• BitCurator packages forensics software (in a Linux virtual machine) to augment digital curation workflows in real-world collecting institutions:

- Reliable acquisition of bitstreams from digital media
- Scalable forensic analysis of disk images / heterogeneous collections of documents
- Metadata acquisition from common file systems
- Preservation metadata and finding aid metadata export
- Data visualization for complex data sources
- Redaction of private and sensitive information
- Identification and removal of duplicate files
Why it’s different

Commercial forensics tools are typically **expensive**, and not always as **reliable** as you’d expect.

Lots of things important to **collecting institutions** that aren’t a **priority** for forensics developers: **access**, **redaction**, **preservation metadata**, etc…

BitCurator packages open source forensics tech into a simple-to-use virtual environment to help you avoid getting bogged down.
The cruising altitude view

Write-blocked physical media acquisition

Creating Digital Forensics XML to describe disk images

Identifying potentially private information

Selective data redaction
What’s new in the past year

BitCurator has had 15 public releases (4 major releases) in the past year. Some highlights (as of 0.5.8):

- BitCurator-specific GUI to assist collections professionals running complex forensic tools.
- One-click mounting of forensically-packaged disk images
- Improved metadata export facilities (PDF, .xlsx, and visualizations)
- PREMIS metadata generation for forensic events and objects
- DFXML improvements and participation in the DFXML working group
- Duplicate file detection and removal
- Byte-level redaction of pattern-matched features (e.g. potentially sensitive information)
Interacting with disk images

Right-click on any forensically-packaged disk image to mount.
Interacting with disk images

![Ubuntu Desktop interface with BitCurator-0.5.8 running, highlighting a disk image labeled 'jo-work-usb-2009-12-11_vol2'.]
Interacting with disk images
One-click tool to:
1. Analyze file systems and export metadata
2. Identify interesting and potentially sensitive features
3. Generate visualizations
4. Automatically generate PREMIS object and event descriptions
Forensic analysis events with UUIDs for each (automatic) step in BitCurator tool (using custom BitCurator PREMIS generation code)
File system metadata...simplified

Per-partition file items with file type, MAC times, MD5 and SHA1 hashes automatically generated via BitCurator GUI
Things can still get messy...

FORENSICS TOOL CRASHED FOR NO REASON

SEALING EVERYTHING, CALLING OTHER CATS

Image credit: http://www.quickmeme.com/
...and we know there are some issues

- Main forensic workflow does not support HFS volumes (only HFS+ - but we have other tools that can help)
- Disk image mount scripts fail on more exotic disks (multiple volumes)
- Really nice GUIs like Autopsy don’t run well (or at all) in Ubuntu LTS (12.04) without some major fiddling.
- VirtualBox doesn’t have USB 3.0 support...yet.
- Distributing large VMs is a bandwidth hog. (we’ve partnered with iBiblio to alleviate this)
- Preservation metadata is easy to generate but hard to make useful for those with limited infrastructure
- Likewise for EAD
- Redaction of private and sensitive information is not a major use case for forensics investigators. We’re rolling our own.
We’re making it easier to upgrade BitCurator and use modules individually.

Scripts with simple YES/NO questions to step you through pulling software from our Git repo, upgrading, and installing dependencies.
Phew! Time to go use it!

Software and docs:  http://wiki.bitcurator.net/

News and blog:  http://www.bitcurator.net/

Source code:  https://www.github.com/kamwoods/bitcurator

The Andrew W. Mellon Foundation