

Kam A. Woods

kamwoods@cs.indiana.edu
http://www.cs.indiana.edu/~kamwoods

1101 N. Union St. Apartment 203
Bloomington, IN 47408
(812) 606-1755

- EDUCATION
- ◇ **Indiana University**, Bloomington, IN.
Ph.D. Computer Science (Ph.D Minor: Computational Linguistics), expected graduation: May 2010.
Written and Oral qualifying exams passed. Disseratation proposal completed. Ph.D Minor requirements fulfilled.
 - ◇ **Indiana University**, Bloomington, IN.
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.
- PUBLICATIONS
- K. Woods and G. Brown. 2009. Assisted Emulation for Legacy Executables. To Appear in *Proceedings of the 5th International Digital Curation Conference*.
- G. Brown and K. Woods. 2009. Born Broken - Fonts and Information Loss in Legacy Digital Documents. *Proceedings of the 6th International Conference on Preservation of Digital Objects*.
- K. Woods and G. Brown. 2009. From Imaging to Access - Effective Preservation of Legacy Removable Media. *Proceedings of Archiving 2009*.
- K. Woods and G. Brown. 2008. Virtualization for Preservation of Executable Art. *Proceedings of the 2008 DOCAM Summit and Symposium*.
- K. Woods and G. 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.
- K. Woods and G. 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.
- R. Girju and K. Woods. 2005. Exploring Contingency Discourse Relations. *Proceedings of the 6th International Workshop on Computational Semantics (IWCS-6)*, Harry Bunt (ed.), The Netherlands.
- RESEARCH INTERESTS
- Development and evaluation of tools and techniques for long-term digital preservation. Development of systems for high-performance automated format migration, low-risk file format identification, and handling of legacy removable media. Development of risk-management tools for access, migration, and accurate rendition of outdated binary formats. Automation of emulation platforms to support continued access to legacy executables.
- ACADEMIC RESEARCH PROJECTS
- ◇ **FDLP CD-ROM virtualization project**. Working with Dr. Geoffrey Brown at Indiana University, I performed primary technical, programming, and analysis tasks on this project, including the development of novel methods for low-risk transfer of legacy materials to archival distributed storage, file distribution and format analysis. Publications based on this work have been accepted to several major international conferences. I helped coordinate with Indiana University Libraries staff during this project, and supervised multiple undergraduate student programmers.
 - ◇ **Legacy document font analysis**. With Dr. Geoffrey Brown, continued existing research into the identification, extraction, and use for quality assurance of font data from legacy binary document formats such as Microsoft Word. As part of this project I constructed a novel terabyte-scale dataset subsequently analyzed using existing software libraries and customized code written on-site. (Spring 2009-Present)

- ◇ **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-Present, Ongoing)
 - ◇ **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
- ◇ **Research Assistant**, Indiana University (Dr. Geoffrey Brown), Computer Science Department
May 2007 – present
Long-term digital preservation research. 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.
 - ◇ **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.
- TECHNICAL SKILLS
- ◇ C/C++, Java, .NET, Python, Perl, Scheme, PHP, Shell Scripting
 - ◇ Long-term digital preservation, Machine Learning, Information Retrieval, and Natural Language Processing
- SERVICE AND ORGANIZATIONS
- Member, ACM (2001-present)
 - Officer, Indiana University Computer Science Graduate Student Association (2005 – 2007)
 - Officer, Swarthmore College Computer Society (2000-2002)