

Kam A. Woods

<https://kamwoods.net/>
kamwoods@email.unc.edu

CB #3360, 100 Manning Hall
Chapel Hill, NC 27599-3360

- EDUCATION**
- ◇ **Indiana University**, Bloomington, IN.
Ph.D. Computer Science (Ph.D. Minor: Computational Linguistics), October 2010. Dissertation title: *Preserving Long-Term Access to United States Government Documents in Legacy Digital Formats*
M.S. Computer Science, May 2008.
 - ◇ **Baylor University**, Waco, TX.
M.S. Computer Science, 2004. Thesis: *Automatic Discovery and Classification of Discourse Contingency Relations*
 - ◇ **Swarthmore College**, Swarthmore, PA.
B.A., Special Major in Computer Science, May 2002.
- PROFESSIONAL EXPERIENCE**
- Research Scientist**, University of North Carolina at Chapel Hill.
October 2013 - present.
- Adjunct Faculty**, University of North Carolina at Chapel Hill.
Spring 2014 - Spring 2017.
- Postdoctoral Research Associate**, University of North Carolina at Chapel Hill.
August 2010 - September 2013.
- RESEARCH AREAS**
- ◇ Machine learning and NLP techniques for archival data analysis and management
 - ◇ Applications of digital forensics to archival data collections.
 - ◇ Privacy and security in data management environments.
 - ◇ Format identification, migration, and risk management.
- RESEARCH GRANTS**
- CARASCAP - Computer-Assisted Redaction and Anonymization of Scholarly Communications and Products (NSF EAGER grant, Grant Reference 2135446), Technical Lead. 2021-2022.
- RATOM - Review, Appraisal, and Triage of Mail (Andrew W. Mellon Foundation, Grant Reference 1802-05537), Co-PI and Technical Lead. 2019-2021.
- BitCurator NLP (Andrew W. Mellon Foundation, Grant Reference 31600663), Co-PI and Technical Lead. 2016-2018.
- BitCurator Access (Andrew W. Mellon Foundation, Grant Reference 31400605), Co-PI and Technical Lead. 2014-2016.
- BitCurator: Tools for Digital Forensics Methods and Workflows (Phase 2) (Andrew W. Mellon Foundation, Grant Reference 11300699), Technical Lead. 2013-2014.
- BitCurator: Tools for Digital Forensics Methods and Workflows (Phase 1) (Andrew W. Mellon Foundation, Grant Reference 31100658), Technical Lead. 2011-2013.
- Curation of a Forensic Data Collection for Education (National Science Foundation, subgrant through award DUE-0919593), Postdoctoral Research Associate. 2010-2011.
- SOFTWARE DEVELOPMENT**
- A representative sample of my current and previous software projects can be found on GitHub at <https://github.com/kamwoods>

REFEREED
CONFERENCE
PAPERS

Kam Woods, Christopher Lee, Oleg Stobbe, Thomas Liebetraut and Klaus Rechert. 2015. Functional Access to Forensic Disk Images in a Web Service. In *Proceedings of the 12th International Conference on Digital Curation (191-195)*. Chapel Hill, NC: University of North Carolina, School of Information and Library Science.

Kam Woods and Christopher A. Lee. 2015. Redacting Private and Sensitive Information in Born-Digital Collections. In *Proceedings of Archiving 2015*. Society for Imaging Science and Technology.

Kam Woods, Alexandra Chassanoff, and Christopher A. Lee. 2013. Managing and Transforming Digital Forensics Metadata for Digital Collections. In *Proceedings of the 10th International Conference on Preservation of Digital Objects (iPRES 2013)*. Lisbon, Portugal.

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

Christopher A. Lee and Kam Woods. 2012. Automated Redaction of Private and Personal Data in Collections: Toward Responsible Stewardship of Digital Heritage. *The Memory of the World in the Digital age: Digitization and Preservation* Vancouver, BC: UNESCO.

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

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

Kam Woods, Christopher A. Lee, Simson Garfinkel, David Dittrich, Adam Russell, and Kris Kearton. 2011. Creating Realistic Corpora for Forensic and Security Education. *ADFSL Conference on Digital Forensics, Security and Law*.

Kam Woods and Geoffrey Brown. 2009. Assisted Emulation for Legacy Executables. *Proceedings of the 5th International Digital Curation Conference*.

Geoffrey Brown and Kam Woods. 2009. Born Broken - Fonts and Information Loss in Legacy Digital Documents. *Proceedings of the 6th International Conference on Preservation of Digital Objects*.

Kam Woods and Geoffrey Brown. 2009. From Imaging to Access - Effective Preservation of Legacy Removable Media. *Proceedings of Archiving 2009*.

Kam Woods and Geoffrey Brown. 2008. Virtualization for Preservation of Executable Art. *Proceedings of the 2008 DOCAM Summit and Symposium*.

Kam Woods and Geoffrey Brown. 2008. Creating Virtual CD-ROM Collections. *Proceedings of the 5th International Conference on Preservation of Digital Objects*. Also in *International Journal of Digital Curation* Vol. 4, No. 2 2009.

Kam Woods and Geoffrey Brown. 2007. Migration Performance for Legacy Data Access. *3rd International Digital Curation Conference*. Also in *International Journal of Digital Curation* Vol. 3, No 2. 2008.

Roxana Girju and Kam Woods. 2005. Exploring Contingency Discourse Relations. *Proceedings of the 6th International Workshop on Computational Semantics (IWCS-6)*, Harry Bunt (ed.), The Netherlands.

WHITE
PAPERS AND
REPORTS

Christopher A. Lee, Camille Tyndall Watson, Jamie Patrick-Burns, Sangeeta Desai, Elisia Kinder, Antoine De Torcy, and Kam Woods. Email Curation in Context: A Report of the Review, Appraisal, and Triage of Mail (RATOM) Project. March 30, 2022.

Christopher A. Lee, Porter Olsen, Alexandra Chassanoff, Kam Woods, Matthew Kirschenbaum, and Sunitha Misra. From Code to Community: Building and Sustaining BitCurator through Community Engagement. September 30, 2014.

Sunitha Misra, Christopher A. Lee, and Kam Woods. A Web Service for File-Level Access to Disk Images. *Code4Lib Journal* 25 (2014).

Christopher A. Lee, Kam Woods, Matthew Kirschenbaum, and Alexandra Chassanoff. From Bitstreams to Heritage: Putting Digital Forensics into Practice in Collecting Institutions. *BitCurator Project White Paper*. December 2013.

Christopher A. Lee, 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).

POSTERS
AND
RESEARCH
REPORTS

Kam Woods. 2014. Filling in the Gaps: Improving Data Curation Education and Research Outcomes in the BitCurator Project (Poster). Presented at *ALISE 2014 Annual Conference*. Philadelphia, PA.

Alexandra Chassanoff, Kam Woods, and Christopher Lee. 2013. Mapping Digital Forensics Metadata to Preservation Events using BitCurator (Poster). Presented at *Joint Annual Meeting of the Society of American Archivists 2013*. New Orleans, LA.

Christopher A. Lee and Kam Woods. BitCurator (Poster). 2013. Presented by Christopher A. Lee at *The 10th International Conference on Preservation of Digital Objects*. Lisbon, Portugal.

Christopher A. Lee and Kam Woods. 2011. BitCurator: Tools for Digital Forensics Methods and Workflows in Real-World Collecting Institutions (Poster). *7th International Digital Curation Conference*. Bristol, UK. December 6-8.

Kam Woods. 2011. Digital Forensics Tools and Realistic Corpora for Long-Term Data Preservation, Access, and Education (Poster). *First Annual Postdoctoral Poster Competition*. Chapel Hill, NC.

Christopher A. Lee and Kam Woods. 2011. Digital Acquisition Learning Laboratory: A White Paper (Research Report).

BOOKS AND
CHAPTERS

Book Chapter: Alexandra Chassanoff, Kam Woods, and Christopher A. Lee *Digital Preservation Metadata Practice for Disk Image Access*, in *Digital Preservation Metadata for Practitioners*. Angela Dappert, Rebecca Squire Guenther, and Sebastien Peyrard (Eds.). Springer, January 2016.

Proceedings: Christopher A. Lee, Jonathan Crabtree, Leo Konstantelos, Nancy McGovern, Yukio Maeda, Maureen Pennock, Helen Tibbo, Kam Woods, and Eld Zierau, eds. *Proceedings of the 12th International Conference on Digital Preservation*. Chapel Hill, NC: University of North Carolina, School of Information and Library Science, 2015.

CONFERENCE
SESSIONS,
PANELS, AND
SYMPOSIA

Talk: Digital Curation and Digital Forensics at UNC SILS – Recent Activities. *Open Preservation Foundation Annual Meeting 2016*. The Hague, Netherlands. June 1, 2016.

Talk: BCA-Webtools: Accessing and Visualizing Disk Images in a Web Browser. *CurateGear 2016*. Chapel Hill, North Carolina. January 14, 2016.

Talk: Bitstream Confidential - Considerations and Approaches to Curating Potentially Sensitive Personal Data in Collections. Kam Woods (with Christopher A. Lee in absentia). *Personal Digital Archiving 2015*. Los Angeles, CA. April 24, 2015.

Lightning Talk: BitCurator - What's New in the Past Year. Kam Woods. *CurateGEAR Symposium*. Chapel Hill, NC. January 8, 2014.

Panel: All Your Bits Aren't Belong To Us: Opportunities and Challenges of Personally Revealing Information in Digital Collections. Matt Kirschenbaum, Cal Lee, Naomi Nelson, and Kam Woods. *Personal Digital Archiving Conference*. College Park, MD, February 21-22, 2013.

Talk: BitCurator: Data Forensics for Digital Libraries. *LAUNC-CH Conference*. Chapel Hill, NC, March 5 2012.

Talk: Data Triage and Data Analytics for Personal Digital Collections. *Personal Digital Archiving Conference*. San Francisco, CA, February 23-24, 2012.

Talk: BitCurator: Tools for Digital Forensics Methods and Workflows in Real-World Collecting Insitutions. *CurateGEAR 2012*. Chapel Hill, NC. January 6, 2012.

Demonstration: docXtractor - Metadata Extraction for Legacy Microsoft Word Documents. *Practical Tools for Digital Preservation: A Hackathon*. York, UK, September 27-29, 2011.

Panel: *At the Nexus of Analog and Digital: A Symposium for Preservation Educators*. Ann Arbor, MI, June 5-7, 2011.

Talk: Creating Realistic Corpora for Forensic and Security Education. *ADFSL Conference on Digital Forensics, Security, and Law* (ADFSL 2011). Richmond, VA, May 25-27, 2011.

Talk: Archival Applications of Digital Forensics Tools and Techniques. *Personal Digital Archiving*. San Francisco, CA, February 24-25, 2011.

Talk (with Christopher A. Lee): Curation of a Disk Image Collection to Support Education (Poster). *6th International Digital Curation Conference*. London, UK, December 6-8, 2010.

Participant: Digital Forensics Research Workshop, Portland, OR, August 2-4, 2010.

Talk (with Geoffrey Brown): Born Broken: Fonts and Information Loss in Legacy Digital Documents. *6th International Conference on Preservation of Digital Objects*. October 5-7, 2009.

Talk (with Geoffrey Brown): Creating Virtual CD-ROM Collections. *5th International Conference on Preservation of Digital Objects*. September 29-October 1, 2008.

Talk (with Geoffrey Brown): Migration Performance for Legacy Data Access. *3rd International Digital Curation Conference*. December 11-13, 2007

WORKSHOPS
AND
WEBINARS

BitCurator Consortium Webinar on BitCurator Access Webtools, September 16, 2016.

Digital Forensics for Archivists (Advanced), Baton Rouge, LA, April 6-8, 2016

BitCurator and BitCurator Access: A Quick Overview. MetaArchive-BitCurator Roundtable (webinar). Online. May 5, 2015

BitCurator: Digital Forensics for Collecting Institutions. Introductory talk for SERI Educational Webinar, "Getting Bits off Disks". Online. January 14, 2014

An Overview of the BitCurator Project: Tools for Digital Forensics Methods and Workflows in Real-World Collecting Institutions. A.R.T. (Archivists Roundtable of Metropolitan New York) and LLF Program Presentation. Online. November 19, 2013

OPF Webinar - Capturing and Analyzing Forensic Disk Images with BitCurator. Taught for the Open Planets Foundation Webinar Series. Online at <http://openplanetsfoundation.org/events/opf-webinar-capturing-and-analyzing-forensic-disk-images-bitcurator>. August 9, 2013

Adding Digital Forensics Tools to your Workflow using BitCurator (with Christopher A. Lee). Archiving 2013 Workshop T1A. Washington, DC. April 2, 2013

Digital Forensics for Archivists, Parts 1 and 2. Taught for the Society of American Archivists Digital Archives Specialist Curriculum & Certificate Program. Chicago, IL. September 11-13, 2013

Digital Forensics for Archivists, Parts 1 and 2. Taught for the Society of American Archivists Digital Archives Specialist Curriculum & Certificate Program. Ann Arbor, MI. June 24-25, 2013

OPF Hackathon - Capturing and Analyzing Forensic Disk Images with BitCurator. Chapel Hill, NC. June 3-5, 2013.

OPF Hackathon: A Practical Approach to Disk Images & Digital Forensics. Copenhagen, Denmark. May 15-17, 2013.

Digital Forensics and Preservation Workshop (Digital Preservation Coalition). Oxford, UK. June 28th, 2011.

PROFESSIONAL
SERVICE AND
LEADERSHIP

Digital Forensics for Archivists course instructor, Society of American Archivists Specialist Curriculum & Certificate Program (2012-present)

Member, ACM (2001-present)

Member, Curation and Archives Research Group, UNC Chapel Hill (2010-present)

Officer, Indiana University Computer Science Graduate Student Association (2005 – 2007)

Officer, Swarthmore College Computer Society (2000-2002)

RESEARCH
AND
SOFTWARE
PROJECTS

- ◇ **CARASCAP - Computer-Assisted Redaction and Anonymization of Scholarly Communications and Products** Technical Lead on NSF (EAGER) funded project to develop software that can identify and redact sensitive information within research-related datasets, documents, and communications. (Summer 2021 - Summer 2022)
- ◇ **RATOM - Review, Appraisal, and Triage of Mail** Technical Lead on Andrew W. Mellon Foundation funded project to develop improved archival email management tools and workflows using ML and NLP libraries. (Fall 2019 - Fall 2021)
- ◇ **BitCurator NLP** Co-PI and technical lead on Andrew W. Mellon Foundation funded project to apply modern natural language processing techniques to born digital collections, particularly those focused on entity extraction and topic modeling. (Fall 2016 - Fall 2018)
- ◇ **BitCurator Access** Co-PI and technical lead on Andrew W. Mellon Foundation funded project to provide access to forensically-packaged materials, perform private and sensitive data analysis and redaction, and build automated emulated environments. (Fall 2014 - 2016)
- ◇ **BitCurator: Tools for Digital Forensics Methods and Workflows in Real-World Collecting Institutions.** Technical and programming lead on Andrew W. Mellon funded project to build, test, and analyze systems and software for incorporating digital forensics methods into data preservation and archiving systems. (Fall 2010 - Fall 2014)
- ◇ **FDLP CD-ROM virtualization.** Technical and programming lead under supervision of Dr. Geoffrey Brown. Developed tools to ensure low-risk transfer of legacy materials to archival distributed storage, analyze file distribution, and perform format analysis. Coordinated with Indiana University Libraries staff; trained and supervised multiple undergraduate student programmers. (Summer 2007-Summer 2010)
- ◇ **Legacy document font analysis.** Identification, extraction, and use for quality assurance of font data from legacy binary document formats. Constructed a novel terabyte-scale dataset subsequently analyzed using existing software libraries and customized code written to integrate with Microsoft Office (.NET development) and perform standalone analysis in cross-platform environments. (Spring 2009-Summer 2010)
- ◇ **Format migration and risk analysis.** Continued development of high-throughput, low-risk server-side migration procedures using off-the-shelf open source packages such as OpenOffice and associated APIs. Exploration of automated techniques to identify high-risk components of migrated documents. (Summer 2007-Summer 2010)
- ◇ **Independent research (Indiana University).** Interlingua research performed with data from the Preliminary Mayan Etymological Dictionary. Automated approaches for ontology construction and mining of semantic information from sparse data. Software prototype and dataset constructed under supervision of Dr. Michael Gasser. (Spring 2007-Fall 2008)

RESEARCH
AND
TEACHING
POSITIONS

- ◇ **Research Scientist**, University of North Carolina at Chapel Hill School of Information and Library Science
October 2013 - present
Currently: Technical Lead on the RATOM project. Previously: Co-PI and technical lead on the BitCurator NLP and BitCurator Access projects. In this role I have led research and software development activities, as well as coordinating with members of the BitCurator Consortium to support dissemination and support of the tools.
- ◇ **Postdoctoral Research Associate**, University of North Carolina at Chapel Hill School of Information and Library Science
August 2010 - October 2013
Technical lead on the BitCurator (Phase 1 and 2) projects. In this role I led research and software development activities, and promoted community engagement and education.
- ◇ **Adjunct Faculty**, University of North Carolina at Chapel Hill School of Information and Library Science
Spring 2014 - present
- ◇ **Research Assistant**, Indiana University Computer Science Department
May 2007 - August 2010
Long-term digital preservation research with Dr. Geoffrey Brown. In this role I also acted as mentor to undergraduate programmers and data acquisition hires over the course of three years, coordinating development of prototype software systems for data preservation and access.
- ◇ **Instructor**, University of North Carolina at Chapel Hill, School of Information and Library Science
January 2011 - May 2011, August 2013 - Present
Course instructor:
 - INLS 690-141: Digital Forensics for Curation of Digital Collections (Spring 2017)
 - INLS 465: Understanding IT for Digital Collections (Spring 2014)
 - INLS 690-141: Digital Forensics for Curation of Digital Collections (Fall 2013)
Co-taught with Christopher A. Lee.
 - INLS 490-141: Acquiring Information from Digital Storage Media (Spring 2011)
Co-taught with Christopher A. Lee.
- ◇ **Associate Instructor**, Indiana University, Computer Science Department
August 2005 - May 2007
Teaching assistant:
 - A110: Introduction to Computers and Computing
Taught laboratory sessions emphasizing digital literacy skills, including training students from diverse backgrounds in advanced document creation and management and in basic web development.
 - A201: Introduction to Programming
Introduction to programming in Python. Designed laboratory activities, assignments, and tests in course designed primarily for non-Computer Science majors. Across three semesters assisted in improving course design, introducing students to programming techniques using virtual worlds, basic algorithm implementation and analysis, and group programming techniques.
Course instructor:
 - A202: Introduction to Programming II
Introductory object-oriented programming in Java. Building on an existing course design, taught an accelerated summer course introducing students to user-defined functions and types, recursion and iteration, parameter-passing mechanisms, and interface design.