

Dr. CHIRAG SHAH

Work Address

School of Communication & Information (SC&I)
Rutgers, The State University of New Jersey
4 Huntington St, New Brunswick, NJ 08901
Email: chirags@rutgers.edu

Home Address

49 Mary St
Bordentown NJ 08505
Phone: (732) 915-4090
Homepage: <http://comminfo.rutgers.edu/~chirags>

Education and Work

2006-2010	PhD in School of Information & Library Science (SILS) at UNC Chapel Hill. Dissertation on Collaborative Information Seeking under Prof. Gary Marchionini.
2003-2006	MS in Computer Science at UMass Amherst. Research Assistant under Prof. W. Bruce Croft at Center for Intelligent Information Retrieval (CIIR).
2002-2003	Research Assistant at Center For Intelligent Internet Research (CFIIR), Dept. of CSE, Indian Institute of Technology (IIT) Bombay, India.
2000-2002	MTech in Computer Science at Indian Institute of Technology (IIT) Madras, India.
1996-2000	BE in Computer Engineering at DDIT Nadiad, India.

Professional Experience

Since Fall 2009	Director of Web Technologies, Texifter (http://www.texifter.com) - a startup for developing tools and support for extracting and analyzing information from online sources.
Summer 2007	Internship at Fuji Xerox Palo Alto Lab (FXPAL) at California. Work on TRECvid as well as Collaborative Information Retrieval.
Summer 2006	Internship at National Institute of Informatics (NII) at Tokyo, Japan. Work on Story Link Detection (SLD) and TREC's Blog track.
Summer 2004	Internship at Honda Research Institute at Mountain View, CA. Work on automated plan building for ASIMO robot.
2003-2006	At UMass Amherst. Various issues on IR, data mining, and social networking with Prof. W. Bruce Croft and Prof. James Allan of CIIR and Prof. David Jensen of Knowledge Discovery Lab.
Jan. 2002- July 2003	At IIT Bombay under Prof. Pushpak Bhattacharyya: Tata Consultancy Services (TCS) sponsored project on enhancing text representation for improving information retrieval. Also worked on TREC's HARD track.
Apr. 2002- Dec. 2002	At IIT Bombay under Prof. Pushpak Bhattacharyya: Ministry of Information Technology, India sponsored project of text clustering for information retrieval.
June 2001- Jan. 2002	Work on using neural network models including Self-Organizing Map (SOM) and constraint satisfaction model for text classification that could help in improving information retrieval.
July 2000- Apr. 2001	Intel sponsored project on spoken documents retrieval for Indian languages at IIT Madras under Prof. Yegnanarayana.
Dec. 1999- Apr. 2000	Building a translation system for Indian languages using United Nations University (UNU) proposed standard called Universal Networking Language (UNL) at IIT Bombay under Prof. Pushpak Bhattacharyya.

Associations and Professional Activities/Services

Services to the research community

- Chair for the Information Retrieval track at HLT-NAACL 2009 Student Research Workshop.
- Reviewer for IP&M Special Issue on Collaborative Information Seeking (2009).
- Reviewer for a book on Collaborative Information Behavior, edited by Jonathan Foster (2009).
- Program committee member for SIGIR 2009 and SIGIR 2010 conferences.
- Student volunteer for SIGIR 2009 conference.
- Program committee member for JITP 2009 and JITP 2011 conferences.
- Chair for a session on "Archiving and Web Tools for DL" at JCDL 2008 conference.
- Reviewer for JCDL 2008 conference.
- Reviewer for CHI 2008, CHI 2009, and CHI 2010 conferences.
- Reviewer for CIKM 2007 conference.

Services to SILS at UNC Chapel Hill

- Faculty search committee member (2008-2009).
- Full time doctoral students representative (2008-2009).

Professional associations

- ASIST student member since 2008.
- ACM student member since 2000.
- IEEE student member during 1997-2002.
- Microsoft Certified Professional since 1999.
- Chairman of IEEE Student Branch at Nadiad, India during 1998-1999.

Selected Awards and Honors

Nov. 2009	SIG USE award for interdisciplinary conference travel (\$200).
April 2009	Student travel grant from Microsoft Research to attend CHI 2009 at Boston.
March 2009	3rd prize (\$500) for the paper at "YouTube and the 2008 Elections Cycle" conference.
Sept. 2008	NSF award for \$448,071 beginning Sept 1, 2008 and running for 3 years titled "Result Space Support for Personal and Group Information Seeking Over Time" (proposal written with my advisor Gary Marchionini).
July 2008	Best paper award (\$1000) at SIGIR 2008.
July 2008	Student travel grant from ACM and Google to attend SIGIR 2008 at Singapore.
June 2008	Student travel grant from ACM to attend JCDL 2008 at Pittsburgh.
June 2007	Student travel grant from ACM to attend SIGIR 2007 at Amsterdam, The Netherlands.
2006-2007	Lester Asheim Merit-based Fellowship at SILS, UNC Chapel Hill.
Since Fall 2006	Doctoral research assistantship. UNC Chapel Hill.
2003-2006	Merit-based research assistantship for Masters. UMass Amherst.
June 2004	Student travel grant from ACM to attend SIGIR 2004 at Sheffield, UK.
May 2002	3rd prize in IEEE's regional Student Paper Contest based on paper competition in IEEE Region 10, which includes a number of Asian countries. Name inducted in IEEE Student Paper Contest Hall of Fame.
May 2001	Scholarships by IIT Madras Alumni Association at USA and IEEE for attending ICASSP 2001 at Salt Lake City, Utah, USA.
Sept. 1999	Best Student Paper award in IEEE's regional paper contest. Gujarat, India.
July 1997	Best Student Paper award at Birla Institute of Technology & Science (BITS) Pilani, India.
Jan. 1997	2nd prize at national level in student paper contest. Chennai, India.
Aug. 1995	Prime Minister Award for outstanding academic achievement. New Delhi, India.
July 1994	2nd in State Board Exam, Gujarat, India.
March 1993	<i>Golden Boy of School</i> award. Madresa, Navsari, India.

Publications

Book Chapters (refereed)

- **Shah, Chirag**. Designs for Systems to Support Collaborative Information Behavior (in press). To be published in Foster, Jonathan (Ed.) *Collaborative Information Behavior*. IGI Golbal, Hershey PA.

Journal Articles (refereed)

- **Shah, Chirag**, Pickens, Jeremy, and Golovchinsky, Gene (in press). Role-Based Results Redistribution for Collaborative Information Retrieval. To be published in *Journal of Information Processing and Management (IP&M)*.
- **Shah, Chirag** (2010). Supporting Research Data Collection from YouTube with TubeKit. *Journal of Information Technology & Politics*, 7(2-3), 226-240.
- Kelly, Diane, Fu, Xin, and **Shah, Chirag** (2010). Effects of Position and Number of Relevant Documents Retrieved on Users' Evaluations of System Performance. *ACM Transactions on Information Science (TOIS)*, 28(2), 9:1-29.
- Sayre, Ben, Bode, Leticia, Shah, Dhavan, Wilcox, Dave, and **Shah, Chirag** (in press). Message Flows in a Digital Age: Tracking Attention to Prop. 8 in Social Media and Conventional News. To be published in *Journal of Policy & Internet*.
- Thorson, Kjerstin, Borah, Porismita, Ekdale, Brian, Namkoong, Kang, and **Shah, Chirag** (in press). YouTube and Proposition 8: a case study in video activism. To be published in *Information, Communication, and Society (iCS)*.
- **Shah, Chirag**, Oh, Sanghee, and Oh, Jung Sun (2009). Research agenda for Social Q&A. *Library & Information Science Research (LISR)*, 31(4), 205-209.
- **Shah, Chirag** (2009). ContextMiner: Supporting the Mining of Contextual Information for Ephemeral Digital Video Preservation. *International Journal of Digital Curation (IJDC)*, 4(2).
- **Shah, Chirag** (2009). Mining Contextual Information for Ephemeral Digital Video Preservation. *International Journal of Digital Curation (IJDC)*, 4(1).
- **Shah, Chirag** and Eguchi, Koji (2008). Improving document representation for Story Link Detection by Modeling Term Topicality. *IPSJ Transactions on Databases*, 1 (3), 11-19.
- **Shah, Chirag**, Oh, Jung Sun, and Oh, Sanghee (2008). Exploring characteristics and effects of user participation in online social Q&A sites. *First Monday*, 13(9).

Journal/Magazine Articles (invited)

- **Shah, Chirag** (2009). What Do You Look Like on YouTube? *Politics Magazine*. July 2009.
- **Shah, Chirag** (2009). ContextMiner: Explore Globally, Aggregate Locally. *IEEE Computer*. March 2009.
- **Shah, Chirag** (2009). TubeKit - A Query-based YouTube Crawling Toolkit. *Bulletin of IEEE Technical Committee on Digital Libraries*, 5(1), Spring 2009.
- **Shah, Chirag**, and Marchionini, Gary (2008). Capturing Relevant Information for Digital Curation. *Bulletin of IEEE Technical Committee on Digital Libraries*, 4(1), Spring 2008.
- **Shah, Chirag**, and Marchionini, Gary (2008). ContextMiner: A Tool for Digital Library Curators. *Bulletin of IEEE Technical Committee on Digital Libraries*, 4(1), Spring 2008.
- **Shah, Chirag** (2008). Understanding System Implementation and User Behavior in a Collaborative Information Seeking Environment. *Bulletin of IEEE Technical Committee on Digital Libraries*, 4(2), Fall 2008.

Conference Papers (refereed)

- **Shah, Chirag**, and Pomerantz, Jeffrey (2010). Evaluating and Predicting Answer Quality in Community QA. *Proceedings of ACM SIGIR 2010 Conference*. Geneva, Switzerland: July 19-23, 2010.
- Sayre, Ben, Bode, Leticia, Wilcox, Dave, Shah, Dhavan V., and **Shah, Chirag** (2009). Message flows in a digital age: tracking attention to Prop. 8 in social media and conventional news. *Proceedings of Association of Internet Researchers (AOIR) Conference*. Milwaukee, Wisconsin: October 8-10, 2009.
- Edgerly, Stephanie, Vraga, Emily, Fung, Timothy, Moon, Tae Joon, Yoo, Woo Hyun, Veenstra, Aaron, and **Shah, Chirag** (2009). YouTube as a public sphere: the Proposition 8 debate. *Proceedings of Association of Internet Researchers (AOIR) Conference*. Milwaukee, Wisconsin: October 8-10, 2009.
- Lee, Christopher A., Marciano, Richard, Hou, Chien-yi, **Shah, Chirag** (2009). Mainstreaming Preservation through Slicing and Dicing of Digital Repositories: Investigating Alternative Service and Resource Options for ContextMiner Using Data Grid Technology. *Proceedings of iPRES 2009*. San Francisco: October 5-6, 2009.
- Marchionini, Gary, **Shah, Chirag**, Lee, Christopher A., Capra, Robert (2009). Query Parameters for Harvesting Digital Video and Associated Contextual Information. *Proceedings of IEEE ACM Joint Conference on Digital Libraries (JCDL) 2009*. Austin, Texas: June 15-19, 2009.
- **Shah, Chirag** (2009). Supporting Research Data Collection from YouTube with TubeKit. *Proceedings of YouTube and 2008 Election Cycle in the United States*. Amherst, MA: April 16-17, 2009. **(3rd prize)**
- **Shah, Chirag**, Marchionini, Gary, and Kelly, Diane (2009). Learning Design Principles for a Collaborative Information Seeking System. *Proceedings of ACM SIGCHI Conference 2009*. Boston, MA: April 4-9, 2009.
- **Shah, Chirag**, and Marchionini, Gary. Hunting for Hip, Hipsters, and Happenings on YouTube (2008). *Proceedings of ASIST 2008 Annual Meeting*. Columbus, OH: October 24-29, 2008.
- Oh, Sanghee, Oh, Jung Sun, and **Shah, Chirag** (2008). The Use of Information Sources by the Internet Users in Answering Questions. *Proceedings of ASIST 2008 Annual Meeting*. Columbus, OH: October 24-29, 2008.
- Pickens, Jeremy, Golovchinsky, Gene, **Shah, Chirag**, Qvarfordt, Pernilla, and Back, Maribeth (2008). Algorithmic Mediation for Collaborative Exploratory Search. *Proceedings of ACM SIGIR 2008 Conference*. Singapore: July 20-24, 2008. **(Best paper award)**
- Capra, Robert, Lee, Christopher A., Marchionini, Gary, Russell, Terrell, **Shah, Chirag**, and Stutzman, Fred (2008). Selection and Context Scoping for Digital Video Collections: An Investigation of YouTube and Blogs. *Proceedings of IEEE ACM Joint Conference on Digital Libraries (JCDL) 2008*. Pittsburgh, PA: June 16-20, 2008.
- Kelly, Diane, **Shah, Chirag**, Sugimoto, Cassidy R., Bailey, Earl W., Clemens, Rachel A., Irvine, Ann K., Johnson, Nicholas A., Ke, Weimao, Oh, Sanghee, Poljakova, Anezka, Rodriguez, Marcos A., van Noord, Megan G., and Zhang, Yan (2008). The effects of performance feedback on users' evaluations of an interactive IR system. *Proceedings of the 2nd Symposium on Information Interaction in Context (IliX)*. London, UK: Oct. 14-17, 2008.
- **Shah, Chirag** and Eguchi, Koji (2007). Use of Topicality and Information Measures to Improve Document Representation for Story Link Detection. *Proceedings of European Conference on Information Retrieval (ECIR) 2007*. Rome, Italy: April 2-5, 2007.
- **Shah, Chirag** and Croft, W. Bruce (2004). Evaluating High Accuracy Retrieval Techniques. *Proceedings of ACM SIGIR 2004*. Sheffield, UK: July 25-29, 2004.
- **Shah, Chirag** and Bhattacharyya, Pushpak (2003). Improving Document Vectors Representation using Semantic Links and Attributes. *Proceedings of International Conference on Natural Language Processing (ICON) 2003*. Mysore, India: December 2003.
- **Shah, Chirag**, Chaudhary, Bhoopesh, and Bhattacharyya, Pushpak (2002). Constructing Better Document Vectors Using Universal Networking Language (UNL). *Proceedings of International Conference on Knowledge Based Systems (KBCS) 2002*. Mumbai, India: December 2002.
- **Shah, Chirag** and Bhattacharyya, Pushpak (2002). A Study for Evaluating the Importance of Various Parts of Speech (POS) for Information Retrieval (IR). *Proceedings of International Conference on Universal Knowledge and Language (ICUKL) 2002*. Goa, India: November 2002.

Conference Posters and Demos (refereed)

- Oh, Jung Sun, **Shah, Chirag**, and Oh, Sanghee (2009). User Participation Patterns Over Time in Yahoo! Answers. *Poster in Proceedings of ASIST 2009 Annual Meeting*. Vancouver, BC, Canada: November 6-11, 2009.
- Lee, Christopher A., Marciano, Richard, Hou, Chien-yi, and **Shah, Chirag** (2009). From harvesting to cultivating: transformation of a Web collecting system into a robust curation environment. *Poster in Proceedings of IEEE ACM Joint Conference on Digital Libraries (JCDL) 2009*. Austin, Texas: June 15-19, 2009.
- **Shah, Chirag** (2009). ContextMiner - explore globally, aggregate locally. *Demo in Proceedings of IEEE ACM Joint Conference on Digital Libraries (JCDL) 2009*. Austin, Texas: June 15-19, 2009.
- **Shah, Chirag** (2009). ContextMiner - Collect Different. Demo at DigCCurr 2009. Chapel Hill, NC: April 1-3, 2009.
- **Shah, Chirag** (2009). TubeKit - A Query-based YouTube Crawling Toolkit. *Demo in Proceedings of IEEE ACM Joint Conference on Digital Libraries (JCDL) 2008*. Pittsburgh, PA: June 16-20, 2008.
- Marchionini, Gary, Tibbo, Helen, **Shah, Chirag**, and Lee, Christopher A (2007). Telling the Whole Story: Selecting and Collecting Web-Based Videos for Archival Collections. *Poster in Proceedings of Digital Curation Conference (DCC)*. Washington, DC, USA: December 2007.
- **Shah, Chirag**, Kelly, Diane, and Fu, Xin (2007). Making Mind and Machine Meet - A Study of Combining Cognitive and Algorithmic Relevance Feedback. *Poster in Proceedings of ACM SIGIR 2007 Conference*. Amsterdam, The Netherlands: July 23-27, 2007.
- Fu, Xin, Kelly, Diane, and **Shah, Chirag** (2007). Using Collaborative Queries to Improve Retrieval for Difficult Topics. *Poster in Proceedings of ACM SIGIR 2007 Conference*. Amsterdam, The Netherlands: July 23-27, 2007.
- **Shah, Chirag** and Marchionini, Gary (2007). DiscoverInfo: A Tool for Discovering Information with Relevance and Novelty. *Demo in Proceedings of ACM SIGIR 2007 Conference*. Amsterdam, The Netherlands: July 23-27, 2007.
- **Shah, Chirag** and Marchionini, Gary (2007). Capturing Relevant Information for Digital Curation. *Poster in Proceedings of IEEE ACM Joint Conference on Digital Libraries (JCDL) 2007*. Vancouver, BC, Canada. June 18-23, 2007.
- **Shah, Chirag** and Marchionini, Gary (2007). ContextMiner: A Tool for Digital Curator. *Demo in Proceedings of IEEE ACM Joint Conference on Digital Libraries (JCDL) 2007*. Vancouver, BC, Canada: June 18-23, 2007.
- **Shah, Chirag**, W. Bruce Croft and David Jensen (2006). Representing Documents with Named Entities for Story Link Detection (SLD). *Poster in Proceedings of International Conference on Information and Knowledge Management (CIKM) 2006*. Arlington VA, USA: October 2006.
- **Shah, Chirag** and Khan, A. Nayeemulla (2001). Spoken Document Retrieval (SDR) for Broadcast News in Indian Languages. *Poster in Proceedings of IEEE's International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2001*. Salt Lake City UT, USA: May 2001.

Workshop Papers (refereed)

- **Shah, Chirag** and Marchionini, Gary (2009). Query reuse in exploratory search tasks. *Poster at Workshop on Human-Computer Interaction and Information Retrieval (HCIR) 2009*. Washington, DC: October 23, 2009.
- **Shah, Chirag** (2009). Lessons and Challenges for Collaborative Information Seeking (CIS) Systems Developers. *Proceedings of Collaborative Information Behavior Workshop at GROUP 2009*. Sanibel Island, Florida: May 10, 2009.
- Marchionini, Gary, Capra, Robert, and **Shah, Chirag** (2008). Focus on results: personal and group information seeking over time. *Poster at Workshop on Human-Computer Interaction and Information Retrieval (HCIR) 2008*. Seattle, WA: October 23, 2008.
- **Shah, Chirag** (2008). Toward Collaborative Information Seeking (CIS). *Proceedings of Collaborative Exploratory Search workshop at JCDL 2008*. Pittsburgh, PA: June 20, 2008.

- **Shah, Chirag** and Marchionini, Gary (2007). Preserving 2008 US Presidential Election Videos. *Proceedings of International Web Archiving Workshop (IWAWS)*. Vancouver, BC, Canada: June 23, 2007.
- **Shah, Chirag** and Gupta, Rakesh (2005). Building Plans for Household Tasks from Distributed Knowledge. *Poster in Proceedings of IJCAI 2005 Workshop - Modeling Natural Action Selection (MNAS)*. Edinburgh, UK: July 2005.

Doctoral Colloquia (refereed)

- **Shah, Chirag** (2008). Understanding System Implementation and User Behavior in a Collaborative Information Seeking Environment. *Doctoral Consortium at SIGIR 2008*. Singapore: July 20, 2008.
- **Shah, Chirag** (2008). *Understanding System Implementation and User Behavior in a Collaborative Information Seeking Environment. Doctoral Consortium at JCDL 2008*. Pittsburgh, PA: June 16, 2008.

Panels (refereed)

- Radford, Marie L., Connaway, Lynn Silipigni, Pomerantz, Jeffrey, **Shah, Chirag**, Oh, Jung Sun, Oh, Sanghee, and Gazan, Rich (2009). Opportunities, Threats, & Theoretical Approaches: Research in Traditional and Social Virtual Reference Quality. *Peer-reviewed panel at ASIST 2009 Annual Meeting*. Vancouver, BC, Canada: November 6-11, 2009.
- Shachaf, Prina, Rosenbaum, Howard, Abels, Eileen, Radford, Marie, Connaway, Lynn, Gazan, Rich and **Shah, Chirag** (2009). Social Reference and Digital Reference: Online question answering practices in two diverse communities. *Peer-reviewed panel at ASIST 2009 Annual Meeting*. Vancouver, BC, Canada: November 6-11, 2009.

Technical Reports (unrefereed)

- Eguchi, Koji and **Shah, Chirag** (2006). Opinion Retrieval Experiments using Generative Models (Notebook Version). *TREC 2006*. NIST, Gaithersburg: November 2006.
- **Shah, Chirag** (2006). Story Link Detection (SLD) Using Relational Knowledge Discovery. *MS Thesis*. UMass Amherst: April 2006.
- Nallapati, Ramesh, and **Shah, Chirag** (2005). Evaluating the Quality of Query Refinement Suggestions in Information Retrieval. *Technical report #IR-521*. CIIR, UMass Amherst: April 2005.
- Connell, Margaret, Feng, Ao, Kumaran, Giridhar, Raghavan, Hema, **Shah, Chirag**, and Allan, James (2004). UMass at TDT 2004 (Notebook Version). *Topic Detection and Tracking (TDT) 2004*. NIST, Gaithersburg: December 2004.
- Ramakrishnan, Ganesh, Bellare, Kedar, **Shah, Chirag**, and Paranjpe, Deepa (2003). Generic Text Summarization Using WordNet for Novelty and HARD (Notebook Version). *TREC 2003*. NIST, Gaithersburg: December 2003.
- **Shah, Chirag** (2002). Text Documents Classification using Self-Organizing Map (SOM) for Information Retrieval. *MTech Thesis*. IIT Madras, India: January 2002.
- **Shah, Chirag** (2002). Automatic Organization of Text Documents in Categories Using Self-Organizing Map (SOM). *IEEE's Regional Student Paper Contest. (3rd prize)*
- **Shah, Chirag** (2000). EnConversion System from English to Universal Networking Language (UNL). *BE Project Dissertation*. IIT Bombay, India: April 2000.

Grants

- **NSF award for \$448,071** (IIS #0812363) beginning Sept 1, 2008 and running for 3 years titled "Result Space Support for Personal and Group Information Seeking Over Time" (proposal written with my advisor Gary Marchionini).

Patents

- Gupta, Rakesh, and **Shah, Chirag**. *Building Plans For Household Tasks From Distributed Knowledge*. U.S. Patent Application No.: 11/296,020. Issued on May 6, 2008.

Talks and Presentations

Contributed Talks

- *Cashing on the Wisdoms of Crowds for Question Answering*. ASIST 2009 Panel on Social Reference and Digital Reference: Online Question Answering Practices in Two Diverse Communities (SIGs SI, BWP, USE). Vancouver, BC, Canada. November 10, 2009.
- *The Same, Yet Different Comparing Studies of "Traditional" Digital Reference and Social Q&A*. Presentation with Sanghee Oh on ASIST 2009 Panel on Opportunities, Threats, & Theoretical Approaches: Research in Traditional and Social Virtual Reference Quality (SIG HCI). Vancouver, BC, Canada. November 9, 2009.
- *Supporting Research Data Collection from YouTube with TubeKit*. At YouTube and the 2008 Election Cycle in the United States Conference. April 16, 2009. Amherst, MA. **(3rd prize winner)**
- *Hunting for Hip, Hipsters, and Happenings on YouTube*. ASIST 2008 presentation. Columbus, Ohio. October 27, 2008.
- *Understanding System Implementation and User Behavior in a Collaborative Information Seeking Environment*. SIGIR 2008 Doctoral Consortium. Singapore. July 20, 2008.
- *Understanding System Implementation and User Behavior in a Collaborative Information Seeking Environment*. JCDL 2008 Doctoral Consortium. Pittsburgh, PA. June 16, 2008.
- *From Searching to Exploring, From Finding to Understanding, and From Seeing to Visualizing the Information*. University Research Day Poster **(2nd prize winner)**. UNC Chapel Hill. March 4, 2008.
- *Preserving 2008 US Presidential Election Videos*. International Web Archiving Workshop (IWA) 2007. Vancouver, BC, Canada. June 23, 2007.
- *Use of Topicality and Information Measures to Improve Document Representation for Story Link Detection*. European Conference on Information Retrieval (ECIR) 2007. Rome, Italy: April 3, 2007.
- *Topic Linking With Blogs - TREC 2006 Blog Track: Open Task*. TREC 2006 at NIST. November 14, 2006.
- *Evaluating High Accuracy Retrieval Techniques*. SIGIR 2004. Sheffield, UK. July 2004.

Invited Talks

- *Beyond Google: A Quest to Understand Information Seeking Across Time and People*. CRADLE Talk. UNC Chapel Hill. December 4, 2009.
- *Awareness in Collaborative Information Seeking*. SILS Research Symposium. October 30, 2009.
- *Toward a model for Collaborative Information Seeking, Synthesis, and Sense-making: A Work in Progress*. University of Maryland. October 13, 2009.
- *Toward a model for Collaborative Information Seeking, Synthesis, and Sense-making: A Work in Progress*. Pennsylvania State University. October 12, 2009.
- *System-mediated and user-mediated Collaborative Information Seeking (CIS)*. University of Sheffield, UK. August 21, 2009.
- *System-mediated and user-mediated Collaborative Information Seeking (CIS)*. Microsoft Research Cambridge, UK. August 19, 2009.
- *A Story of Context-Capturing and Cross-Country Coding*. University of Wisconsin at Madison. August 12, 2009.
- *Republic of Web: A web of, by, and for the people*. CRADLE talk. UNC Chapel Hill. November 14, 2008.
- *Collaborative Information Seeking: Counting on the Wisdom of Crowd*. INLS 200 guest lecture. UNC Chapel Hill. October 23, 2008.

- *Collaborative Information Seeking: Cashing on the Wisdom of Crowd*. School of Computer Sciences, Universiti Sains Malaysia. Penang, Malaysia. July 30, 2008.
- *Collaborative Information Seeking: An Art of Making 1+1>2*. INLS 490 guest lecture. UNC Chapel Hill. April 21, 2008.
- *Collaborative Information Retrieval: All for One and One for All*. INLS 200 guest lecture. UNC Chapel Hill. March 24, 2008.
- *We, the People on YouTube: Understanding Democracy of, by, and for the Online Media Sharing*. University Research Day Talk. UNC Chapel Hill. March 4, 2008.
- *Hunting for Hip, Hipsters, and Happenings on YouTube - A ContextMiner Story*. CRADLE Talk. UNC Chapel Hill. September 21, 2007.
- *Finding Connectors, Mavens, and Salesmen on YouTube*. Yahoo! Research Berkeley, CA. August 10, 2007.
- *Collaborative Exploratory Search (CES) - A Beginning*. Summer Interns Presentation at FXPAL, CA. August 7, 2007.
- *Two is a company: A step toward building a Collaborative Exploratory Search (CES) environment*. WHIP Talk at FXPAL, CA. June 6, 2007.
- *Mining Contextual Information for Ephemeral Digital Video Preservation: A Beginning*. Research Issues and Questions Seminar. UNC Chapel Hill. April 24, 2007.
- *Digital Museums and User Interaction*. With Amy Morgan and Stephanie Adamson in Human Information Interaction class. UNC Chapel Hill. April 24, 2007
- *Investigating the Effects of Faceted Interfaces - Research Proposal*. Human Computer Interaction (HCI) Seminar. UNC Chapel Hill. December 14, 2006.
- *Relevance Revisited*. Research Issues & Questions Seminar. UNC Chapel Hill. November 21, 2006.
- *Use of Topicality and Information Measures to Improve Representation for Story Link Detection (SLD)*. National Institute of Informatics (NII), Tokyo, Japan. August 16, 2006.
- *Using Named Entities Based Representation for Story Link Detection (SLD)*. UMass Amherst. May 5, 2006.
- *Story Link Detection (SLD) using Named Entity Graphs*. NorthEastern University, Boston. January 6, 2006.
- *A formal study of Terminological Feedback*. UMass Amherst. January 13, 2005.
- *Model Selection in Unsupervised Tracking*. UMass Amherst. January 6, 2005.
- *Building Plans for Household Tasks from Distributed Knowledge*. Honda Research Institute (HRI), Mountain View, CA. August 13, 2004.

Tools Developed

(All open-source, and available to the research community for free under creative commons licenses)

- *Coagmento* - a collaborative information seeking, synthesis, and sense-making system (<http://www.coagmento.org>)
This tool has been tested with cognitive walkthroughs, pilot runs, and lab studies. It is now deployed as an online service and being used by several groups, including the students of a graduate HCI class.
- *ContextMiner* - a tool for curators and researchers for collecting data as well as contextual information by doing automated crawls from various sources, including YouTube, Flickr, Twitter, blogs, and the open web (<http://www.contextminer.org>)
ContextMiner has been in use by more than 300 individuals and organizations around the world, helping them in a variety of research projects, including masters and PhD theses.
- *InfoExtractor* - A tool for extracting structured information from social web. (<http://www.infoextractor.org>)
- *TubeKit* - a toolkit for building crawlers for YouTube (<http://www.tubekit.org>)
In about two years time, TubeKit has been downloaded nearly 500 times, and has served as a critical tool for data collection and analysis pertaining to research on online video participation.

- *XCreator* - a web-based toolkit for creating browser plug-ins that helps in collecting user's browsing data. (<http://www.xcreator.org>)
This toolkit can be very useful for creating tools for doing user studies that rely on monitoring user's online activities.
- *DiscoverInfo* - a collection visualization interface for 1898 North Carolina election documents (<http://idl.ils.unc.edu/~chirag/DiscoverInfo/>)
- *DIToolkit* - a toolkit to crawl, index, and visualize a collection (<http://idl.ils.unc.edu/~chirag/DIToolkit/>)

Data Collections Developed

- YouTube Collections - consist of nearly 100,000 YouTube videos, collected for different topics including presidential elections and cancer research, over two years with a number of attributes recorded almost every day. The collection includes actual videos, user comments, and user profiles.
- Yahoo! Answers - consists of more than 2,000,000 questions in 25 different categories, along with nearly 15,000,000 answers and close to 1,000,000 user profiles.
- Links Data - includes inlinks and outlinks data crawled from the web about a number of talk-show videos. There are nearly 1,000,000 links extracted from such data, creating a huge graph of connections.

Teaching Experience

Fall 2009	Taught self-designed graduate level online course on "Information Retrieval Systems Design and Implementation" (INLS490-154W). Created video podcasts, an online quiz system, and other material for web-based teaching. (http://www.inforetrieval.org)
Fall 2008 & Spring 2009	Taught a new self-designed graduate level course on "Information Retrieval Systems Design and Implementation" (INLS490-154)
Spring 2008	Guest lecture in INLS 490 class.
Spring 2008	Guest lecture in INLS 200 class.
Spring 2008	Took EDUC 757: College Teaching Seminar for a formal training in college teaching
Spring 2001	Teaching assistant for Computer Organization (graduate level) course under Prof. S. Raman at IIT Madras, India.
Fall 2000	Teaching assistant at Ushamartine Academy of Communication Technology (UACT), India for training professionals from various commercial firms such as Wipro, TCS, Infosys, etc.
Summer 2000	Course for professionals on building e-commerce applications using various technologies such as ASP, Windows NT, IIS, etc. at Pro-Link Academy, Navsari, India.
Summer 1999	Seminars to students and professionals on web technologies at National Institute of Information Technology (NIIT), Navsari, India.
Summer 1998	Taught Programming with Java to professionals at Paradise Computers, Navsari. Conducted a course on providing network solutions with Microsoft BackOffice at Databank, Navsari, India.
Summer 1997	Conducted a course on website creation at Databank, Navsari, India for professionals.

Dissertation Committee

1. Prof. Gary Marchionini (Chair). School of Information and Library Science, UNC Chapel Hill.
2. Prof. Diane Kelly. School of Information and Library Science, UNC Chapel Hill.
3. Prof. Barbara Wildemuth. School of Information and Library Science, UNC Chapel Hill.
4. Prof. Deborah Barreau. School of Information and Library Science, UNC Chapel Hill.
5. Dr. Susan Dumais. Microsoft Research.

Collaborators

The following is a list of some of my past and present collaborators, giving an idea of the kind of projects I have been involved in (other than my dissertation) and the kind of organizations and disciplines I have been working with.

Current collaborators

- | | |
|-------------------|---|
| Since Summer 2006 | Fuji Xerox Palo Alto Lab (FXPAL) California on research related to algorithmically mediated collaborative exploratory search. |
| Since Spring 2008 | UMass Amherst (Political Science) for studying online social participation around a number of political issues. |
| Since Fall 2009 | UNC Greensboro (Media Studies) for studying dissemination and aggregation of information around issues such as environment and food. |
| Since Fall 2009 | UNC Chapel Hill (SILS), University of Pittsburgh (iSchool) and University of Texas at Austin (iSchool) for studying health-related information seeking in online social Q&A services, such as Yahoo! Answers and Internet Public Library (IPL). |
| Since Summer 2009 | University of Wisconsin Madison (Journalism and Mass Communication, Political Science) for studying various political (e.g., proposition 8) and health (e.g., cancer survival) issues in online social networking sites. |
| Since Fall 2008 | UNC Chapel Hill (SILS, Journalism), WRAL TV Station, The Carrboro Citizen Newspaper for further development and deployment of ContextMiner. |

Past collaborators

- | | |
|---------------------------|---|
| Fall 2006-
Summer 2009 | Library of Congress and Internet Archives for preserving presidential elections 2008 videos and contextual information. |
| Summer 2008-
Fall 2008 | University of Pittsburgh (Social and Urban Research) for understanding social dynamics on information seeking and sharing sites, such as YouTube and blogs. |
| Summer 2006-
Fall 2008 | National Institute of Informatics (NII), Tokyo and University of Kobe, Japan for topic tracking, event detections, and sentiment analysis in blogs. |