

Using Forensics Tools and Methods for Data Curation

Lead instructors: Cal Lee and Kam Woods

Main Topics:

- Definition and landscape of digital forensics
- Main opportunities and advantages for applying digital forensics to the curation of collections
- Available tools and methods
- Current and future developments in applying digital forensics to digital collections

Learning Objectives:

- Understand the relevance of earlier discussions about levels of representation when acquiring and processing born-digital materials
- Identify various forms of data that may be "hidden" on the physical storage medium
- Recognize the role of write blockers and creation of disk images to prevent accidental manipulation of volatile data
- Identify the data that a file system uses to manage files
- Identify and compare alternative strategies for providing public access to data from disk images

Recommended Readings on Basic Concepts and Motivations:

Garfinkel, Simson, and David Cox. "Finding and Archiving the Internet Footprint." Paper presented at the First Digital Lives Research Conference: Personal Digital Archives for the 21st Century, London, UK, February 9-11, 2009.

<http://simson.net/clips/academic/2009.BL.InternetFootprint.pdf>

John, Jeremy Leighton. "Adapting Existing Technologies for Digitally Archiving Personal Lives: Digital Forensics, Ancestral Computing, and Evolutionary Perspectives and Tools." Paper presented at iPRES 2008: The Fifth International Conference on Preservation of Digital Objects, London, UK, September 29-30, 2008.

Kirschenbaum, Matthew G., Richard Ovenden, and Gabriela Redwine. "Digital Forensics and Born-Digital Content in Cultural Heritage Collections." Washington, DC: Council on Library and Information Resources, 2010. <http://www.clir.org/pubs/reports/pub149/pub149.pdf>

Lee, Christopher A., Matthew Kirschenbaum, Alexandra Chassanoff, Porter Olsen, and Kam Woods. "BitCurator: Tools and Techniques for Digital Forensics in Collecting Institutions." D-Lib Magazine 18, No. 5/6 (May/June 2012). <http://www.dlib.org/dlib/may12/lee/05lee.html>

Lee, Christopher A., and Kam Woods. "Automated Redaction of Private and Personal Data in Collections: Toward Responsible Stewardship of Digital Heritage." In *Proceedings of Memory of the World in the Digital Age: Digitization and Preservation: An International Conference on Permanent Access to Digital Documentary Heritage, 26-28 September 2012, Vancouver, British Columbia, Canada*, edited by Luciana Duranti and Elizabeth Shaffer, 298-313: United Nations Educational, Scientific and Cultural Organization, 2013.
<http://www.ils.unc.edu/caltee/p298-lee.pdf>

Lee, Christopher A., Kam Woods, Matthew Kirschenbaum, and Alexandra Chassanoff. "From Bitstreams to Heritage: Putting Digital Forensics into Practice in Collecting Institutions." September 30, 2013. <http://www.bitcurator.net/docs/bitstreams-to-heritage.pdf>

Ross, Seamus, and Ann Gow. "Digital Archaeology: Rescuing Neglected and Damaged Data Resources." London: British Library, 1999. <http://eprints.erpanet.org/47/>

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. <http://www.ils.unc.edu/calleep/p57-woods.pdf>

Other Related Readings:

AIMS Working Group. "AIMS Born-Digital Collections: An Inter-Institutional Model for Stewardship." 2012. <http://www2.lib.virginia.edu/aims/whitepaper/>

Carrier, Brian. *File System Forensic Analysis*. Boston, MA: Addison-Wesley, 2005.
Elford, Douglas, Nicholas Del Pozo, Snezana Mihajlovic, David Pearson, Gerard Clifton, and Colin Webb. "Media Matters: Developing Processes for Preserving Digital Objects on Physical Carriers at the National Library of Australia." Paper presented at the 74th IFLA General Conference and Council, Québec, Canada, August 10-14, 2008.

Farmer, Dan, and Wietse Venema. *Forensic Discovery*. Upper Saddle River, NJ: Addison-Wesley, 2005.

Garfinkel, Simson L. "AFF: A New Format for Storing Hard Drive Images." *Communications of the ACM* 49, no. 2 (2006): 85-87. <http://simson.net/clips/academic/2006.CACM.AFF.pdf>

Garfinkel, Simson L. "Digital forensics research: The next 10 years." *Digital Investigation* 7 (2010): S64-73. <http://dfwrs.org/2010/proceedings/2010-308.pdf>

Garfinkel, Simson L. "Forensic feature extraction and cross-drive analysis." *Digital Investigation* 3S (2006): S71-81. <http://simson.net/clips/academic/2006.DFRWS.pdf> [Specifically: Sections 1-3, p.S71-75]

John, Jeremy Leighton. "Adapting Existing Technologies for Digitally Archiving Personal Lives: Digital Forensics, Ancestral Computing, and Evolutionary Perspectives and Tools." Paper presented at iPRES 2008: The Fifth International Conference on Preservation of Digital Objects, London, UK, September 29-30, 2008.
http://www.bl.uk/ipres2008/presentations_day1/09_John.pdf

John, Jeremy Leighton. "Digital Forensics and Preservation." DPC Technology Watch Report 12-03. Digital Preservation Coalition, November 2012
http://www.dpconline.org/component/docman/doc_download/810-dpctw12-03.pdf

John, Jeremy Leighton, Ian Rowlands, Peter Williams, and Katrina Dean. "Digital Lives: Personal Digital Archives for the 21st Century >> An Initial Synthesis." Version 0.2. March 3, 2010. <http://britishlibrary.typepad.co.uk/files/digital-lives-synthesis02-1.pdf>.

Jones, Keith J., Richard Bejtlich, and Curtis W. Rose. *Real Digital Forensics: Computer Security and Incident Response*. Upper Saddle River, NJ: Addison-Wesley, 2006.
Kirschenbaum, Matthew G. *Mechanisms: New Media and the Forensic Imagination*. Cambridge, MA: MIT Press, 2008.

Kirschenbaum, Matthew G., Erika Farr, Kari M. Kraus, Naomi L. Nelson, Catherine Stollar Peters, Gabriela Redwine, and Doug Reside. "Approaches to Managing and Collecting Born-Digital Literary Materials for Scholarly Use." College Park, MD: University of Maryland, 2009.

Lavagnino, John. "The Analytical Bibliography of Electronic Texts." Paper presented at the Joint Annual Conference of the Association for Literary and Linguistic Computing and the Association for Computers and the Humanities, Bergen, Norway, 1996.

Lee, Christopher A. "Computer-Supported Elicitation of Curatorial Intent." In *Dagstuhl Seminar Proceedings 10291, Automation in Digital Preservation*, edited by Andreas Rauber, Jean-Pierre Chanod, Seamus Ross, and Milena Dobрева. 2010.

Petzold, Charles. *Code: The Hidden Language of Computer Hardware and Software*. Redmond, WA: Microsoft Press, 1999. [Bytes and Hex (180-189)]

Radtisch, Markus, Peter May, Asger Askov Blekinge, Per Møldrup-Dalum, and Alan Akbik. "Characterisation Technology, Release 1 and Release Report." Scalable Preservation Environments (SCAPE) Project, 2012.
http://www.scape-project.eu/wp-content/uploads/2012/05/SCAPE_D9.1_SB_v1.0.pdf

Thomas, Susan, Renhart Gittens, Janette Martin, and Fran Baker. "Workbook on Digital Private Papers." 2007. Paradigm Project. <http://www.paradigm.ac.uk/workbook/introduction/index.html>.

Thomas, Susan, and Janette Martin. "Using the Papers of Contemporary British Politicians as a Testbed for the Preservation of Digital Personal Archives." *Journal of the Society of Archivists* 27, no. 1 (2006): 29-56.

Underwood, William E., and Sandra L. Laib. "PERPOS: An Electronic Records Repository and Archival Processing System." Paper presented at the International Symposium on Digital Curation (DigCCurr 2007), Chapel Hill, NC, April 18-20, 2007.

Underwood, William, Marlit Hayslett, Sheila Isbell, Sandra Laib, Scott Sherrill, and Matthew Underwood. "Advanced Decision Support for Archival Processing of Presidential Electronic Records: Final Scientific and Technical Report." Technical Report ITTL/CSITD 09-05. October 2009.

White, Ron and Timothy Edward Downs. *How Computers Work*. 9th Edition. How It Works Series. Indianapolis, IN: Que, 2007. [Data Storage (158-75, 182-3, 186-7)]

Woods, Kam and Geoffrey Brown. "Creating Virtual CD-ROM Collections." In *Proceedings of the 5rd International Conference on the Preservation of Digital Objects*. Also in *International Journal of Digital Curation* Vol. 4, No. 2. 2009.

Woods, Kam, and Geoffrey Brown. "From Imaging to Access - Effective Preservation of Legacy Removable Media." In *Archiving 2009: Preservation Strategies and Imaging Technologies for Cultural Heritage Institutions and Memory Organizations: Final Program and Proceedings*, 213-18. Springfield, VA: Society for Imaging Science and Technology, 2009.
<http://www.digpres.com/publications/woodsbrownarch09.pdf>

Woods, Kam and Geoffrey Brown. "Migration Performance for Legacy Data Access." In *Proceedings of the 3rd International Digital Curation Conference*. Also in *International Journal of Digital Curation* Vol. 3, No. 2. 2008.

Woods, Kam and Christopher A. Lee. "Acquisition and Processing of Disk Images to Further Archival Goals." In *Proceedings of Archiving 2012* (Springfield, VA: Society for Imaging Science and Technology, 2012), 147-152.

<http://www.ils.unc.edu/callee/archiving-2012-woods-lee.pdf>

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.

<http://www.ils.unc.edu/callee/p239-woods.pdf>

Woodyard, Deborah. "Data Recovery and Providing Access to Digital Manuscripts." Paper presented at the Information Online 2001 Conference, Sydney, Australia, January 16-18, 2001.