LOW-RISK DATA EXTRACTION FROM PHYSICAL MEDIA

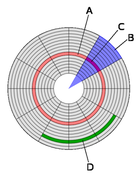
Many of the methods currently used to extract data from fixed and removable digital media (including floppy and hard disks, CD-ROMs and other optical media, and flash drives) are error-prone and labor intensive. One goal of the BitCurator project is to provide and demonstrate the hardware, software tools, and expertise required to efficiently extract bitstreams (raw disk images) from physical media in a manner that preserves all of the context necessary for future access.

Devices and media often arrive at collecting institutions in poorly documented or inconsistent states. Forensic extraction of disk images can assist both in preserving the contents and in establishing context and interrelationships between devices and users.

Even simple actions, such as plugging a device into a host computer for processing, can result in irrevocable changes to writeable filesystems on that device. Hardware write blockers are can be used to ensure no changes are inadvertently made to source media.

DATA TRIAGE, ANALYTICS, AND REPORT GENERATION

A number of powerful open source tools have been developed by the digital forensics community to process and extract information from disk images. However, few of these tools are currently user friendly. BitCurator will provide "one pass" analytics to aggregate the output of these tools into simple, human-readable reports that identify sensitive information, potential preservation problems, and describe the contents of the filesystem through visualizations.



Forensically packaged bitstreams contain metadata describing the acquisition process and data from unused portions of the disk in addition to filesystem contents. This information can assist in making informed decisions about how to process the materials, and provide valuable provenance and chain of custody support.

Open source digital forensics tools developed by Garfinkel, Carrier, and others provide scalable, high-performance processing of digital materials. Adapting these tools to the needs of collecting institutions requires interface revision and additional automation for users without forensics experience.

Human readable reports are necessary for data triage and rapid response. These reports should be easily customizable; providing information of file type distribution, private, sensitive, and protected data, and users, environment, and filesystem activity.

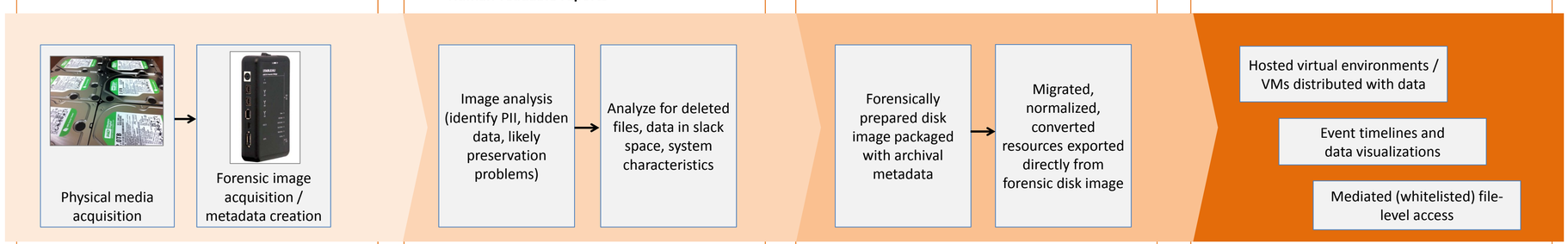
- Users, environment, and filesystem activity
- File type distribution
- Private, sensitive, and protected data

- Bulk Extractor
- Fiwalk
- The Sleuth Kit
- DFXML
- BitCurator plugins + wrappers

BitCurator

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

FORENSIC AUGMENTATION OF DIGITAL CURATION WORFLOWS



Reducing risk in the acquisition process

Use of hardware write-blockers

Acquisition metadata

Compression and storage of large disk image files in forensic container formats

Identification and redaction or protection of private and sensitive information

Repair of damaged filesystems (log recovery, extraction of data from truncated files,)

Locating hidden and "lost" data in unallocated space on storage devices

Human-readable reports

Object and environment monitors

Metadata crosswalks (Digital Forensics XML to archival metadata)

Dependency notification

Creating "views" of the raw filesystem; preparing access control lists for local and remote access

Sandboxed VMs with dynamic permissions overlays to provide access at varying levels of authorization

FIND OUT MORE!

Our blog, project events, links to software sources, and information on project personnel.

http://www.bitcurator.net/

@bitcurator

UNC School of Information and Library Science | MITH Maryland Institute for Technology in the Humanities | The Andrew W. Mellon Foundation | DCC

7th International Digital Curation Conference
December 6-8, 2011 – Bristol, UK

Christopher A. Lee (PI) and Kam Woods (Technical Lead)
Research performed under the direction of C. Lee at UNC Chapel Hill and Matthew Kirschenbaum (Co-PI) at the Maryland Institute for Technology and the Humanities, University of Maryland