Disk Image Repository Considerations

Digital Forensics Research Workshop

August 1-3, 2011

New Orleans, LA

Cal Lee

University of North Carolina at Chapel Hill

The Andrew W. Mellon Foundation

UNC
SCHOOL OF INFORMATION
AND LIBRARY SCIENCE

Archivists are pretty good at dealing with this kind of stuff:

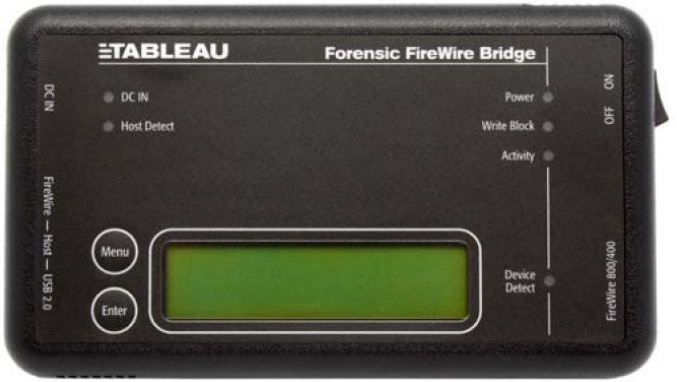
Source: The Processing Table: Reflections on a manuscripts internship at the Lilly Library. https://processingtable.wordpress.com/tag/archival-processing/

What about when they get this kind of stuff?

Source: Simson Garfinkel

Source: “Digital Forensics and creation of a narrative.” *Da Blog: ULCC Digital Archives Blog*. http://dablog.ulcc.ac.uk/2011/07/04/forensics/

Looks like they might need some of this kind of stuff:



AFFLIB

Open Source Computer Forensics Software

Perhaps you've seen some places that look like this:

El Paso County Sheriff's Office (Colorado)

http://shr.elpasoco.com/Law+Enforcement+Bureau/Investigations+Division/Computer+Crime+Lab.htm

How about these places?

Stanford University Libraries and Academic Information Resources (SULAIR)

Bodleian Library, Oxford University

British Library, London

Repositories (Archives, Libraries) are Getting Digital Media:

- Floppy disks in boxes
- CDs
- Entire machines
- …

What the Archivists can Get from Digital Forensics

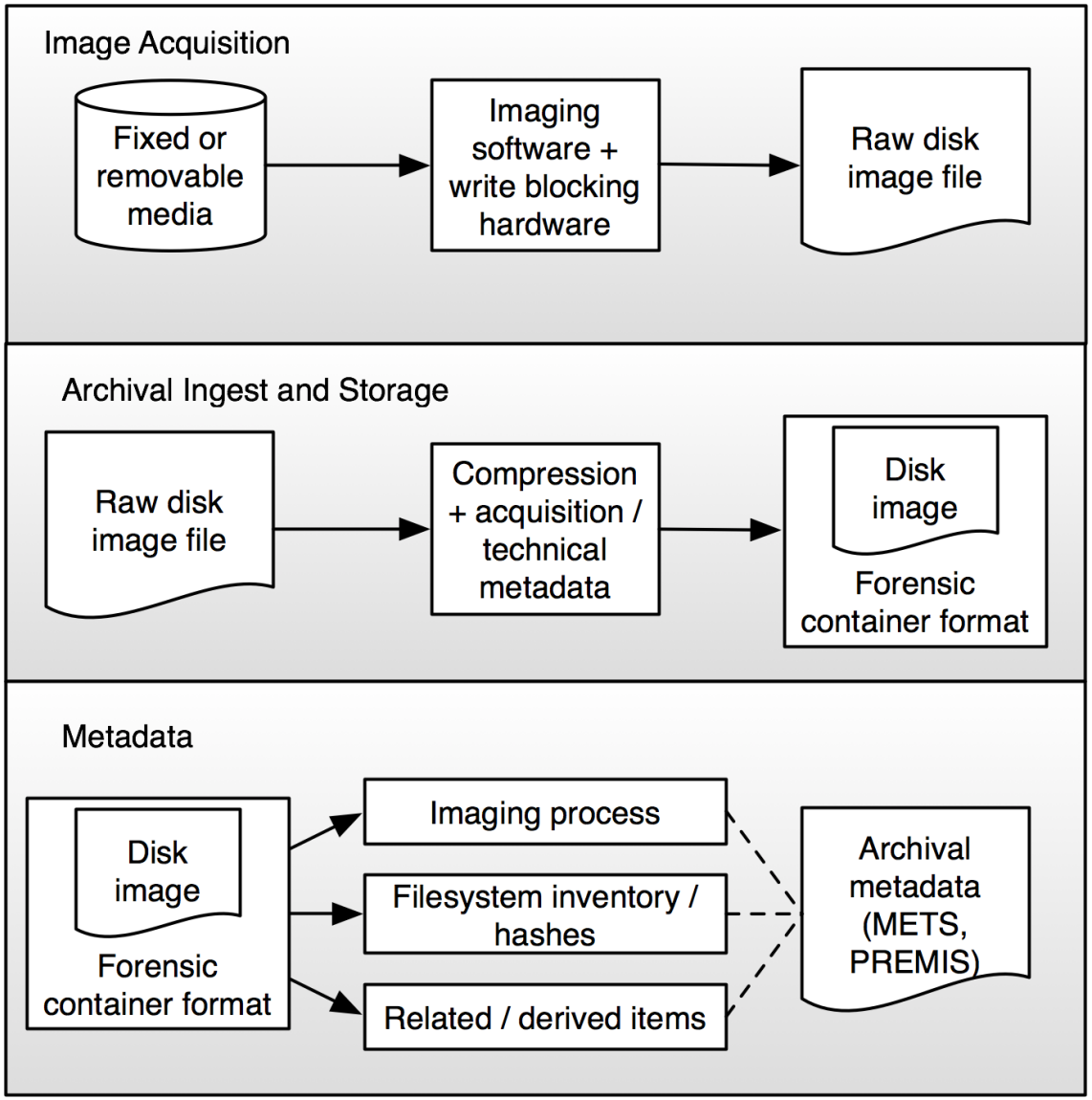
- Recover data when layers of technology fail or are no longer available
- Capturing evidence from places that are not always immediately visible
- Ensuring that actions don't make irreversible changes to essential characteristics (e.g. MAC values)
- Attending to order of volatility
- Documenting what we do, so others will know what we might have changed
- Taking advantage of the information associated with files to ensure that users of the files understand their context of creation

Particular Archival Considerations

- Long-term **preservation** of forensic images, including interoperability across applications
- Provision of **access** to images that contain sensitive information (e.g. zeroing out portions, hashes as surrogates for bitstreams, on-the-fly redaction)
- Conventions for **describing** and annotating images
- **User interfaces** for making sense of content at multiple levels of representation (e.g. sector level data, registry keys, parts of a .PST, entire discrete files)
- Distinct **scalability** issues (unit of analysis = collection, rather than a case)

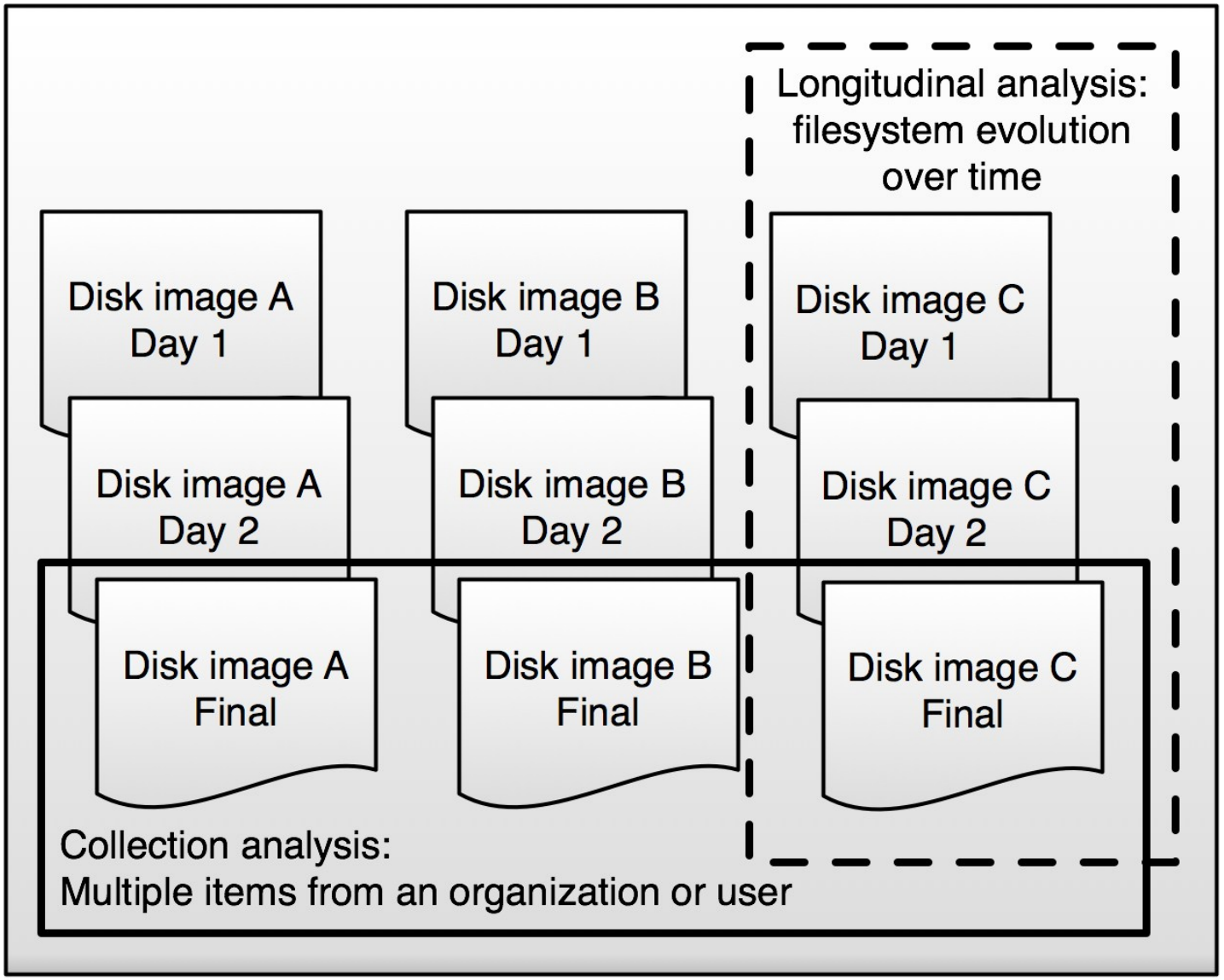
Curation of Disk Images

Storage Media Acquisition and Handling Profile for Digital Repositories*



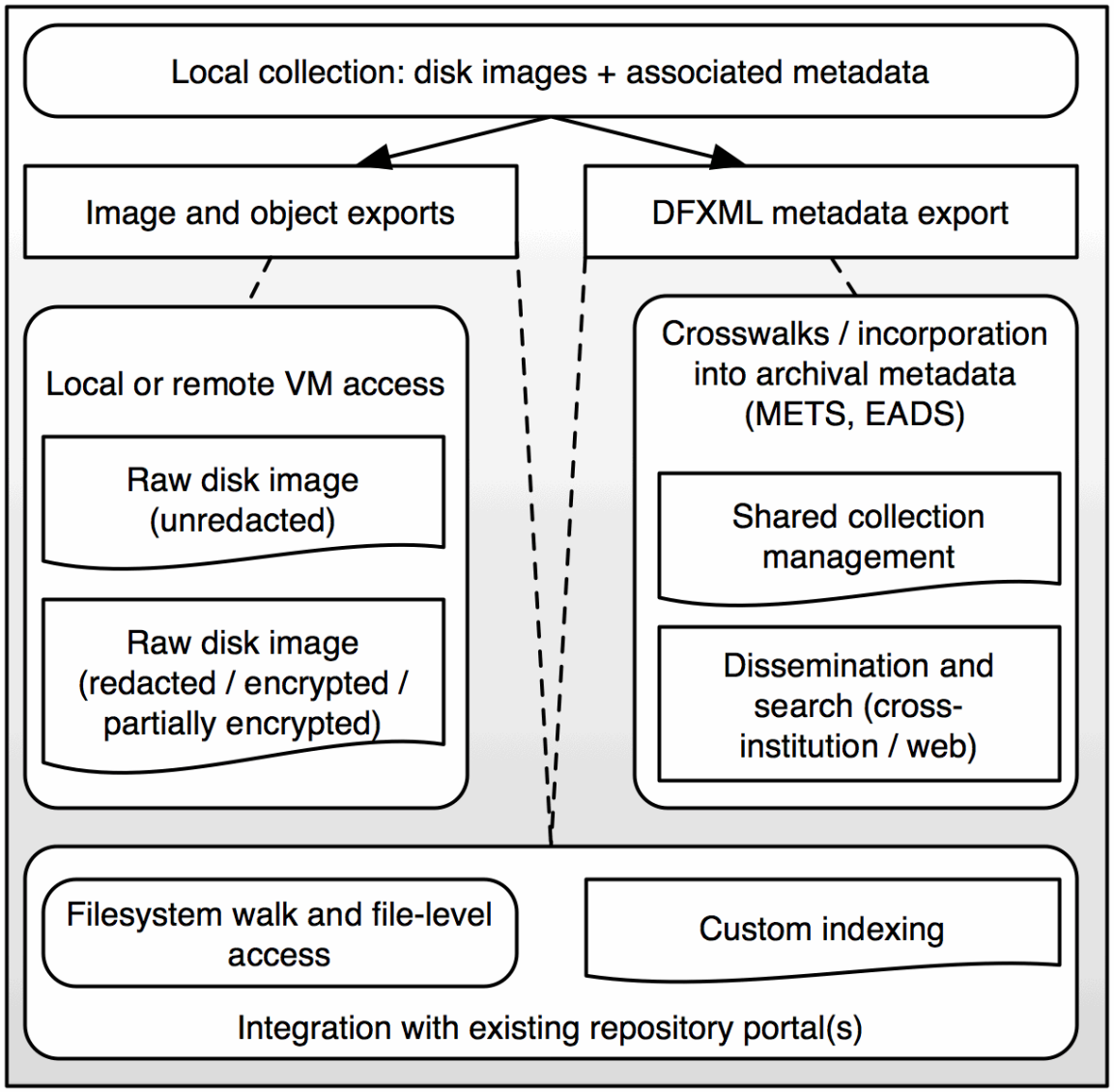
*Woods, Kam, Christopher A. Lee, and Simson Garfinkel. “Extending Digital Repository Architectures to Support Disk Image Preservation and Access.” In *Proceedings of the 2011 Joint Conference on Digital Libraries* (forthcoming).

Modes of Disk Image Collection Analysis*



*Woods, Kam, Christopher A. Lee, and Simson Garfinkel. “Extending Digital Repository Architectures to Support Disk Image Preservation and Access.” In *Proceedings of the 2011 Joint Conference on Digital Libraries* (forthcoming).

Distribution Paths*



*Woods, Kam, Christopher A. Lee, and Simson Garfinkel. “Extending Digital Repository Architectures to Support Disk Image Preservation and Access.” In *Proceedings of the 2011 Joint Conference on Digital Libraries* (forthcoming).

Vision for the Future

- Widespread incorporation of the right bits and pieces of forensics methods into routine processing of acquisitions by collecting institutions
- BitCurator – a modular software environment that implements various batch processes on bitstreams to support two contexts:
  - Collecting institutions that are already generating forensic images of removable media
  - Individuals and institutions that are just getting started (BitCurator in a Box)

Financial Support:

- National Science Foundation: Creating Realistic Forensic Corpora for Undergraduate Education and Research (DUE-0919593)
- Andrew W. Mellon Foundation: Digital Acquisition Learning Laboratory

Thank you!