

## CURRICULUM VITAE

**DR. CATHERINE LESLEY BLAKE**

[clblake@illinois.edu](mailto:clblake@illinois.edu)

<http://people.lis.illinois.edu/~clblake>

January 14, 2016

---

### CONTACT INFORMATION

Graduate School of Library and Information Science  
University of Illinois at Urbana-Champaign, GSLIS 303, MC-493      Phone: +1 217-333-0115  
501 E Daniel St, Champaign IL 61820-3302      FAX: +1 217- 244-3302

---

### EDUCATION

- 2003      Doctor of Philosophy in Information and Computer Science (ICS)**  
Information Access and Management, University of California, Irvine  
**Co- Advisors:** Wanda Pratt (ICS & Medical Informatics) and Dennis Kibler (ICS)  
**Committee:** Michael Pazzani (ICS), Paul Dourish (ICS), and Catherine Carpenter (Medicine)  
**Title:** Information Synthesis: A Mixed-Initiative Meta-Analytic Approach to Facilitate Knowledge Discovery from Scientific Text
- 1999      Master of Science in Information and Computer Science**  
Artificial Intelligence, University of California, Irvine
- 1995      Master of Computer Science**  
Software Engineering, University of Wollongong, Australia
- 1994      Bachelor of Computer Science**  
Software Engineering, University of Wollongong, Australia

---

### EMPLOYMENT

- Associate Director**      **2011 –**  
Center for Informatics Research in Science & Scholarship (CIRSS)
- Associate Professor (with tenure)**      **2009 –**  
Graduate School of Library and Information, University of Illinois at Urbana Champaign, IL  
Courtesy Appointment, Medical Information Science, UIUC (2010-  
Courtesy Appointment, Department of Computer Science, UIUC (2009-2012)
- Assistant Professor**      **2004 – 2009**  
School of Information and Library Science, University of North Carolina, Chapel Hill, NC
- Graduate Student Researcher**      **1997 – 2003**  
School of Information and Computer Science, University of California, Irvine, CA
- Librarian for the UCI Machine Learning Repository**      **1998 – 2003**  
School of Information and Computer Science, University of California, Irvine, CA
- Research Scientist**      **1995 – 1998**  
The Broken Hill Proprietary Company Limited Research (BHP Research)  
Advanced Information Systems Group, Melbourne Research Laboratories, Victoria, Australia.
- Applications Programmer Analyst**      **1990 – 1995**  
The Broken Hill Proprietary Company Limited Information Technology (BHP Information  
Technology), Wollongong, New South Wales, Australia.

## HONORS AND AWARDS

---

- Honorable Mention – Poster at the iConference, Fort Worth TX (2013)
- Faculty Fellowship – Environmental Change Institute, UIUC (2010-2011)
- John Wiley & Sons Best Paper Award(s) – Journal of the American Society for Information Science and Technology (2007)
- Faculty Fellowship – Renaissance Computing Institute (RENCI) (2007)
- UNC New Faculty Development Award – UNC Chapel Hill (2006)
- UNC Chapel Hill Nominee for the Microsoft New Faculty Fellowship (2005)
- Dissertation Fellowship – UC Irvine School of Information and Computer Science (2003)
- Dissertation Award – California Breast Cancer Research Program (2002)
- Dissertation Fellowship – UC Irvine Dept of Information and Computer Science (1997)
- Education Abroad Program Scholarship – University of Wollongong, Australia (1994)

## BOOK CHAPTERS

---

- Blake, C. (2015). Biological responses to chemical exposure: Case studies in how to manage ostensible inconsistencies using the Claim Framework. In *Inconsistency Robustness (Studies in Logic)* Edited by Carl Hewitt and John Woods. <http://www.amazon.com/Inconsistency-Robustness-Studies-Logic-Hewitt/dp/1848901593>
- Stanton, J., Palmer, C.L., Blake, C., & Allard, S. (2012) Chapter 6 - Interdisciplinary Data Science Education, In Xiao, N. & McEwen, LR Special Issues in Data Management, American Chemical Society, 97-113. <http://pubs.acs.org/doi/abs/10.1021/bk-2012-1110.cb006>
- Blake, C. (2010) Text Mining, In *Annual Review of Information Science and Technology*, 45(1): 123-155. <http://onlinelibrary.wiley.com/doi/10.1002/aris.2011.1440450110/full>

## PEER-REVIEWED JOURNAL, CONFERENCE AND WORKSHOP PAPERS

---

- Gabb, H.A. and Blake, C. (In Press). An informatics approach to cumulative chemical exposure from consumer products: A case study for asthma-associated and endocrine disrupting chemicals. *Environmental Health Perspectives*.
- Lucic, A. and Blake, C. (2015) Digital Library Support for High Level Learning: A Situated Case Study and Evaluation in Plant Science, *Annual Meeting of the Association for Information Science and Technology*, St. Louis, Missouri, Nov 6-10, 2015. <http://dl.acm.org/citation.cfm?id=2857103>
- Zheng, W. and Blake, C. (2015) Using distant supervised learning to identify protein subcellular localizations from full-text scientific articles, *Journal of Biomedical Informatics*. 57:134–144 <http://www.sciencedirect.com/science/article/pii/S1532046415001537>
- Blake, C. and Lucic, A. (2015) An automated approach to identify endpoints to support the systematic review process. *Journal of Biomedical Informatics*. 56:42-56 <http://www.sciencedirect.com/science/article/pii/S1532046415000830>
- Guo, J., Lu, Y., Mori, T. and Blake, C. (2015) Expert-Guided Contrastive Opinion Summarization for Controversial Issues, *3rd International Workshop on Natural Language Processing for Social Media, 24th International World Wide Web Conference*, Florence, May 19<sup>th</sup>, 2015 <http://dl.acm.org/citation.cfm?id=2743038>
- Blake, C. (2015), Evidence-based discovery, *iConference*, Newport Beach, CA, March 23-27, 2015. <https://www.ideals.illinois.edu/handle/2142/73452>

- Gabb, H., Lucic, A., Blake, C. (2015), A Method to Automatically Identify the Results from Journal Articles, *iConference*, Newport Beach, CA, March 23-27,2015.  
<https://www.ideals.illinois.edu/handle/2142/73645>
- Lucic, A. & Blake, C. (2015) A Syntactic Characterization of Authorship Style Surrounding Proper Names, LLC, The Journal of Digital Scholarship in the Humanities, 30(1):53-70, Available online, June 29, 2013 <http://dsb.oxfordjournals.org/content/digitalsh/30/1/53.full.pdf>
- Blake, C. (2014) Biological responses to chemical exposure: Case studies in how to manage ostensible inconsistencies using the Claim Framework, *Inconsistency Robustness'14*, Stanford University July 29-31, 2014. [http://cirss.lis.illinois.edu/Documents/Publications\\_docs/Blake\\_2014c.pdf](http://cirss.lis.illinois.edu/Documents/Publications_docs/Blake_2014c.pdf)
- Twidale, M., Blake, C. and Gant, J. (2013) Towards a Data Literate Citizenry, *iSchool Conference* Fort Worth, TX, February 2013 247-257.  
<https://www.ideals.illinois.edu/bitstream/handle/2142/38385/189.pdf?sequence=4>
- Blake, C. (2012) Using Secondary Information to inform Public Policy, *Internet, Politics, Policy, Big Data Big Challenges*, Sept 20-21,2012 UK. <http://ipp.oii.ox.ac.uk/2012/programme-2012/track-b-policy/panel-1b-using-big-data-to-inform-public/catherine-blake-using-secondary>
- Hoon Park, D. & Blake, C. (2012) Identifying comparative sentences in full-text scientific articles. *Association of Computational Linguistics, Proceedings of the Workshop on Detecting Structure in Scholarly Discourse*, July, pp 1-9, Jeju South Korea. <http://www.aclweb.org/anthology/W12-4301>
- Blake, C, Zheng W (2011) In Search of Protein Locations, *Association for Computational Linguistics/Human Language Technology Conference (ACL/HLT), Proceedings of the Biological Natural Language Process (BioNLP) Workshop*, pp. 101-102, Portland OR.  
<http://www.aclweb.org/anthology/W11-0212>
- Blake, C. Zheng, W., Painter, K., & Weyerhaeuser, W. (2010) The Role of Semantics in Recognizing Textual Entailment. *Proceedings of the Text Analysis Conference, Recognizing Textual Entailment (RTE6)*, Maryland, DC.  
<http://www.nist.gov/tac/publications/2010/participant.papers/UIUC.proceedings.pdf>
- Zheng, W. & Blake, C. (2010) Automatic extraction of location relations from text, *American Society for Information Science and Technology Annual Meeting*, Oct, Pittsburgh, PA.  
<https://www.ideals.illinois.edu/bitstream/handle/2142/15046/wuzheng.pdf?sequence=2>
- Blake, C. (2010) Beyond genes, proteins, and abstracts: Identifying scientific claims from full-text biomedical articles, *Journal of Biomedical Informatics*, 43(2):173-189.  
<http://www.sciencedirect.com/science/article/pii/S1532046409001476>
- West, S.L., Blake, C., Liu, Z., McKoy, J.N, Oertel, M.D., Carey, T.S. (2009) Reflections on Using Electronic Medical Record Data for Clinical Research. *Health Informatics Journal*, 15(2):108-121.  
<http://jhi.sagepub.com/content/15/2/108.full.pdf+html>
- Blake, C. (2007) The Role of Sentence Structure in Recognizing Textual Entailment. *Proceedings of the ACL-PASCAL Workshop on Textual Entailment and Paraphrasing*, p101-6, Prague, Czech Republic. <http://dl.acm.org/citation.cfm?id=1654557>
- Blake, C., Kampov, J., Orphanides, A., West, D., & Lown, C., (2007) UNC-CH at DUC 2007: Query Expansion, Lexical Simplification, and Sentence Selection strategies for Multi-Document Summarization, Presentation at *Document Understanding Conference (DUC) 2007*, Rochester, NY.  
<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.132.4035>
- Blake, C. & Pratt, W. (2006). Collaborative Information Synthesis I: A Model of information behaviors of scientists in Medicine and Public Health. *Journal of the American Society for Information Science and Technology*, 57(13):1740-9. **JASIST Best Paper Award**  
<http://onlinelibrary.wiley.com/doi/10.1002/asi.20487/abstract>

- Blake, C. & Pratt, W. (2006). Collaborative Information Synthesis II: Recommendations for information systems that support synthesis activities. *Journal of the American Society for Information Science and Technology*. 57(14):1888-95. **JASIST Best Paper Award**  
<http://onlinelibrary.wiley.com/doi/10.1002/asi.20486/abstract>
- Blake,C (2006) A Comparison of Document, Sentence and Term Event Spaces, In *Proceedings of the Joint 21st International Conference on Computational Linguistics (COLING) and the 44th Annual Meeting of the Association for Computational Linguistics (ACL)*, p601-8, Sydney, Australia. (Paper acceptance rate 22.3%) <http://dl.acm.org/citation.cfm?id=1220251>
- Blake,C and Rendall, M., (2006) Scientific Discovery: A View from the Trenches, In Ljupco Todorovski, Nada Lavrac, Klaus P. Jantke (Eds.): *Lecture Notes in Computer Science, Discovery Science, 9th International Conference*, p41-52 Barcelona, Spain. (Long paper acceptance rate 27%)  
[http://link.springer.com/chapter/10.1007/11893318\\_8](http://link.springer.com/chapter/10.1007/11893318_8)
- Blake,C., West, D., Luo,L., Marchionini, G. (2005) Cataloging On-Line Health information: A Content Analysis of the NC Health Info Portal *Proceedings of the Annual American Medical Informatics Association Symposium (AMIA)*, p56-60, Washington DC. (Paper acceptance rate 37%)  
<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1560486/>
- Blake,C (2005) Information Synthesis: A New Approach to Explore Secondary Information in Scientific Literature, In *Proceedings of the 5<sup>th</sup> ACM/IEEE-CS Joint Conference on Digital Libraries*, p56-64, Denver, CO. (Paper acceptance rate 31%)  
[http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=4118515&tag=1](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4118515&tag=1)
- Blake, C. & Pratt, W. (2002). Collaborative Information Synthesis. In Grove, A.(Ed)*Proceedings of the 65<sup>th</sup> Annual Conference of the American Society for Information Science and Technology (ASIST)*, p44-56, Philadelphia, PA. <http://onlinelibrary.wiley.com/doi/10.1002/meet.1450390105/abstract>
- Blake, C. & Pratt, W. (2002). A Semantic Approach to Identify Candidate Treatments from Existing Medical Literature. In Sanda M. Harabagiu and Vinay Chaudhri (Eds.) *Mining Answers from Texts and Knowledge Bases: Papers from the 2002 Spring Symposium*. Technical Report SS-02-06. American Association for Artificial Intelligence, p9-13, Menlo Park, CA.  
<https://www.aaai.org/Papers/Symposia/Spring/2002/SS-02-06/SS02-06-003.pdf>
- Blake, C. (2002). Information Synthesis: A Process used by Scientists in Medicine and Public Health to Overcome Information Overload. Paper in the Doctoral Forum of the *Fourth International Conference on Conceptions of Library and Information Science: Emerging Frameworks and Methods (CoLIS 4)*, Seattle, WA.
- Blake, C. & Pratt, W. (2001). Better Rules Fewer Features: A Semantic Approach to Selecting Features from Text. In *Proceedings of the Institute of Electrical and Electronics Engineers Data Mining Conference (IEEE-DM)*, p59-66, San Jose, CA.  
[http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=989501&tag=1](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=989501&tag=1)
- Blake, C. & Pratt, W. (2001). Better Rules Fewer Features: A Semantic Approach to Selecting features from text. In *Proceedings of the Institute of Electrical and Electronics Engineers Data Mining Conference Workshop on Text Mining (TextDM)*, (paper accepted for both the main conference and the workshop) San Jose, CA. [http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=989501&tag=1](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=989501&tag=1)
- Blake, C. & Pratt, W. (2000). Multiple Categorization of Search Results: An Extension to Dynamic Categorization. In *Proceedings of the Annual American Medical Informatics Association Symposium (AMIA)*, p81-85, Los Angeles, CA.  
<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2243812/pdf/procamiasymp00003-0116.pdf>

## PEER-REVIEWED ABSTRACTS, POSTERS, PANELS, AND DOCTORAL FORUMS

---

- Souden, M. Blake, C., Twidale, M., Anderson, C. Stelmack, J., (2015) Making sense of big data: Online question answering practices supporting healthcare data re-use. *8th Annual Conference on the Science of Dissemination and Implementation: Optimizing Personal and Population Health*, December 14-15, Washington D.
- Sherman, G., Blake, C. & Lee, J. (2015) Identifying population characteristics from tables in full text articles, *American Medical Informatics Association Symposium*, Nov 14 - 18, San Francisco, CA.
- Blake, C., Souden, M., Anderson, C.L., Twidale, M., and Stelmack, J.E. (2015) Online Question Answering Practices Supporting Healthcare Data Re-use, *Annual Meeting of the American Society for Information Science and Technology (ASIS&T)*, November 6-10, St Louis, MO.  
<http://dl.acm.org/citation.cfm?id=2857186>
- Evans, C., Diesner, J. and Blake, C. (2015), Email Data Analysis as an Alternate Lens into Historical Events, *HASTAC Conference*, May, Lansing, MI.
- Blake, C and Guo, J. (2014) Information behaviors at the edge of reason: the role of uncertainty, science, and culture on environmental policy, *Annual Meeting of the American Society for Information Science and Technology (ASIS&T)*, Seattle, WA.  
<http://onlinelibrary.wiley.com/doi/10.1002/meet.2014.14505101164/abstract>
- Svetlozara, S., Cohen, D. and Blake, C (2014) Exploring Cultural Differences in Language Usage: The Case of Negation, *Annual Meeting of the American Society for Information Science and Technology (ASIS&T)*, October 31 – November 5, Seattle, WA.  
<http://onlinelibrary.wiley.com/doi/10.1002/meet.2014.14505101086/abstract>
- Blake, C. and Gabb, H. (2014) Parameter tuning: Exposing the gap between data curation and effective data analytics, *Annual Meeting of the American Society for Information Science and Technology (ASIS&T)*, October 31 – November 5, Seattle, WA.  
<http://onlinelibrary.wiley.com/doi/10.1002/meet.2014.14505101138/abstract>
- Lee, J. and Blake, C. (2014) Information integration: a case study of air quality in Chicago and St. Louis, *Annual Meeting of the American Society for Information Science and Technology (ASIS&T)* October 31 – November 5, Seattle, WA.  
<http://onlinelibrary.wiley.com/doi/10.1002/meet.2014.14505101154/abstract>
- Ahmed, S., Blake, C., Williams, K., Lenstra, N. and Liu, Q. (2013) Identifying Claims In Social Science Literature Poster at *iConference*, Fort Worth TX. **Honorable mention**  
<https://www.ideals.illinois.edu/bitstream/handle/2142/42061/485.pdf?sequence=2>
- Blake, C., Stanton, J., & Saxenian, A. (2013) Filling the Workforce gap in Data Science and Data Analytics, Workshop at the *iConference*, Fort Worth TX.  
<https://www.ideals.illinois.edu/bitstream/handle/2142/42501/424.pdf?sequence=4>
- Lucic, A., and Blake, C. (2012) Characterizing Authorship Style Using Linguistic Features, *Digital Humanities*, July 19, Hamburg Germany. <http://www.db2012.uni-hamburg.de/conference/programme/abstracts/characterizing-authorship-style-using-linguistic-features/>
- AbdelRahman, S.E. & Blake, C. (2012) UIUC\_RuleBased: A Rule-based Human Interpretation System for Semantic Textual Similarity Task *Workshop on Semantic Evaluation (SemEval)*, *First Joint Conference on Lexical and Computational Semantics, Collocated with NAACL-HLT* June 7-8, Montreal, Canada <http://dl.acm.org/citation.cfm?id=2387726>
- Qin, J., Crowston, K.G., Lesk, M., Blake, C., Mostafa, J., (2012) Data Science and Analytics: What Is in It for iSchools? Workshop panel at the *iConference*, Toronto, Canada.
- Stanton, J., Palmer, C., Blake, C., Farmer, L., & Allard, S. (2012) Brainstorming Data Science @ iSchools, *iConference 2012 Alternative Event Submission*.

- Lucic, A. & Blake, C. (2011) Comparing the Similarities and Differences between Burton Pike's and Stephen Mitchell's translation of The Notebooks of Malte Laurids Brigge by Rainer Maria Rilke, *Digital Humanities (DH 2011)*, July, Palo Alto, CA.  
<http://dh2011abstracts.stanford.edu/xtf/view?docId=tei/ab-221.xml;query=;brand=default>
- Zheng, W. & Blake, C. (2010) Automatic extraction of location relations from text, Abstract at the *iConference*, Champaign IL.  
<https://www.ideals.illinois.edu/bitstream/handle/2142/15046/wuzheng.pdf?sequence=2>
- D'Ignazio, J., Blake, C., & Sonnenwald, D. (2009) Bridging between Scientific Disciplines: Educational Strategies to Meet the Challenge of Managing Information across the Sciences Panel at the 73<sup>rd</sup> Annual Meeting of the American Society for Information Science and Technology (ASIS&T), November 6-11, Vancouver, BC, Canada.  
<http://onlinelibrary.wiley.com/doi/10.1002/meet.2009.1450460138/abstract>
- Blake, C (2008) Beyond Genes, Proteins, and Abstracts: A Framework to Capture Scientific Claims, Abstract in the 2008 *Microsoft eScience Workshop*, Indianapolis, IN.
- Blake, C (2007) In Support of e-science - Shifting from Information Retrieval to Information Synthesis, Abstract in the 2007 *Microsoft eScience Workshop* at RENCI, Chapel Hill, NC.  
<http://people.lis.illinois.edu/~clblake/Papers/EsSciencePres.pdf>
- Blake, C., Nassar, N. & Jones, R. (2007) Using Concepts and Entailment for Passage Retrieval from Biomedical Literature, Abstract at the 2007 *Microsoft eScience Workshop* at RENCI, Chapel Hill, NC. <http://people.lis.illinois.edu/~clblake/Papers/EsSciencePost.pdf>
- Kraus, S., Blake, C. & West, S.L (2007) Information Extraction from Medical Notes, In K.A. Kuhn, J.R. Warren and T.-Y. Leong (Eds.) Abstract in the *Proceedings of the 12th World Congress on Health (Medical) Informatics – Building Sustainable Health Systems (MedInfo)*, p1662-4, Brisbane, Australia.  
<http://people.lis.illinois.edu/~clblake/Papers/KrausBlakeWestMEDINFO2007.pdf>
- Wildemuth, B.M, Blake, C.L., Spurgin, K., Oh, S. Zhang, Y. (2006) Patients' Perspectives on Personal Health Records: An Assessment of Needs and Concerns, Abstract in *Critical Issues in eHealth Research Conference: Toward Quality Patient Centered Care*, Bethesda, MD.  
<http://people.lis.illinois.edu/~clblake/Papers/EHealthAbstract.doc>
- Tenopir, C, Brown, A., Brown, C., Blake, C. (2006), "How Chemists are Really Finding and Using Information in our Digital Environment", Panel at the *Proceedings of 69<sup>th</sup> Annual Meeting of the American Society for Information Science and Technology (ASIST)*, p38, Austin, TX.  
<http://onlinelibrary.wiley.com/doi/10.1002/meet.1450430138/abstract>
- Luo, L. West, D., Marchionini, G. Blake, C (2005) A Study of Annotations for a Consumer Health Portal. Abstract in the *Proceedings of the 5<sup>th</sup> ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL)*, p388, Denver, CO. <http://dl.acm.org/citation.cfm?doid=1065385.1065491>
- Blake, C., & Anderson, C., (2005) The Shift from Information Retrieval to Synthesis, Panel at the *The First i-Conference of the i-school Community Bringing Disciplines to Confront Grand Challenges*, State College, PA. <http://people.lis.illinois.edu/~clblake/Papers/BlakeAnderson.pdf>
- Blake, C (2004) A Text Mining Approach to Enable Detection of Candidate Risk Factors. In M. Fieschi and E. Coiera (Eds.), Abstract in *Proceedings of the 11<sup>th</sup> World Congress on Medical Informatics – Building High Performance Health Care Organizations (Medinfo)*, p1528, San Francisco, CA.
- Pratt, W, Srinivasan, P., Smalheiser, N. & Blake, C (2004) Mining the Literature to Promote Biomedical Discoveries Panel at *The 11<sup>th</sup> World Congress on Medical Informatics (Medinfo) – Building High Performance Health Care Organizations*, San Francisco, CA.
- Blake, C. (2003) A Technique to Resolve Contradictory Answers. Abstract in M. Maybury (ed) *New Directions in Question Answering, Papers from the AAAI Spring Symposium*, Technical Report SS-03-

07, American Association for Artificial Intelligence, p49-52, Stanford, CA.

<http://www.aaai.org/Library/Symposia/Spring/2003/ss03-07-008.php>

Blake, C. & Pratt, W. (2002). Automated Information Extraction and Analysis for Information Synthesis. Abstract in the *Proceedings of the Annual American Medical Informatics Association Symposium (AMIA)*, San Antonio, TX. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2244281/>

## RESEARCH ACTIVITIES WITHOUT PEER REVIEW

---

Blake, C. and Gabb, H. (2015) Parameter tuning: Exposing the gap between data curation and effective data analytics, Poster at the *Graduate School of Library and Information Science Research Showcase*. <https://www.asis.org/asist2014/proceedings/submissions/posters/306poster.pdf>

Guo, J., Lu, Y., Mori, T. and Blake, C. (2015) Expert-Guided Contrastive Opinion Summarization for Controversial Issues, Poster at the *Graduate School of Library and Information Science Research Showcase*.

Lee, J. and Blake, C. (2015) Information integration: a case study of air quality in Chicago and St. Louis, Poster at the *Graduate School of Library and Information Science Research Showcase*.

Svetlozara, S., Cohen, D. and Blake, C. (2015) Exploring Cultural Differences in Language Usage: The Case of Negation, Poster at the *Graduate School of Library and Information Science Research Showcase*.

Ludäscher & Blake, C. (2015) CIRSS – Active Current Research Projects Poster at the, *Graduate School of Library and Information Science Research Showcase*, University of Illinois.

Blake, C. (2014) Socio-technical Data Analytics, Presentation at the *Graduate School of Library and Information Science Research Showcase*, University of Illinois, IL, February 2014

Lucic, A. & Blake, C. (2014) Automatically Summarizing Medical Literature, Poster at the *Graduate School of Library and Information Science Research Showcase*, University of Illinois, IL.

Surbeck, E. & Blake, C. (2014) The Impact of Ground Ozone on Asthma: A Case Study using Project Indicator in Champaign-Urbana Poster at the *Graduate School of Library and Information Science Research Showcase*, University of Illinois, IL, February 2014.

Palmer, C. & Blake, C. (2011-2014) CIRSS – Active Current Research Projects Poster at the, *Graduate School of Library and Information Science Research Showcase*, University of Illinois.

Blake, C., Gasser, L. & Downie, S. (2011) Institutional Organizations that Foster Interdisciplinary Education and Research, *Innovation Summit*, University of Illinois, IL.

Blake, C. (2008) Uncovering the Creative Processes in Chemistry: Implications for Discovery System Design, *Symposium on Computational Approaches to Creativity in Science*, Stanford, March 29–30, 2008.

West, S., Liu, Z., McKoy, N., Oertel, M., Blake, C., Schwartz, B., Ochart, F., Carey, T. (2007) Use of electronic medical records and administrative claims data for assessing type 2 diabetes care, *UNC DEiIDE project report for Grant AHRQ 290-05-0040-1*.

Blake, C. (2003) Information Synthesis: A Mixed-Initiative Meta-Analytic Approach to Facilitate Knowledge Discovery from Scientific Text, *Doctoral Thesis in Information and Computer Science, Specialization: Information Access and Management*, University of California, Irvine.

Blake, C. (2001). Text Mining – A Survey. *Manuscript and Oral Presentation to Advance to Ph.D. Candidacy*, University of California, Irvine, CA.

Blake, C. & Merz, C. (1998) UCI Repository of machine learning databases, Department of Information and Computer Science, University of California at Irvine, CA.

Blake, C., Keogh, E. and Merz, C. J. (1999) UCI repository of machine learning databases. Department of Information and Computer Science, University of California at Irvine, CA.

Blake, C. (1996). An Overview of Quinlan's C4.5 Algorithm. *BHP TechNote* BHP/CP/N/015, Melbourne Research Laboratories, Australia.

Blake, C. (1996). Data Access and Data Integrity in a Geographically Distributed Environment- 'Global Information Warehouse' *BHP TechNote* BHP/CP/N/029, Melbourne Research Laboratories, Australia.

## **RESEARCH REPORTED IN THE LOCAL AND POPULAR PRESS**

---

GSLIS Magazine (2105) Blake builds Claim Framework to analyze and synthesize medical research, <https://www.uillinois.edu/common/pages/DisplayFile.aspx?itemId=221303>

Impact Illinois 2015 (2015) SODA for real-world solutions, p8-9  
<https://www.uillinois.edu/common/pages/DisplayFile.aspx?itemId=221303>

Goldberg, B. (2012) A snapshot of librarianship's never-ending metamorphosis, What's New in LIS Schools, American Library Association. <http://americanlibrariesmagazine.org/2012/08/21/whats-new-in-lis-schools/>

New Faces of Illinois (2010) Sept 17, 2009. <http://news.illinois.edu/ii/09/0917/newfaces.html>

R. Pordes et al (2008), New science on the Open Science Grid, 125, pg 012070, Journal of Physics: Conference Series

Green, Karen (2008) Feature – Claim jumping through scientific literature, Issue 76, International Science Grid (subsequently mentioned in "American Libraries Direct." June 4, 2008).  
<http://renci.org/research/claim-jumping/>

Heavey, A. (2007). Open Science Grid Research Highlights July 2007 – June 2008.

## **PRESENTATIONS - INVITED SPEAKER**

---

“Summarizing findings with the Claim Framework”. Department of Computing and Information Systems, University of Melbourne, July 2015.

“Collaborative Data Analytics”, Invited speaker, Department of Computer and Systems Sciences, Stockholm University. May 2014.

“Representing Biological Responses to Chemical Exposures using the Claim Framework”, Invited Speaker at University of Wisconsin, Milwaukee, March 2014

“Living on the edge of science: A pilot study of concept formation in toxicology literature” Invited Speaker at the Symposium on Cognitive Systems and Discovery Informatics, Carnegie Mellon University's Silicon Valley Campus, San Jose, CA, June 21-22, 2013

“Using Secondary Information to inform Evidence-Based Discovery”, *Health Information Technology Center Workshop*, University of Illinois, NCSA November 2, 2012.

“Claim Jumping : Bridging disciplinary boundary's using the Claim Framework” Making Sense of Biological Systems: Using Knowledge Mining to Improve and Validate Models of Living Systems, Bozeman Montana, August 23-25, 2012.

“How Informatics Can Help To Address Grand Challenges in Health and the Environment” Bioinformatics and Medical library IT interns, June 2012.

“Informatics methods to scale up Risk Assessments” Invited Speaker at Environmental Protection Agency (EPA), March 2012.

“How informatics can help to address grand challenges in health and the environment”, Illinois Informatics Institute Spring Lecture Series, February 2012.

“Beyond genes, proteins, and abstracts: Identifying scientific claims from full-text biomedical articles”, Invited Speaker at Computational Bioscience Program at the University of Colorado Denver, July 2011.



- “Towards Evidence-Based Discovery”, Graduate School of Library and Information Science (GSLIS), American Society for Information Science & Technology Student Chapter, Jan 2011.
- “Towards Evidence-Based Discovery”, Environmental Change Institute Symposium, November 2010.
- “Information Synthesis”, Invited Speaker, Accountability through Algorithm: Developing the Field of Computational Journalism organized by James T. Hamilton and Fred Turner, and hosted by the Center for Advanced Study in the Behavioral Sciences, Summer 2009.
- “Uncovering the Creative Processes in Chemistry: Implications for Discovery System Design”, Symposium on Computational Approaches to Creativity in Science, Stanford, March, 2008.
- “Evidence-Based Discovery”, Invited Speaker at the University of Illinois, School of Library and Information Science, December 2008.
- “Claim Jumping through Scientific Literature”, Invited Speaker at the Renaissance Computing Institute for Glaxo Smith Cline, October 2008.
- “Reinventing Science Librarianship – Education for new Roles”, Invited Speaker at the Reinventing Science Librarianship: Models for the Future forum, co-sponsored by the ARL and the Coalition for Networked Information, October 2008.
- “Text Mining Scientific Literature”, Guest panelist on Prospects for Science on the Open Science Grid at the OSG All Hands Consortium Meeting, Chapel Hill, March 2008.
- “Information Synthesis”, Invited Speaker at the School of Information University of Michigan, April 2008.
- “Uncovering Creative Processes in Chemistry: Implications for Discovery System Design”, Invited Speaker at the NSF Symposium on Computational Approaches to Creativity in Science, San Jose California, March 2008.
- “Information Synthesis”, Invited Speaker at Arizona State University, School of Computing and Informatics, March 2008.
- “Text Mining”, Invited panelist at the XXVII Annual Charleston Conference, Nov 2007.
- “Text Mining”, Invited speaker at the School of Biology, University of North Carolina at Greensboro, 2006.
- “Text Mining”, Invited speaker at the Department of Computer Science, University of North Carolina at Chapel Hill, 2004.
- “Text Mining”, Invited speaker at the Department of Pharmacy, University of North Carolina at Chapel Hill, 2004.
- “Information Synthesis”, Invited speaker at the Department of Emergency Medicine, University of North Carolina at Chapel Hill, 2004.
- “Information Synthesis”, Invited speaker at the Department of Chemistry, University of North Carolina at Chapel Hill, 2002.
- “Information Synthesis”, Invited speaker at the School of Library & Information Science and the College of Nursing, University of Iowa, 2002.
- “Information Synthesis”, Invited speaker at the Department of Computer Science, University of Missouri, 2002.
- “Information Synthesis”, Invited speaker at the School of Information and Library Science , University of North Carolina at Chapel Hill, , 2002.

#### **PRESENTATIONS – CONFERENCES AND WORKSHOPS**

---

- “Evidence-based discovery”, iConference, Newport Beach, CA, March 23-27, 2015.

- “Socio-technical Data Analytics” Research Showcase at the University of Illinois, Graduate School of Library and Information Science , April 2014.
- “Biological responses to chemical exposure: Case studies in how to manage ostensible inconsistencies using the Claim Framework”, Inconsistency Robustness’14, Stanford University July 29-31, 2014.
- “Using Secondary Information to inform Public Policy”, *Internet, Politics, Policy, Big Data Big Challenges*, UK September 20-21, 2012.
- “Text Mining in a Complex World”, Graduate School of Library and Information Science (GSLIS) Corporate Round Table, March 2011.
- “Claim Framework”, Research Showcase at the University of Illinois, Graduate School of Library and Information Science , April 2010.
- “Beyond Genes, Proteins, and Abstracts: A Framework to Capture Scientific Claims”, Microsoft eScience Workshop, Indianapolis, IN, December 2008.
- “In Support of e-science - Shifting from Information Retrieval to Information Synthesis”, Microsoft eScience Workshop at RENCI, Chapel Hill, NC, 2007.
- “UNC-CH at DUC 2007: Query Expansion, Lexical Simplification, and Sentence Selection strategies for Multi-Document Summarization, Presentation at *Document Understanding Conference (DUC) 2007*, Rochester, NY, 2007.
- “A Comparison of Document, Sentence and Term Event Spaces”, Paper presentation at the Joint 21st International Conference on Computational Linguistics (COLING) and the 44th Annual Meeting of the Association for Computational Linguistics (ACL), Sydney, Australia, 2006.
- “Scientific Discovery: A View from the Trenches”, Paper presentation at the 9th International Conference on Discovery Science, Barcelona, Spain, 2006.
- “Information Synthesis: A New Approach to Explore Secondary Information in Scientific Literature”, Paper presentation at the ACM/IEEE-CS Joint Conference on Digital Libraries, Denver, CO, 2005.
- “Cataloging On-Line Health information: A Content Analysis of the NC Health Info Portal”, Paper presentation at the Annual American Medical Informatics Association Symposium (AMIA), Washington DC, 2005.
- “The Shift from Information Retrieval to Synthesis”, Panel at The First i-Conference of the i-school Community Bringing Disciplines to Confront Grand Challenges, State College, PA, 2005.
- “How Chemists are Really Finding and Using Information in our Digital Environment”, Panel at the 69<sup>th</sup> Annual Meeting of the American Society for Information Science and Technology (ASIST), Austin, TX, 2005
- “Mining the Literature to Promote Biomedical Discoveries Panel” at the 11<sup>th</sup> World Congress on Medical Informatics (Medinfo) – Building High Performance Health Care Organizations, San Francisco, CA, 2004.
- “Information Synthesis: A Process used by Scientists in Medicine and Public Health to Overcome Information Overload”. In Doctoral Forum of the Fourth International Conference on Conceptions of Library and Information Science: Emerging Frameworks and Methods (CoLIS 4), Seattle, WA, 2002.
- “Collaborative Information Synthesis”, Paper presentation at The 65<sup>th</sup> Annual Conference of the American Society for Information Science and Technology (ASIST), Philadelphia, PA, 2002.

- “A Semantic Approach to Identify Candidate Treatments from Existing Medical Literature”, Paper presentation at the Mining Answers from Texts and Knowledge Bases AAAI Spring Symposium, Menlo Park, CA, 2002.
- “Better Rules Fewer Features: A Semantic Approach to Selecting Features from Text”, Paper presentation at The Institute of Electrical and Electronics Engineers Data Mining Conference (IEEE-DM), San Jose, CA, 2001 .
- “The Nation’ Breast Cancer Research Portfolio: A View from 30,000 ft.” Presentation at the Avon Symposium, University of California, Irvine, CA.
- “Multiple Categorization of Search Results: An Extension to Dynamic Categorization”, Paper presentation of the Annual American Medical Informatics Association Symposium (AMIA), Los Angeles, CA, 2000.

#### **FUNDED RESEARCH - CURRENT**

---

- A Big Data-Theoretic Approach to Quantify Organizational Failure Mechanisms in Probabilistic Risk Assessment
  - National Science Foundation
  - Principal Investigator: Zahra Mohaghegh, Co-PIs: Cheri Ostroff (Management, Univ of South Australia) and Catherine Blake
  - Total Award Amount: \$899,663
  - Total Award Period Covered: May 5, 2015 – May 4, 2020
- Information Forum requirements to promote knowledge capture, knowledge sharing, and community interaction on the VHA Data Portal (renewal)
  - Veterans Affairs
  - Principal Investigators: Catherine Blake and Maria Souden, Co-PI Michael Twidale
  - Total: \$111,978
  - Period: October 1, 2015 – September 30, 2016
- Developing a model for Sociotechnical Data Analytics (SODA) Education
  - Institute for Museum and Library Services (IMLS) RE-05-12-0054-12, building institutional capacity category. The grant supports the development of a SODA specialization (curriculum, website, and promotional materials) which was approved in October 2012 and the first cohort graduated in Spring 2013.
  - Principal Investigator: Catherine Blake
  - Total: \$712,338 (Award: \$498,777; Matching: \$213,561)
  - Period: July 1, 2012 – June 30, 2015 (No Cost extension to June 30, 2016)
- Digital Innovation Leadership
  - University of Illinois
  - PI: Jon Gant, Co-PIs, Extension, Illinois Extension, Team member: Blake, et al.
  - Total Award Amount \$300,000
  - Total Award Period Covered: 2 years
  - Period: September 15, 2014 – August 14, 2016

#### **FUNDED RESEARCH - COMPLETE**

---

- Information Forum requirements to promote knowledge capture, knowledge sharing, and community interaction on the VHA Data Portal
  - Veterans Affairs
  - Principal Investigators: Catherine Blake and Maria Souden, Co-PI Michael Twidale
  - Total: \$92,346
  - Period: August 15, 2014 – August 14, 2015

- Collaborative Information Seeking and Data Use
  - University of Illinois
  - Principal Investigator: Michael Twidale, Co-PI Catherine Blake
  - Total: \$7,000
  - Period: January 2014 – July 2014
- Towards Evidence-Based Discovery
  - National Science Foundation, Division of Information and Intelligent Systems (IIS), Computer & Information Science & Engineering (CISE), #0812522
  - Principal Investigator: Catherine Blake
  - Total: \$449,317
  - Period: September 2008 – August 2011 – No cost extension to August 2012
- Evidence-Based Discovery in Environmental Literature
  - Faculty Fellowship program, Environmental Change Institute, UIUC
  - Principal Investigator: Catherine Blake
  - Total: \$10,000
  - Period: April 2010 – May 2011
- The Future of IT for State Farm.
  - State Farm Insurance Companies (#2010-06961)
  - Principal Investigator: Dr. Jon Gant, Co-PIs: Catherine Blake, Michael Twidale
  - Total: \$44,295
  - Period: 1 August 2010 – 31 January 2011
- Linking Data Collected by Hospitals and Ambulance Services
  - Gillings Innovation Labs
  - Principal Investigator : David Richardson, UNC School of Public Health
  - Total: \$595,000 (Blake, 10% effort)
  - Period: August 2008 – July 2010
- Developing the Field of Computational Journalism
  - Center for Advanced Study in the Behavioral Sciences (CASBS) in Palo Alto, CA.
  - Principal Investigators: James Hamilton (Duke University), Fred Turner (Stanford)
  - One-week summer workshop, Summer 2009
- DigCCurr II: Extending an International Digital Curation Curriculum to Doctoral
  - Institute of Museum and Library Services
  - Principal Investigator: Helen Tibbo; Co-Investigators: Cal Lee, Gary Marchionini, Catherine Blake, Jeffrey Pomerantz, Jane Greenberg, Javed Mostafa, Diane Kelly
  - Total: \$878,634 (Blake, 2% effort)
  - Period: September 2008 – August 2012
- Clinical and Translational Science Award (CTSA)
  - National Institutes of Health
  - Principal Investigator: Formally Paul Watkins, now Etta Pisano
  - Total: \$61,000,000 (Blake, Research assistant support 2008-9)
  - Period: June 2008 – May 2018
- Claim Jumping through Scientific Literature
  - Renaissance Computing Institute (RENCI) Faculty Fellows Program
  - Principal Investigator : Catherine Blake
  - Total: \$50,000 + in-kind programmer support
  - Period: July 2007 – June 2008

- Automated Information Extraction from Electronic Health Records
  - Computing Research Association Multidisciplinary Research Opportunity for Women (MRO-W) Program
  - Principal Investigator : Catherine Blake, Co-PI Sue West
  - Total: \$22,500
  - Period: Sept 2007 – August 2008
- An Online Text Mining Evaluation Environment (OnTeMEE)
  - UNC Junior Faculty Award
  - Principal Investigator : Catherine Blake
  - Total: \$7,500
  - Period: January 2007 – December 2007
- Text Mining Chemistry Literature
  - The Center for Environmentally Responsible Solvents and Processes, STC Program of the National Science Foundation under agreement CHE-9876674
  - Principal Investigator: Catherine Blake
  - Total: \$68,652
  - Period: August 2004 – June 2006
- Text Mining Biomedical Literature
  - IBM SUR Grant - Principal Investigators: John Smith (Computer Science), Paul Jones (ibiblio) and John Reuning (ibiblio), Catherine Blake
  - Total: \$17,115 (Blake's portion only)
  - Period: August 2004 – June 2006
- Annotation Project
  - Microsoft Research
  - Principal Investigators: Gary Marchionini, Paul Solomon and Cathy.Marshall
  - Total: \$100,000 (Blake, Summer salary)
  - Period: January 2004 – June 2005
- Using Scientific Text to Identify Breast Cancer Risk-Factors.
  - California Breast Cancer Research Program Dissertation Award
  - Principal Investigator: Catherine Blake
  - Total: \$29,136
  - Period July 2002 – June 2003
- Using Scientific Text to Identify Breast Cancer Risk Factors
  - University of California, Irvine Multi-Investigator Grant
  - Principal Investigators: Wanda Pratt and Tammy Tengs
  - Period June 2001 – June 2002

## **TEACHING RECORD**

---

### **University of Illinois at Urbana Champaign, iSchool**

- Text Mining (LIS590TX), Spring 2010, Fall 2010 (online), Spring 2012, 2013, 2014, 2015
- Evidence-based Discovery (LIS590AG)
  - List of instructors ranked as excellent (online) Fall 2014
  - Core in the SODA specialization Fall 2013 (Inaugural offering), Fall 2014, 2015 (hybrid, both on campus and online)
- Foundations in Socio-technical Data Analytics (LIS590AD)
  - Core in the SODA specialization, Spring 2013 (Inaugural offering), Spring 2014, Spring 2015(Co-taught with V.Stodden)

- Socio-technical Data Analytics Practicum (LIS591)
  - Developed practicum requirements for the SODA specialization, coordinated experiences with industry and academic partners
  - Spring 2013, Fall 2013, Summer 2014, Fall 2014, Summer 2015, Fall 2015
- Introduction to Databases(LIS490DB)
  - List of instructors ranked as excellent Fall 2010
  - Fall 2009, 2010, 2011, 2012, Spring 2010 (online) & Spring 2011(online), Fall 2015
- Foundations Information Processing in LIS (LIS452)
  - Fall, 2011 with Wu Zheng, new course on Java
- Doctoral Research Supervision
  - Ana Lucic, Graduate School of Information and Library Science, *Comparison relations from text* (Advisor, Fall 2012-current)
  - Henry Gabb, Graduate School of Information and Library Science, *Topic TBD* (Advisor, Fall 2012-current)
  - Jinlong Guo, Graduate School of Information and Library Science, *Topic TBD* (Advisor, Fall 2013-current)
  - Joo Ho Lee, Graduate School of Information and Library Science, *Topic TBD* (Advisor, Fall 2013-current)
  - Wu Zheng, Graduate School of Information and Library Science, *Automatic detection of location relationships*. (Advisor since Spring 2010, ABD)
  - Tiffany C. Chao, Graduate School of Information and Library Science, *Methods metadata: curating scientific research data for reuse* (Advisor: Carole Palmer, committee members: Catherine Blake, Jane Greenberg, Michelle M. Wander, Graduated Spring 2015)
  - V.G.Vinod Vydiswaran, Computer Science, *Modeling and predicting trustworthiness of online textual information*. (Advisors C.Zhai & D. Roth (CS), Committee member: Catherine Blake, Wai-Tat Fu (CS), Peter L. Pirolli (Palo Alto Research Center) Graduated Fall 2012)
  - Dae Hoon Park (2011/12), Computer Science, Primary Advisor: C.Zhai
  - Brant Chee, GSLIS, Committee Member, *Pharmacovigilance: Exploring Alternative Data Sources with Machine Learning and Visualization Techniques* (Advisor: B.Schatz, Graduated Fall 2011)
  - Qui Wei, GSLIS, Committee Member, *Information Fusion in Taxonomic Descriptions* (Advisor: B.Heidorn, Graduated Spring 2011)
- Doctoral Research – Proposal
  - Ana Lucic (2015), Blake role, chair and director of research
  - Mumtaz Vauhkohen (2015), Education , Blake role committee Member
  - Wu Zheng (2013), Blake role, chair and director of research
  - Tiffany Chao (2012), Primary Advisor: C.Palmer Blake role committee Member
- Doctoral Research – Field Exam
  - Henry Gabb (2015), Primary Advisor: C. Blake
  - Jinsoek Kim (2015), Primary Advisor J. Diesner
  - Shubhanshu Mishra (2015), Primary Advisor J. Diesner
  - Ana Lucic (2014), Primary Advisor: C.Blake
  - Ahmed,Shameem (2013), Primary Advisor: K.Williams

- Peter Organisciak (2012), Primary Advisor: M.Efron
- Wu Zheng (2010), Primary Advisor: C. Blake
- Doctoral Research – Annual reviews
  - Wu Zheng (2010-current), Primary Advisor: C. Blake
  - Ana Lucic (2011 – current), Primary Advisor: C.Blake
  - Henry Gabb (2012 – current), Primary Advisor: C.Blake
  - Kenney Guo (2013 – current), Primary Advisor: C.Blake
  - Joocho Lee (2013 – current), Primary Advisor: C.Blake
  - Caryn Anderson (2015) , Primary Advisor: C.Blake
  - Aseel Addawood (2015) , Primary Advisor: C.Blake
  - Garrick Sherman (2015), Primary Advisor M. Efron
  - Ming Jiang (2015), Primary Advisor: J. Diesner
  - Jinseok Kim (2014), Primary Advisor: J. Diesner
  - Tiffany Chao (2012-current), Primary Advisor: C.Palmer
  - Cheryl Thompson (2012-2014), Primary Advisor: C.Palmer
  - Kahyun Choi (2012), Primary Advisor: S.Downie
- Certificate of Advanced Study (CAS)
  - Ana Lucic (Summer 2011, graduated with CAS degree in Spring 2012)
  - Eric Johnson (Fall 2011, Spring 2012)
  - Walker Weyerhaeuser (Spring 2011, Fall 2011)
- Funded MS Research Assistants(degree, graduation year), semester supervised
  - Maxwell Isenholt (Fall 2014, Spring 15) SODA Research Assistantship
  - Sveta Stoytcheva (Summer 2014), SODA Research Assistantship
  - Max McKittrick (Fall 2014 as undergraduate, now in MS program), SODA RA
  - Shan Chen (Spring 13), SODA research assistantship
  - Zhaoyu Wang (Spring 13), SODA research assistantship
  - David Failing (2013/14), SODA research assistantship
  - Elizabeth Surbeck (2013/14), SODA research assistantship
  - Amber Brookshire (Spring 2011), MLS, 2012
  - Kyle Painter (MLS, 2011), Summer 2010, Fall 2010, Spring 2011
  - Aaron Fleisher (MLS, 2011), Summer 2010
  - Walker Weyerhaeuser (MLS, 2011), Summer 2010
- MS Practicum (degree, graduation year), semester supervised
  - Yidan Sun (Bioinformatics, 2016, Summer 2015) SODA practicum
  - Andrew Stephens (MS, 2016, Summer 2015) SODA Practicum at Waterborne Environmental (<http://www.waterborne-env.com/>)
  - Kristen Lundberg (MS, 2015, Spring 2015) SODA practicum
  - Sophie Young (MS, 2015, Spring 2015) SODA practicum (with Jon Gant)
  - Minhao Jiang (MS, 2015, Spring 2015) SODA practicum
  - Jessica Tieman (MS 2014, Summer 2014), SODA practicum
  - Svetlozara Stoytcheva (MS 2015, Summer 2014), SODA practicum
  - Hengyi Fu (MS 2013, Spring 2013)
  - Qiyuan Lui (MS, 2013), SODA practicum
  - Shan Chen (MS, 2013), SODA practicum
  - Zhaoyu Wang (MS, 2013), SODA practicum
  - Erin Griffin Sackett-Hermann (MLS, 2011), Summer 2011, with Melissa Cragin

- Independent Study Supervised (degree, graduation year) semester supervised
  - JooHo Lee (PhD), Spring 2014
  - Kenney Guo (PhD), Spring 2014
  - Andrea Thomer (PhD) Fall 2013
  - Cheryl Thompson (PhD) Fall 2013
  - Jamie Keen (BA) Spring 2013
  - Ana Lucic (PhD) Spring 2013
  - Henry Gabb (PhD) Fall 2013
  - Tom Lynch (MLS, 2012) Fall 2011
  - Dae-Hoon Park (CS PhD, 2014) Fall 2011
  - Eric Johnson (MLS, 2012) Fall 2011
  - Ellen Dennis (MLS, 2011) Spring 2011
  - Wu Zheng (PhD) Fall 2009
  - Candler Armstrong (PhD, Sociology) Fall 2009

### University of North Carolina at Chapel Hill, School of Information and Library Science

- Introduction to Database Concepts and Applications (INLS523), Spring and Fall 2004, Fall 2005, Spring and Fall 2006, Spring and Fall 2008
- Text Mining (INLS613), Spring 2006, Fall 2007, Spring 2009
- Evidence-Based Discovery (INLS890), Spring 2009
- Information Tools (INLS461), Fall 2005, Fall 2006
- Doctoral Research Supervised
  - Wiriya Phokhwang, *Information Needs and Uses of nurses in Thailand*, UNC, College of Nursing, Ph.D., 2008. Role: Committee Member.
  - John MacMullen *Contextual analysis of variation and quality in human-curated gene ontology annotations*, UNC, School of Information and Library Science, Ph.D., 2007. Role: Committee Member.
  - Megan Winget UNC, School of Information and Library Science, Ph.D., 2006, Role: Committee Member prior to advancing to candidacy.
  - Robert Fu UNC, MS Statistics, 2007, Ph.D. Information and Library Science, Role: Research advisor 2004
  - Chen Qi UNC, Computer Science, Ph.D. 2009, Role: Research Advisor 2004
  - Nancy Baker UNC, School of Information and Library Science, Ph.D., 2010, Role: Research advisor 2004/2005
  - Leo Cao UNC, School of Information and Library Science, Ph.D., 2011, Role: Research advisor Fall 2005
  - Anita Lunga (Computer Science, Duke University) Ph.D. 2008, Role: Research advisor Fall 2005
- Master's Papers Supervised
  - Ryan Jones, *Evaluation of the Effectiveness of Cosine Similarity in Predicting Relevance between Paired Citing and Cited Sentences*, Spring 2009
  - Amol Bapat, *A system to extract abbreviation-expansion pairs from biomedical literature*, Spring 2009
  - Monte Evans *A New Approach to Journal and Conference Name Disambiguation through K-Means Clustering of Internet and Document Surrogates*, Spring 2009



- Benjamin B Pennell, *Effectiveness of Text Representations in the Automatic Classification of Regional Game Design Trends in Video Game Reviews*, Fall 2008
- Andreas K. Orphanides, *Extraction of natural-language dates and comparison of dates in hypothesis and text to identify negative textual entailment*, Spring 2008
- Sarah K. Peterson, *My Kingdom for a (Digital) Horse: Perceptions of Virtual Property in Second Life*, Spring 2008
- Stephanie C. Adamson, *Soliciting user contribution in the modern digital library: a critique framework for evaluating methods and a case study recommendation for a digital library of historical materials*, Spring 2008
- Kaye M Balke, *Overdue for an Overhaul: The Redesign of the Durham County Public Library's Web Site*, Spring 2007
- Bryan E Howard, *Evaluation of Text Classification Accuracy*, Fall 2007
- E. Ashley Rogers Brown, *Building a Better Mousetrap: Capturing New Data in ISI Journal Citation Reports and Local Journal Utilization Reports to Support Academic Collection Managers*, Spring 2006
- Independent Study Supervised (degree, graduation year), semester supervised
  - Scott Kraus (MSIS, 2010) , Fall 2008
  - Monte Evans (MSIS, 2009), Fall 2008
  - Amol Bapat (MSIS, 2009), Fall 2008
  - Adrienne MacKay (MSIS, 2008), Spring 2008
  - Cory Lown (MSIS, 2007), Spring 2007
  - Julia Kampov (MSLS, 2007), Spring 2007
  - Sayan Chakraborty (MSIS, 2006), Spring 2006
  - Margaret Harrelson (MSIS, 2007), Fall 2005
  - Ashley Brown (MSLS, 2005), Fall 2005
  - Lewis Churchill (Ph.D., 2009), Spring 2004
- Funded Graduate Research Assistants
  - Carrie Dundas-Lucca (MSLS, 2010), Spring 2009
  - Christopher Doty (MSIS, 2010) , Fall 2008- Spring 2009
  - Annie Chen (MSIS, 2010), Fall 2008- Spring 2009
  - Ryan Jones (MSLS, 2009), Fall 2007- Spring 2008
  - Amol Bapat (MSIS, 2009), Fall 2007- Spring 2008
  - Joshua Purvis (MSIS, 2007), Spring 2006
  - Toshiba Burns-Johnson (MSIS, 2007), Fall 2005-Spring 2007
  - Meredith Randall (MSLS, 2007), Fall 2005-Spring 2007
  - Chris Maier (MSIS, 2006), Fall 2004 - Spring 2005
- Funded Undergraduate Research Assistants
  - Laura Dickson, Public Health, Fall 2007-Spring 2008
  - Anna Mansour, Information and Library Science, Fall 2007-Spring 2008
  - Lynnelle Fowler-McDonald, Fall 2007
- Field Experiences Supervised (degree, graduation year) , semester supervised
  - Vedana Vaidhyanathan (MSLS, 2007), Fall 2007
  - Kaye Balke (MSIS, 2007), Fall 2006

- Ashley Brown (MSLS, 2006), Spring 2005

### **University of California, Irvine, School of Information and Computer Science**

- Instructor of problem solving and programming techniques - Summer 2000 and 1999
- Software engineering Teaching Assistant - Winter '95, '98, Fall '94, '97, Spring '95

### **NATIONAL SERVICE**

---

#### **Program Committee Activities**

- Paper co-chair, American Society for Information Science & Technology Annual Meeting, Copenhagen, Denmark, October 14-18, 2016
- Program Committee member, iConference Program Committee, Berlin, Germany, 2014; Irvine, CA 2015; Philadelphia, PA 2016
- Poster Chair, iConference Fort Worth, TX, 2013
- Program Committee Member, The 7th Linguistic Annotation Workshop & Interoperability with Discourse, ACL Workshop, Sofia, August 8-9, 2013
- Program Committee Member, Detecting Structure in Scholarly Discourse (DSSD) Workshop that will take place in conjunction with ACL 2012, to be held on Jeju Island, Korea on July 8-14, 2012.
- Program Committee Member, 8<sup>th</sup> International conference on Language Resources and Evaluation (LREC) (2012)
- Program Committee Member, 73<sup>rd</sup> American Society for Information Science & Technology Annual Meeting (2010)
- Program Committee Member, 5th International Conference on e-Social Science (2009), National Centre for e-Social Science (UK) and German Social Science Infrastructure Services (GESIS).
- Program Committee Member, Association of Computational Linguistics Workshop on Biology Natural Language Processing (ACL BioNLP) (2007, 2008, 2009, 2010)
- Program Committee Member, Symposium of the American Medical Informatics Association (AMIA) (2006)
- Program Committee Member, Conference on Information and Knowledge Management (CIKM) (2006, 2007)

#### **Scientific Committee**

- 9th International Conference on Language Resources and Evaluation, 2014

#### **External Review Boards**

- National Institute of Environmental Health Sciences (NIEHS) (2015)
- Portuguese Foundation for Science and Technology (FCT) (2012)
- National Science Foundation (NSF) – Social Computational Systems (2011)
- National Institute of Environmental Health Sciences (NIEHS)/ National Institutes of Health (NIH) Review Panel (2008)
- Agency for Healthcare Research and Quality (AHRQ) Review Panel (2007)
- NSF Review Panel – Intelligent Information Integration (2007)
- Maryland Industrial Partnerships Program (MIPS)
- Social Science and Humanities Research Council of Canada

### **Professional Society Leadership**

- Faculty mentor at the iConference doctoral colloquium (2012)
- American Society for Information Science and Technology (ASIST), Chair Elect 2008-2009, chair 2009-2010, SIG-MED (soon to become SIG-HEALTH)

### **Editorial and Advisory Boards**

- Editorial board of the Journal of Biomedical Discovery and Collaboration (2005-current)
- Advisory Board: Curriculum Development: Digital Libraries PIs: E.Fox (Virginia Tech), B.Wildemuth (UNC CH), J. Pomerantz (UNC CH), (2006-2009)
- Advisory Board of NEH proposal by Gregory Crane (Tufts University) (June 2004-2006)
- Advisory Board of IMLS proposal by Kate Oliver (Medical Library, Johns Hopkins University) and Wendy Scott (North Carolina State University Libraries) “Recruiting Expertise: 21st Century Academic Research Library”

### **Conference and Chapter Reviewing**

- American Medical Informatics Association Symposium (2002, 2015)
- Joint Conference on Digital Libraries (2004,2013)
- Symposium American Medical Informatics Association (2001-2009,2011, 2014)
- Biomedical Natural Language Processing (2007, 2008, 2009)
- Conference on Information and Knowledge Management (2004-2005,2007, 2008, 2009)
- Conference of the Special Interest Group on Information Retrieval (SIGIR) (2004-2008)
- American Society for Information Science & Technology (2005, 2010)
- Annual Review of Information Science and Technology (2007)

### **Journal Reviewer**

- Journal of Biomedical Informatics (2008, 2012, 2013, 2015)
- Journal of American Medical Informatics Association (2015)
- Journal of the American Society for Information Science & Technology (2005-current)
- Transactions on Interactive Intelligent Systems (2011)
- Journal of Information Processing and Management (2005-2006, 2011,2012)
- Bioinformatics (2008)
- Information Retrieval (2007)
- Transactions of Information Systems (2007)
- Decision Support Systems (2007)
- Data and Knowledge Engineering (2005)

### **LOCAL SERVICE**

---

#### **University of Illinois, Urbana-Champaign**

- GSLIS Master’s and Certificate of Advanced Study admissions committee (2009/10, 2015/16)
- GSLIS Doctoral Studies Committee (2014/15), DSC Chair (2012/13, 2013/14)
- GSLIS Bioinformatics Coordinator (2012-15)
- Committee on Research and Education at the NCSA (2015)
- University of Illinois Visioning Future Excellence in Health and Wellness (Spring, 2013)
- UIUC Campus Budget Oversight Committee (2010-2012)

- GSLIS Executive Committee (2010/11, 2011/12)
- GSLIS Awards Committee (2011/12)
- UIUC Illinois Informatics Initiative Director Search Committee (Spring 2011)
- GSLIS Research Computing Programmer/Manager Search Committee (2011)
- UIUC Health Education Committee (2010-2011)
- UIUC Institutional Organizations that Foster Interdisciplinary Education and Research, Illinois Innovation Summit (with Gasser and Downie, 2010-11)
- GSLIS Research Showcase Committee Member (2009-2010)

#### **University of North Carolina at Chapel Hill – School of Information and Library Science**

- Co-Chair Ontology and Data Modeling Committee, University of North Carolina's Clinical and Translational Science Award (2008-2009) ([bioivlab.ils.unc.edu/projects/CTSA](http://bioivlab.ils.unc.edu/projects/CTSA))
- UNC Developing Evidence to Inform Decisions about Effectiveness Member (DEcIDE) (2005 – 2009)
- Member of the Center for Environmentally Responsible Solvents and Processes (CERSP) (2005-2007) ([www.nsfstc.unc.edu](http://www.nsfstc.unc.edu))
- Informatics Committee (Fall 2008, Spring 2009)
- Research and Doctoral Committee (Fall 2008, Spring 2009)
- Ad Hoc Personnel Committee (Summer 2008)
- Personnel Committee (Fall 2004, Spring 2005, Fall 2006, Fall 2007, Spring 2008)
- Information Resources and Technology Committee (Fall 2005, Spring 2006)
- Introduced and assessed the database placement exam (Spring 2004-2006)
- Ad-hoc Curriculum Development Committee to review undergraduate (INLS050) and graduate (INLS102) introductory technology courses (Spring 2005)

#### **PROFESSIONAL AFFILIATIONS**

---

- American Society for Information Science and Technology (ASIS&T) 2002–3, since 2006.
- American Artificial Intelligence Association (AAAI) 1996-2009
- American Medical Informatics Association (AMIA) 1999-2009
- Association for Computing Machinery (ACM) since 2001
- Association for Computational Linguistics (ACL) since 2006
- Associate Member of the Lineberger Comprehensive Cancer Center (LCCC) 2004-2009
- Institute of Electrical and Electronics Engineers Inc (IEEE) 2001–2003