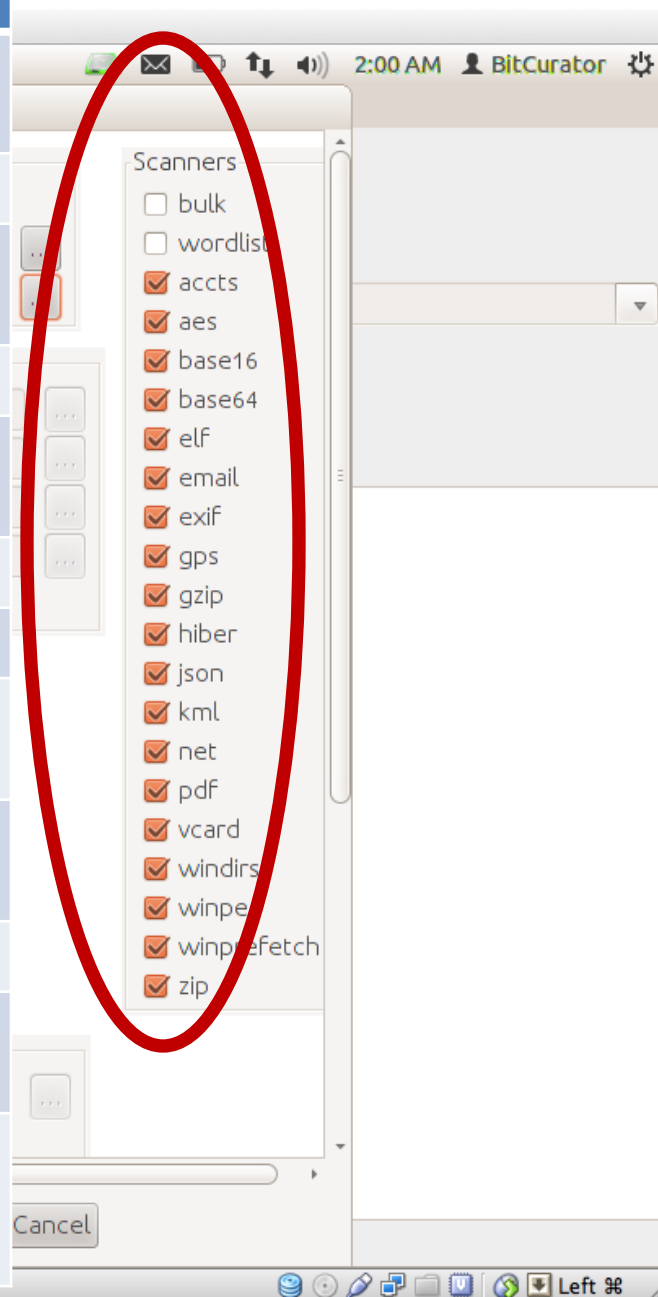
- Use Plugin Directory ...
- Use Scan Option Name

Scanners

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

Restore Defaults Start bulk_extractor Cancel

Scanner	Description
scan-accts	Looks for phone numbers, credit card numbers, etc
scan_base64	Decodes BASE64 text
scan_kml	Detects KML (Keyhole Markup Language) files – used to identify geographic locations
scan_gps	Detects XML from Garmin GPS devices
scan_aes	Detects in-memory AES (Advanced Encryption Standard) keys from the key schedules
scan_json	Detects JavaScript Object Notation files
scan_exif	Detects EXIF structures from JPEG files
scan_zip	Detects and decompresses ZIP files and zlib streams
scan_gzip	Detects and decompresses GZIP files and gzip streams
scan_pdf	Extracts text from some kinds of PDF files
scan_hiber	Detects and decompresses Windows hibernation file fragments
scan_winprefetch	Detects and extracts fields from windows prefetch fields from Windows prefetch files and file fragments



Bulk Extractor Viewer

File Edit View Tools Help



Highlight: Match case

Reports Feature Filter Match case Navigation

bulk_extractor Scan

Image File sampleimage.E01
 Feature Directory bulk-extractor-output

Progress **Done**
 bulk_extractor scan completed. See Status, below, for details.

Options

```
'bulk_extractor'
'-o'
'/home/bcadmin/Desktop/sampleimage.E01'
'/home/bcadmin/Desktop/bulk-extractor-output'
```

Report is Ready

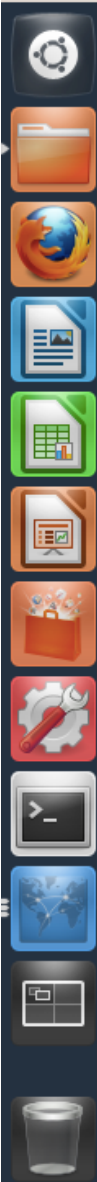
bulk_extractor has completed.
 Report bulk-extractor-output has been opened and is ready for viewing.

OK

Status

```
Elapsed time: 0.4985 sec.
Overall performance: 2.958 MBytes/sec.
Total email features found: 0
Done.
```

Text Hex



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

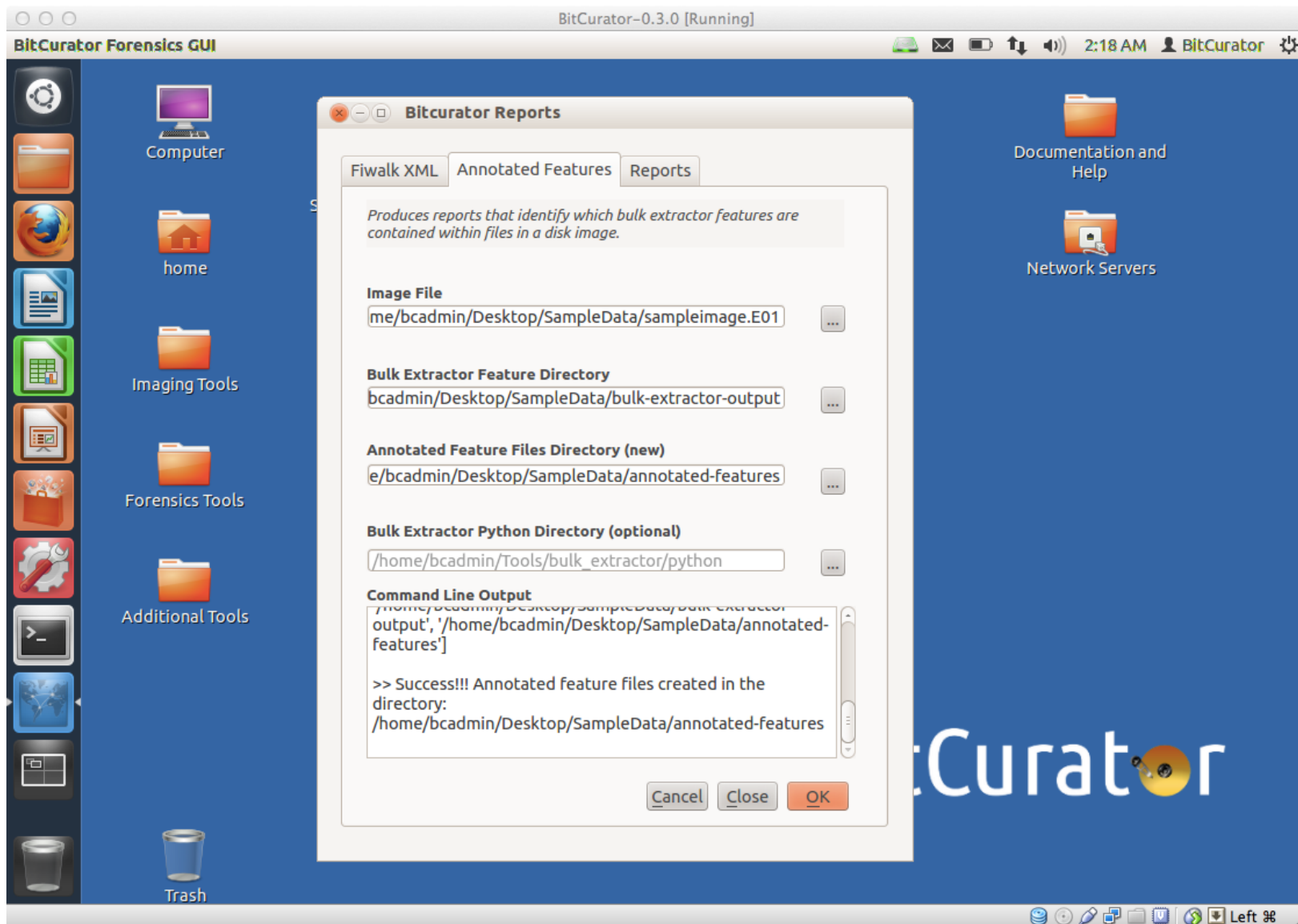
Below the histogram is a table of "Referenced Feature File" and "Referenced Feature":

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

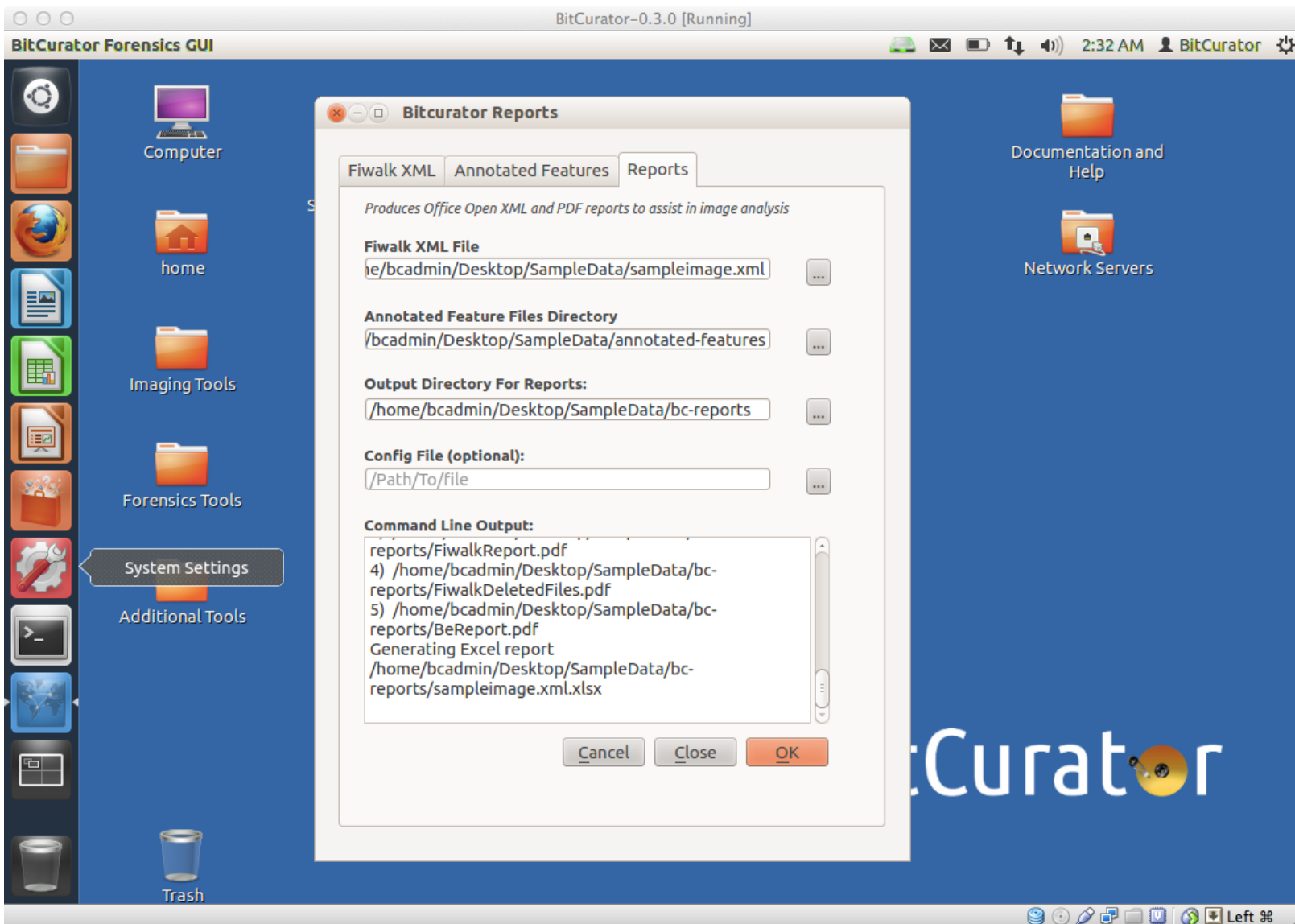
The right pane shows a specific instance in context, titled "Image File sampleimage.E01". The feature file is "email.txt" and the feature path is "42273785". The feature is "privacy@Motorola.com". The image content is a snippet of text from a document, likely a privacy policy, which includes the following text:

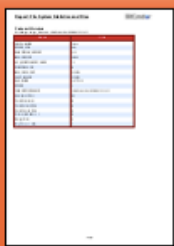
your credit card number, so this information can only be viewed by Motorola. Motorola uses Secure Sockets Layer (SSL) encryption technology, the highest level of security on the Internet. The SSL protocol provides server authentication, data integrity, and privacy on the Web. This security measure helps 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

Matching Bulk Extractor Output (Based on Byte Offsets) to fiwalk Output (Based on Filesystem Location)



Generating BitCurator Reports





1



2



3

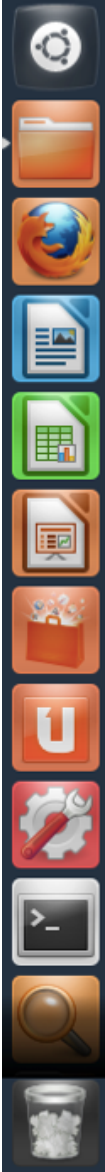
Report: File System Statistics and Files

BitCurator

Technical Metadata

Disk Image: image_filename: charlie-work-usb-2009-12-11.aff

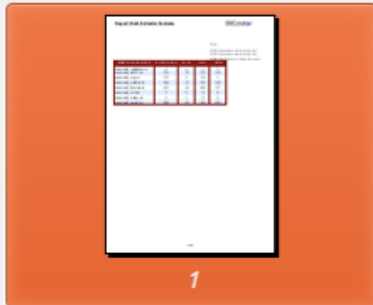
Feature	Value
SECTORSIZE	1024
FTYPE STR	ntfs
PARTITION OFFSET	512
BLOCK SIZE	4096
ACQUISITION SECONDS	73
FIRST BLOCK	0
BLOCK COUNT	258559
LAST BLOCK	258558
PAGESIZE	16777216
FTYPE	1
IMAGE FILENAME	charlie-work-usb-2009-12-11.aff
Number of Files	128
Total Directories	23
Total Deleted Files	0
Total Unused Files	0
Files with Nlinks > 1	0



BeReport.pdf

↑ Previous ↓ Next 1 (1 of 1) Fit Page Width

Thumbnails



1

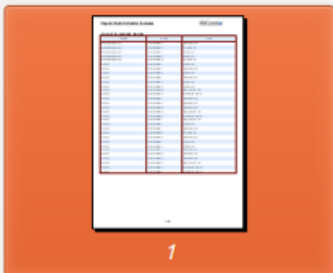
Report: Bulk Extractor Features

BitCurator

*Note:**FIUF: Total features unallocated to files**FIUF: Total features unallocated to files**FICR: Total features in compressed regions*

Bulk Extractor Report Files	Feature Instances	FLTF	FIUF	FICR
annotated_telephone.txt	5	4	1	2
annotated_rfc822.txt	258	39	219	110
annotated_zip.txt	127	8	119	3
annotated_windirs.txt	466	13	453	180
annotated_domain.txt	653	48	605	317
annotated_exif.txt	2	2	0	0
annotated_winpe.txt	1	1	0	0
annotated_email.txt	500	42	458	224

Thumbnails



1



2



3



Report: Bulk Extractor Features

BitCurator

Feature File: annotated_email.txt

Filename	Position	Feature
Email\Charlie_Email.zip	2299 1360-ZIP-115	charlie@m57.biz
Email\Charlie_Email.zip	2299 1360-ZIP-148	terry@m57.biz
Email\Charlie_Email.zip	2299 1360-ZIP-55	pat@m57.biz
Email\Charlie_Email.zip	2299 1360-ZIP-648	pat@m57.biz
Email\Charlie_Email.zip	2299 1858-ZIP-130	terry@m57.biz
Unknown	2299 1858-ZIP-37	pat@m57.biz
Unknown	2299 1858-ZIP-97	charlie@m57.biz
Unknown	2299 2313-ZIP-129	pat@m57.biz
Unknown	2299 2313-ZIP-55	charlie@m57.biz
Unknown	2299 2313-ZIP-712	pat@m57.biz
Unknown	2299 2313-ZIP-732	pat@m57.biz
Unknown	2299 2834-ZIP-108	alix.pery@yahoo.com
Unknown	2299 2834-ZIP-129	rubinfritz31@mail.com
Unknown	2299 2834-ZIP-45	charlie@m57.biz
Unknown	2299 3168-ZIP-133	charlie@m57.biz
Unknown	2299 3168-ZIP-204	charlie@m57.biz
Unknown	2299 3168-ZIP-226	alix.pery@yahoo.com
Unknown	2299 3168-ZIP-247	rubinfritz31@mail.com
Unknown	2299 3168-ZIP-51	alix.pery@yahoo.com
Unknown	2299 3587-ZIP-37	pat@m57.biz
Unknown	2299 3587-ZIP-97	charlie@m57.biz

Nautilus Scripts

- Scripts that can be run using the GNOME file manager called Nautilus (Linux analog to Windows Explorer or Mac OS X Finder)
- Can be used in the BitCurator environment or your own Linux environment

MD5 Hashes of Files (Nautilus Script)

The screenshot displays a Linux desktop environment. On the left, the Nautilus file manager shows a directory tree with the following structure:

- testdocs
 - Presentations
 - Public
 - RandallAir
 - Research
 - Agreement
 - BakSVN
 - BCiaB
 - DOMEX
 - FromLaptop
 - AFF4
 - bak3
 - CALFILES
 - CRADLE
 - DCC2010
 - Dissertation
 - docXtractor
 - DPapers
 - findntuser
 - FromStore
 - JOBAPP
 - M57
 - NewHomepage
 - NSFProp2009
 - PDA2011
 - Scripting and Programming
 - ssh-bak
 - testdocs

The gedit text editor window, titled "md5Results.md5 (~/Research/FromLaptop/testdocs) - gedit", displays the following content:

```
k210f06f75bfc5ed09ef853280010bbd /home/kam/Research/FromLaptop/testdocs/archaeol6-69760_1.docx
bacd0311caaafdd0de8d7c38bf15d64 /home/kam/Research/FromLaptop/testdocs/archaeol21-81016_2.docx
77a39ca3ac7124c7374ad146be17c1b8 /home/kam/Research/FromLaptop/testdocs/archaeol21-81048_1.docx
a3d9c64b823c95a59c9309de6d3842b0 /home/kam/Research/FromLaptop/testdocs/archaeol21-81740_1.docx
7f76beba6d6c1bb30c14a6a35befad3 /home/kam/Research/FromLaptop/testdocs/archaeol21-83991_1.docx
b6279b34c082c6abfa9686e90fb91798 /home/kam/Research/FromLaptop/testdocs/archaeol21-85873_2.docx
b393d82ca450dfe85dbbd82fbfaa4e82 /home/kam/Research/FromLaptop/testdocs/archaeol21-105745_1.docx
7e8933260fea7f3ade57f255f8c102a4 /home/kam/Research/FromLaptop/testdocs/cameron1-88343_1.docx
9ec5781e7d6ca810da7ae3da43c3e1e7 /home/kam/Research/FromLaptop/testdocs/contexto1-66217_1.docx
d08bf83b14c92221b8f7147bec7973c5 /home/kam/Research/FromLaptop/testdocs/contexto1-69471_1.docx
3c85ecd7e128bf4b3ccbba8da3993010 /home/kam/Research/FromLaptop/testdocs/contexto1-79006_1.docx
61ed85e3d39fa22fc43b74ae30836f96 /home/kam/Research/FromLaptop/testdocs/contexto1-80765_1.docx
b19b44cd870acd5d546d829a7b7f2e0d /home/kam/Research/FromLaptop/testdocs/contexto1-85055_1.docx
b5aea481497bf36f936e545653a6a3c2 /home/kam/Research/FromLaptop/testdocs/dennisp1-58597_2.docx
51bf7f5bc623c5ff9db88a154541a42f /home/kam/Research/FromLaptop/testdocs/dennisp1-70846_3.docx
adf11d9c854486c70db3746e252e8e18 /home/kam/Research/FromLaptop/testdocs/dennisp1-76116_2.docx
c28d100f1f064de7fc365a660e470789 /home/kam/Research/FromLaptop/testdocs/dennisp1-83743_1.docx
b9345190327e0debb113e287c258ea7a /home/kam/Research/FromLaptop/testdocs/dennisp1-91925_1.docx
b224123f788d7c0fd7a73a2dfb34a219 /home/kam/Research/FromLaptop/testdocs/national3-99879_1.docx
b4bce311d0c97b4009b3cdc6d50529ae /home/kam/Research/FromLaptop/testdocs/quaterna1-62216_1.docx
e245d76ce027cd79309024ab9e83c6d1 /home/kam/Research/FromLaptop/testdocs/scotiaar1-104626_2.docx
bb0eccd64d98db2660c909c276f30398 /home/kam/Research/FromLaptop/testdocs/scotiaar1-104630_1.docx
5b7bb2476374212e849e1513923c670d /home/kam/Research/FromLaptop/testdocs/scotiaar1-104632_2.docx
638cc2ed4562380fdad3bb87d570250a /home/kam/Research/FromLaptop/testdocs/scotiaar1-104662_10.docx
e440d9f4f6eaa9dc188f9b5fb8ad23c7 /home/kam/Research/FromLaptop/testdocs/scotiaar1-106279_2.docx
4ed53bc15578a68f1ec8cc14ac2a9503 /home/kam/Research/FromLaptop/testdocs/scotiaar1-106334_1.docx
bb1e322e94e13bc91a57be59f7a905ab /home/kam/Research/FromLaptop/testdocs/universi1-84908_1.docx
bb1e322e94e13bc91a57be59f7a905ab /home/kam/Research/FromLaptop/testdocs/universi1-84908_2.docx
5f8ef69582e68b68d2ad9ecf8e37e564 /home/kam/Research/FromLaptop/testdocs/universi1-84914_1.docx
```

The status bar at the bottom of the gedit window shows "Plain Text", "Tab Width: 8", "Ln 1, Col 1", and "INS". The desktop taskbar at the bottom shows several document icons, with the current document "md5Results.md5" selected (12.1 kB).

Quick Start Guide

Most recent version always available at:

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

BitCurator Quick Start Guide v0.3.4

Last updated: September 8, 2013



UNC
SCHOOL OF INFORMATION
AND LIBRARY SCIENCE

MITH MARYLAND INSTITUTE FOR
TECHNOLOGY IN THE HUMANITIES

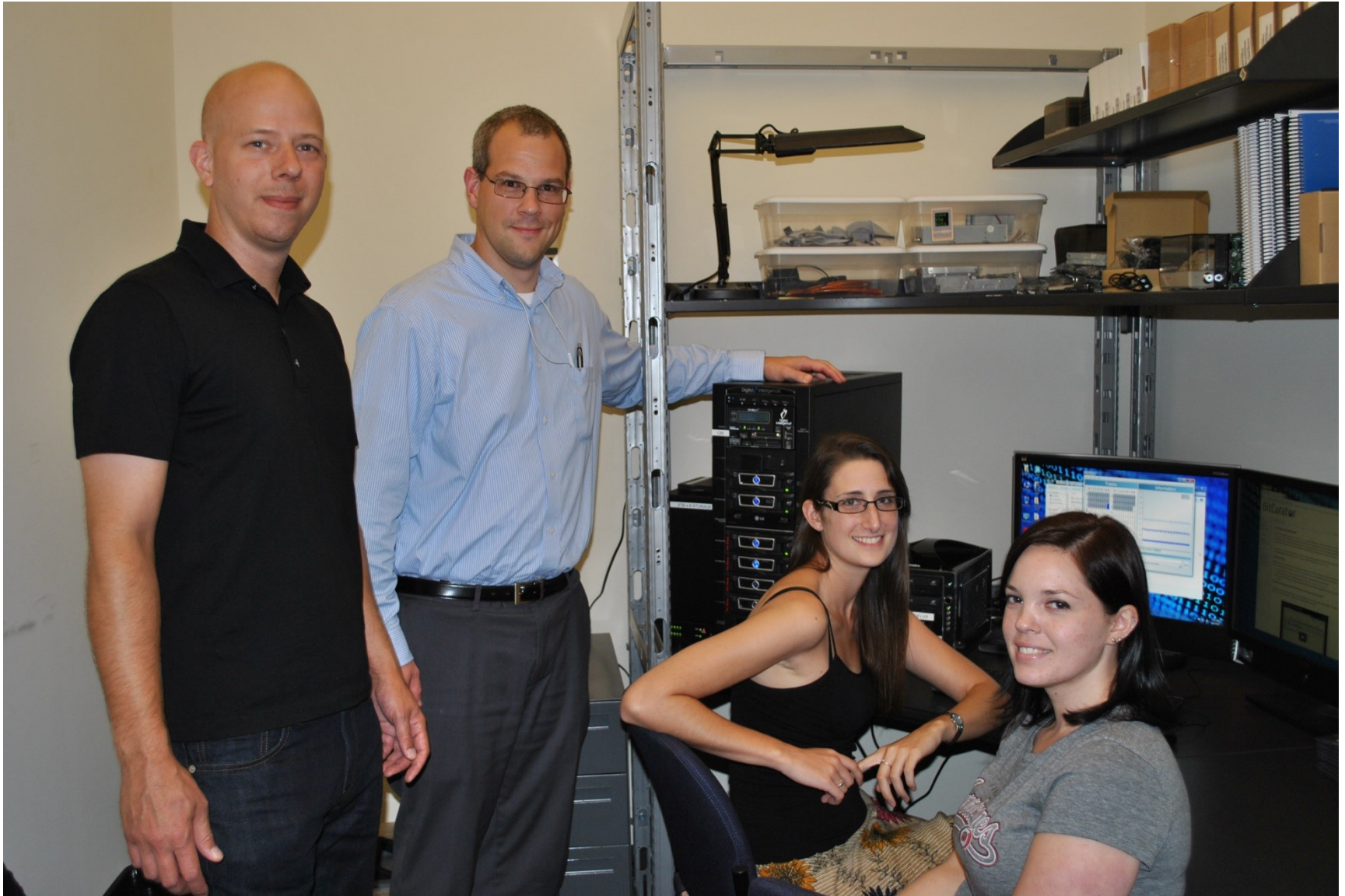
Open Source Software Strategy

- Code released under GPL, v3 – available through GitHub
- Existing code that we incorporate is generally GPL or public domain (government products)
- Packaging elements of the code to be integrated into other environments (e.g. Archivematica)
- Regular contact with individuals and organizations responsible for other development efforts
- Phase 2 of BitCurator involves both further development and significant community engagement, including hiring of Community Lead (Porter Olsen)

Engagement Efforts

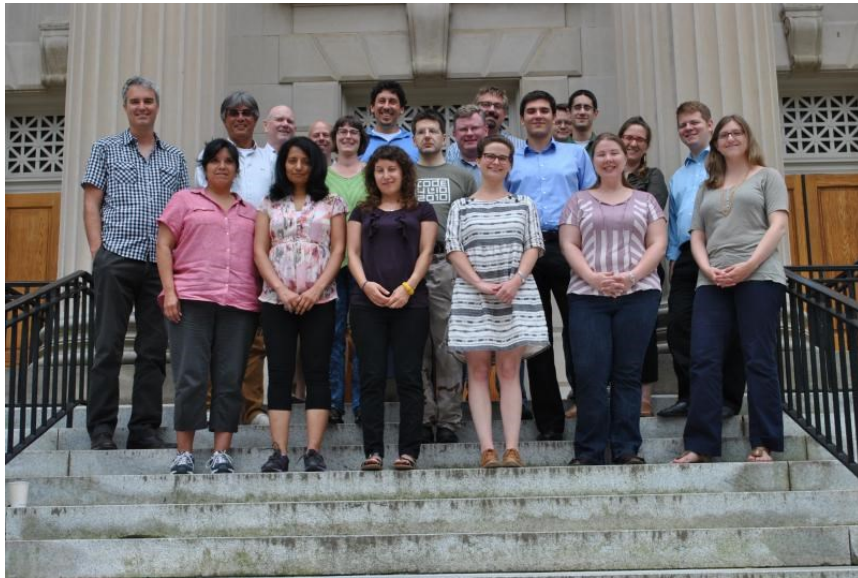
- Oodles of conference papers, presentations, posters, demos, workshops, tutorials, webinars
- BitCurator users mailing list
- Web site, wiki, social media
- Professional education offerings:
 - Two-day class for Digital Archives Specialist (DAS) curriculum of Society of American Archivists (SAA) – Cal Lee and Kam Woods
 - “Born-Digital Materials: Theory & Practice” Course for Rare Book School – Matt Kirschenbaum and Naomi Nelson
 - Courses for graduate students at UNC SILS

Dedicated Digital Forensics Lab at UNC SILS



Hackathon: Tackling Real-World Collection Challenges with Digital Forensics Tools and Methods

- June 3-5, 2013, Chapel Hill, NC
- Collaboration between OPF, UNC SILS and Library of Congress
- Two categories of participants: collection owners, developers
- Participants from Germany, Netherlands, US



Institutions represented:

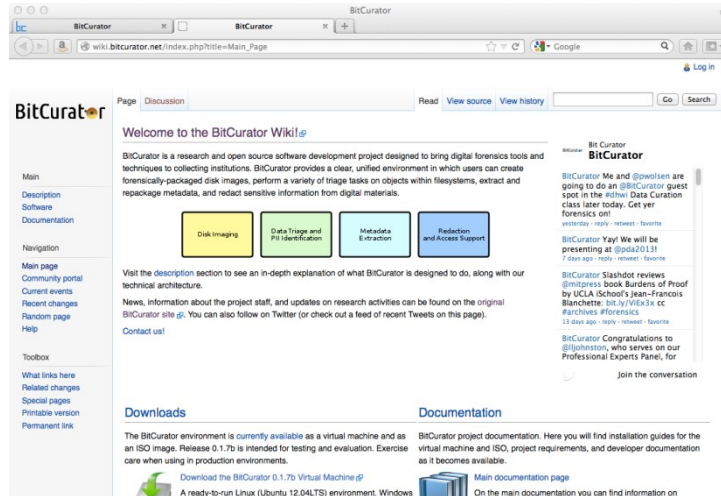
- Baker Library, Harvard Business School
- Duke University
- Florida State University
- Massachusetts Institute for Technology
- Library of Congress
- National Snow and Ice Data Center
- New York Public Library
- Open Planets Foundation
- Sloane Art Library, UNC
- Stanford University Libraries
- UNC Libraries
- University of Freiburg
- UNC SILS
- Wake Forest University

<http://sils.unc.edu/events/2013/hackathon-opf>

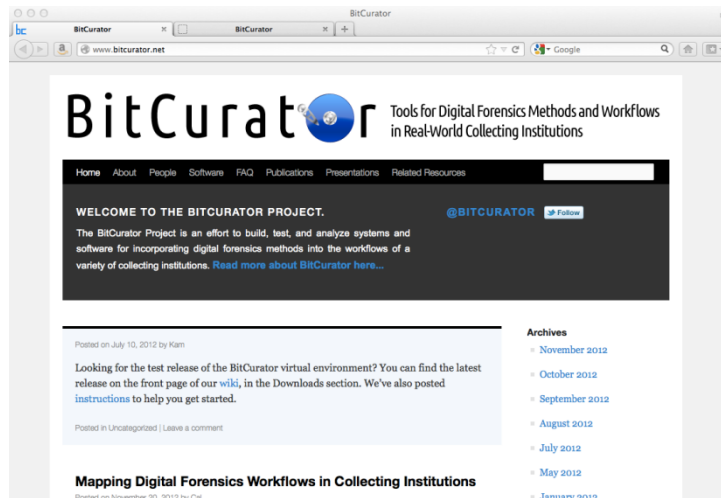
Sustainability Plans

- The primary model we're exploring is a membership-based consortium.
- Looking at existing consortia for potential models (e.g. ArchivesSpace, IIPC)
- Members of Consortium would pay an annual fee, which could entitle them to member benefits such as:
 - Access to dedicated personnel who can assist with integration of BitCurator into their institutions' particular workflows
 - Ability to submit development requests
 - Priority notification of new code releases
 - BitCurator online training events
- This is all very tentative. Draft sustainability plan under development now.

Sources for BitCurator Information:



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