

Prasenjit Mitra

9271 West Calvin Road
Hartstown, PA 16131
Phone: +974 5565 5751
Email: pmitra3@gmail.com

Education

- Aug, 2004 **Stanford University, Stanford, CA**
Doctor of Philosophy, Electrical Engineering, GPA 3.95/4.00
- Dec, 1994 **The University of Texas at Austin, Austin, TX**
Master of Science, Computer Science, GPA 4.00/4.00
- May, 1993 **Indian Institute of Technology, Kharagpur, India**
Bachelor of Technology (Honors), Computer Science and Engineering

Work Experience

- 2014-Present
2014-Present
2010-2014
2003-2010
2003-Present
- Principal Scientist, Qatar Computing Research Institute**
Professor, College of Information Sciences and Technology, Penn State University
Associate Professor, College of Information Sciences and Technology,
Assistant Professor, College of Information Sciences and Technology,
Graduate Faculty, Department of Computer Science and Engineering,
Department of Industrial and Manufacturing Engineering,
The Pennsylvania State University, University Park, PA
- 2008-2013 **(Consulting) Chief Scientist, Global IDs, New York, NY**
Consulting on the research and development of the company's data integration products and intellectual property. Providing a vision for future directions.
- 2002 **Senior Software Engineer, DBWizards, Menlo Park, CA**
Consulted on the research and development of schema matching and ontology matching algorithms.
- 2001-2002 **Senior Software Engineer, Narus, Incorporated, Palo Alto, CA**
Designed and implemented a join algorithm for generating reports from a database of streaming network data.
- 1999-2003 **Research Assistant, Department of Computer Science, Stanford University, Stanford, CA**
Conducted research on data warehouses, data integration and interoperation, and semantic technologies.
- 1997-2000 **Senior Member of Technical Staff, Oracle Corporation, Redwood Shores, CA**

	Designed and implemented the object cache, inheritance and schema evolution in Oracle 8i.
1995 – 97	Member of Technical Staff, Oracle Corporation, Redwood Shores, CA Designed and implemented the connection load balancer and the Oracle Trace Logging utility for the Oracle Parallel Server.
1994	Research Assistant, The University of Texas at Austin, Austin, TX Developed efficient linear algebra algorithms for use in parallel numerical computation.
1993-94	Teaching Assistant, The University of Texas at Austin, Austin, TX Assisted in teaching courses: Spring, 1994 CS 352: Computer System Architecture. Fall, 1993 CS 310: Computer Organization and Programming
1992	Instructor, Mahila Samity Polytechnic Institute, Oil and Natural Gas Corporation, Dehradun, UP, India Taught a course: Computer Systems Analysis and Design.
1992	Intern, Institute for Management Development, Oil and Natural Gas Corporation, Dehradun, UP Developed database software on IBM/CICS VMS.
1991	Intern, Natl. Ctr. of Knowledge Based Computing, Indian Statistical Institute, Kolkata, WB Developed an image-processing algorithm for the automated testing of PCBs.
	Awards and, Honors
2009-2014 2008	NSF CAREER Award , National Science Foundation, U.S.A. IEEE VAST 2008 Challenge , Grand Challenge (Data Integration), led winning team, representing the Pennsylvania State University.
1989-1993	National Talent Search Award and Scholarship , National Center for Educational Research and Training, India
1994	ACM Programming Contest , UT Austin: member of winning team, represented the University of Texas at Austin.

Professional Memberships

Member, Association for Computing Machinery (ACM).

Member, Institute of Electrical and Electronic Engineers (IEEE).

Member, Association for the Advancement of Artificial Intelligence (AAAI).

Dissertations Supervised

M.S.: S. Ray (Microsoft), V. Kalsi (Intel), P. Ayyagari (Bloomberg), A.R. Jaiswal (PSU) Ziming Zhuang (Yahoo! Labs);

Ph.D.: S. Debnath (Microsoft), Q. Tan (Google), B. Sun (eBay Research), Z. Zhuang (Yahoo! Labs), Y. Liu (Assistant Professor at Korean Advanced Institute of Science and Technology), P. Biyani (Yahoo Research), S. Bhatia (IBM Almaden), S. Kataria (Xerox Research), D. Yuan (Google), X. Zhang (Twitter), A. Jaiswal (Reunify);

Postdoctoral Scholar: Q. Zhao (Telephonica Research Labs), Q. He (LinkedIn), C. Caragea (Assistant Professor at University of North Texas).

Publications

Biyani, P., Bhatia, S., Caragea, C., and, Mitra, P. (2014) "Using non-lexical features for identifying factual and opinionative threads in online forums." *Knowledge-Based Systems* 69: 170-178.

Xu, S., Klippel, A., MacEachren, A.M., and, Mitra, P. (2014) "Exploring Regional Variation in Spatial Language Using Spatially Stratified Web-Sampled Route Direction Documents." *Spatial Cognition and, Computation*. 14(4): 255-283 (2014).

Rokach, L., and, Mitra, P.: "Parsimonious citer-based measures: The artificial intelligence domain as a case study." *Journal of the Association for the Information Science and Technology* 64(9): 1951-1959 (2013).

Yuan, D., Mitra, P., and, Giles, C.L. (2013) "Mining and Indexing Graphs for Supergraph Search." *Proceedings of the VLDB Endowment* 6(10): 829-840.

Jaiswal, A.R., Miller, D.J., and Mitra, P. (2013) "Schema matching and embedded value mapping for databases with opaque column names and mixed continuous and discrete-valued data fields." *ACM Transactions on Database Systems*. 38(1): 2.

Yuan, D., and Mitra, P. (2013) "Lindex: a lattice-based index for graph databases." *VLDB Journal* 22(2): 229-252.

Bhatia, S., Caragea, C., Chen, H-H, Wu, Treeratpituk, P., Wu, Z., Khabsa, M., Mitra, P., and, Giles, C.L. (2012) "Specialized Research Datasets in the CiteSeer^x Digital Library." *D-Lib Magazine* 18(7/8).

Bhatia, S., and, Mitra, P. (2012) "Summarizing figures, tables, and algorithms in scientific publications to augment search results." *ACM Transactions on Information Systems* 30(1): 3.

Jaiswal, A.R., Pezanowski, S., Mitra, P., Zhang, X., Xu, S., Turton, I., Klippel, A., and, MacEachren, A.M. (2011) "GeoCAM: A geovisual analytics workspace to contextualize and interpret statements about movement." *Journal Spatial Information Science* 3(1): 65-101.

Sun, B., Mitra, P., Giles, C.L., Mueller, K.T. (2011) "Identifying, Indexing, and Ranking Chemical Formulae and Chemical Names in Digital Documents." *ACM Transactions on Information Systems*. 29(2): 12.

Ok, C., Lee, S., Mitra, P., and Kumara, S. (2010) "Distributed Routing in Wireless Sensor using Welfare Metric," *Information Sciences*, Elsevier Inc., 180(8):1183-1518.

Jaiswal, A.R., Miller, D. J., and Mitra, P. (2010) "Un-Interpreted Schema Matching with Embedded Value Mapping under Opaque Column Names and Data Values," *IEEE Transactions on Knowledge and Data Engineering*, IEEE Computer Society Press. 22(2): 291-304.

Ok, C-S., Lee, S., Mitra, P., and Kumara, S. (2009) "Distributed energy balanced routing for wireless sensor networks", *Computers and Industrial Engineering*, Elsevier Ltd., 57(1): 125-135.

Lu, X., Kataria, S., Brouwer, W.J., Wang, J.Z., Mitra, P. and Giles, C.L. (2009) "Automated Analysis of Images in Documents for Intelligent Document Search", *International Journal on Document Analysis and Recognition*, Springer Berlin/Heidelberg, 12(2): 65-81.

Lee, D., Kang, J., Mitra, P., Giles, C.L. and On, B-W. (2007) "Are Your Citations Clean?: New Scenarios and Challenges in Maintaining Digital Libraries," *Communications of the ACM (CACM)*, 50(12): 33-38.

Li, F., Luo, B., Liu, P., Lee, D., Mitra, P., Lee, W-C, and Chu, C-H. (2007) "In-broker Access Control: Towards Efficient End-to-End Performance of Information Brokerage Systems," *International Journal on Intelligent Control and Systems*, 12(4):283-292.

Afrati, F.N., Li, C., and Mitra, P. (2006) "Rewriting queries using views in the presence of arithmetic comparisons," *Theoretical Computer Science*, 368(1-2): 88-123.

Snow, D. R., Gahegan, M., Giles, C. L., Hirth, K. G., Milner, G. R., Mitra, P., and Wang, J.Z. (2006) "Next Generation Cybertools in Archaeological Science", *Science*, 311(5763): 958-959.

Debnath, S., Mitra, P., Pal, N. and Giles, C. L. (2005) “Automatic Identification of Informative Sections of Webpages”, *IEEE Transactions on Knowledge and Data Engineering, (TKDE)*, 17(9): 1233-1246.

Decker, S., Mitra, P. and Melnik, S. (2000) “Framework for the Semantic Web: An RDF Tutorial,” *IEEE Internet Computing*, 4(6): 68-73.

Peer Reviewed Conference Proceedings

Biyani, P., Bhatia, S., Caragea, C., and Mitra, P., (2015, to appear) “Using Subjectivity Analysis to Improve Thread Retrieval in Online Forums,” *Advances in Information Retrieval - 37th European Conference on IR Research, ECIR 2015. Proceedings. Lecture Notes in Computer Science*. Springer.

Huang, W., Wu, Z., Liang, C., Giles, C.L., Mitra, P., (2015, in press) “A Neural Probabilistic Model for Context Based Citation Recommendation.” *Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence*. AAAI Press.

Ofek, N., Rokach, L., and Mitra, P. (2014) “Methodology for Connecting Nouns to Their Modifying Adjectives.” In Alexander Gelbukh (Ed.), *15th International Conference on Computational Linguistics and Intelligent Text Processing -, CICLing*, Springer, Lecture Notes in Computer Science, Volume 8403:271-284.

Biyani, P. Caragea, C., Mitra, P., and, Yen, J. (2014) “Identifying Emotional and Informational Support in Online Health Communities.” *25th International Conference on Computational Linguistics, COLING*:827-836.

Bhatia, S., Biyani, P., and, Mitra, P. (2014): “Summarizing Online Forum Discussions - Can Dialog Acts of Individual Messages Help?.” In Alessandro Moschitti, Bo Pang, Walter Daelemans (Eds.), *Conference on Empirical Methods in Natural Language Processing, EMNLP*:2127-2131.

Banerjee, S., Caragea, C., and, Mitra, P. (2014) “Playscript Classification and Automatic Wikipedia Play Articles Generation.” *22nd International Conference on Pattern Recognition, ICPR*:3630-3635.

Wu, Z., Wu, J., Khabsa, M., Williams, K., Chen, H-H., Huang, W., Tuarob, S., Ray Choudhury, S., Ororbial, A., Mitra, P., and, Giles, C.L. (2014) “Towards building a scholarly big data platform: Challenges, lessons and opportunities.” *IEEE/ACM Joint Conference on Digital Libraries, JCDL*:117-126.

Huang, W., Wu, Z., Mitra, P., and, Giles C.L. (2014) “RefSeer: A citation recommendation system.” *IEEE/ACM Joint Conference on Digital Libraries, JCDL*:371-374.

Das Gollapalli, S., Qi, Y., Mitra, P., and, Giles, C.L. (2014) “Extracting Researcher Metadata with Labeled Features.” In Mohammed Javeed Zaki, Zoran Obradovic, Pang-Ning Tan, Arindam Banerjee, Chandrika Kamath, Srinivasan

Parthasarathy(Eds.) *14th SIAM International Conference on Data Mining*, SDM:740-748.

Biyani, P., Caragea, C., Mitra, P., Zhou, C., Yen, J., Greer, G. E., and, Portier, K. (2013) “Co-training over domain-independent and domain-dependent features for sentiment analysis of an online cancer support community.” In Jon G. Rokne, Christos Faloutsos (Eds.): *Advances in Social Networks Analysis and Mining*, ASONAM, ACM, New York, NY, 413-417.

Newman, M., Yen, J., Mitra, P., Murphy, W, Jacobson, N., and, Kim, H. (2013) “Supervised Machine Learning Classification of Journals Entries About Emotional Personal Experiences.” *American Medical Informatics Association Annual Symposium*, AMIA.

Wu, Z., Li, Z., Mitra, P., and, Giles, C.L. (2013) “Can back-of-the-book indexes be automatically created?” In Qi He, Arun Iyengar, Wolfgang Nejdl, Jian Pei, Rajeev Rastogi (Eds.), *22nd ACM International Conference on Information and Knowledge Management*, CIKM, ACM, New York, NY:1745-1750).

Oh, S., Lei, Z., Lee, W-C., Mitra, P., and, Yen, J. (2013) “CV-PCR: a context-guided value-driven framework for patent citation recommendation.” In Qi He, Arun Iyengar, Wolfgang Nejdl, Jian Pei, Rajeev Rastogi (Eds.), *22nd ACM International Conference on Information and Knowledge Management*, CIKM. ACM, New York, NY:2291-2296.

Wu, Z., Das Golapalli, S., Li, Z., Mitra, P., Giles, and, C. L. (2013) “Searching online book documents and analyzing book citations.” In Simone Marinai, Kim Marriott (Eds.): *ACM Symposium on Document Engineering, DocEng*. ACM, New York, NY: 81-90.

Wu, Z., Mitra, P., and, Giles, C.L. (2013) “Table of Contents Recognition and Extraction for Heterogeneous Book Documents.” *International Conference on Document Analysis and Recognition (ICDAR)*:1205-1209.

Tuarob, S., Bhatia, S., Mitra, P., and, Giles, C. L. (2013) “Automatic Detection of Pseudo-codes in Scholarly Documents Using Machine Learning. *International Conference on Document Analysis and Recognition (ICDAR)*:738-742.

Choudhury, S. R., Mitra, P., Rokach, L., and, Giles, C. L. (2013) “Figure Metadata Extraction in Digital Documents.” *12th International Conference on Document Analysis and Recognition (ICDAR)*:135-139.

Yuan, D., and, Mitra, P. (2013) “Cross Language Indexing and Retrieval of the Cypriot Digital Antiquities Repository.” *13th ACM Symposium on Document Engineering (DocEng)*, New York, NY: ACM:235-236.

Choudhury, S., Tuarob, S., Mitra, P., Rokach, L., Kirk, A., Szep, S., Pellegrino, D., Jones, S., and, Giles, C.G. (2013) “A figure search engine architecture for a chemistry digital library.” In J. Stephen Downie, Robert H.

McDonald, Timothy W. Cole, Robert Sanderson, Frank Shipman (Eds.), *13th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL)*. ACM, New York, NY:369-370.

Tuarob, S., Mitra, P., and, Giles, C. L. (2013) “A classification scheme for algorithm citation function in scholarly works.” In J. Stephen Downie, Robert H. McDonald, Timothy W. Cole, Robert Sanderson, Frank Shipman (Eds.), *13th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL)*. ACM, New York, NY:367-368.

Caragea, C., Silvescu, A., Mitra, P., and, Giles, C. L. (2013) “Can't see the forest for the trees?: a citation recommendation system.” In J. Stephen Downie, Robert H. McDonald, Timothy W. Cole, Robert Sanderson, Frank Shipman (Eds.), *13th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL)*. ACM, New York, NY: 111-114.

Chen, H.-H., Treeratpituk, P., Mitra, P., and, Giles, C. L. (2013) “CSSeer: an expert recommendation system based on CiteseerX.” In J. Stephen Downie, Robert H. McDonald, Timothy W. Cole, Robert Sanderson, Frank Shipman (Eds.), *13th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL)*. ACM, New York, NY: 381-382.

Gollapalli, S. D., and, Mitra, P., and, Giles, C. L. (2013) “Ranking experts using author-document-topic graphs.” In J. Stephen Downie, Robert H. McDonald, Timothy W. Cole, Robert Sanderson, Frank Shipman (Eds.), *13th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL)*. ACM, New York, NY: 87-96.

Gollapalli, S. D., Caragea, C., Mitra, P., and, Giles, C. L. (2013) “Researcher homepage classification using unlabeled data.” In Daniel Schwabe, Virgílio A. F. Almeida, Hartmut Glaser, Ricardo A. Baeza-Yates, Sue B. Moon (Eds.), *22nd International World Wide Web Conference*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland: 471-482.

Biyani, P., Caragea, C., and, Mitra, P. (2013) “Predicting Subjectivity Orientation of Online Forum Threads.” *Computational Linguistics and Intelligent Text Processing - 14th International Conference, CICLing 2013 Proceedings, Part II. Lecture Notes in Computer Science*. Springer (7817):109-120.

Jiang, Y., Ming, J., Wu, D., Yen, J., Mitra, P., Messner, J.I., and, Leicht, R. (2012) “BIM Server Requirements to Support the Energy Efficient Building Lifecycle,” In *Proceedings of the 2012 ASCE International Conference on Computing in Civil Engineering (ASCE 2012)*, Special Session on *BIM Computer Science Fundamentals*: 365-372.

Caragea, C., Silvescu, A., and, Mitra, P. (2012) “Combining Hashing and Abstraction in Sparse High Dimensional Feature Spaces.” In Jörg Hoffmann, Bart Selman (Eds.), *Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence*. AAAI Press.

Zhang, X., Qiu, B., Mitra, P., Xu, S., Klippel, A., and, MacEachren, A. (2012) “Disambiguating Road Names in Text Route Descriptions using Exact-All-Hop Shortest Path Algorithm.” *20th European Conference on Artificial Intelligence*. IOS Press:876-881.

Oh, S., Lei, Z., Mitra, P., and, Yen, J. (2012) “Evaluating and ranking patents using weighted citations.” In Karim B. Boughida, Barrie Howard, Michael L. Nelson, Herbert Van de Sompel, Ingeborg Sølvsberg (Eds.), *Proceedings of the 12th ACM/IEEE-CS Joint Conference on Digital Libraries, JCDL*, ACM Press, New York, NY:281-284.

Biyani, P., Caragea, C., Singh, A., and, Mitra, P. (2012) “I want what i need!: analyzing subjectivity of online forum threads.” *21st ACM International Conference on Information and Knowledge Management, CIKM*, ACM Press, New York, NY: 2495-2498.

Tuarob, S., Mitra, P., and, Giles, C. L. (2012) “Improving algorithm search using the algorithm co-citation network.” *Proceedings of the 12th ACM/IEEE-CS Joint Conference on Digital Libraries, JCDL*:277-280.

Zhang, Y., Mitra, P., Yen, J., and, Fu, W.-T. (2012) “Panel on social media for consumer health.” *ACM International Health Informatics Symposium, IHI '12*. ACM Press, New York, NY:889-890.

Golapalli, S. D., Mitra, P., and, Giles, C. L. (2012) “Phrase Pair Classification for Identifying Subtopics.” In Ricardo A. Baeza-Yates, Arjen P. de Vries, Hugo Zaragoza, Berkant Barla Cambazoglu, Vanessa Murdock, Ronny Lempel, Fabrizio Silvestri (Eds.), *Advances in Information Retrieval - 34th European Conference on IR Research, ECIR 2012, Barcelona, Spain, April 1-5, 2012. Proceedings. Lecture Notes in Computer Science*. Springer 7224. 489-493.

Huang, W., Kataria, S., Caragea, C., Mitra, P., Giles, C. L., and, Rokach, L. (2012) “Recommending citations: translating papers into references.” In Xue-wen Chen, Guy Lebanon, Haixun Wang, Mohammed J. Zaki (Eds.), *21st ACM International Conference on Information and Knowledge Management, CIKM'12*. ACM Press, New York, NY:1910-1914.

Gollapalli, S. D., Mitra, P. and, Giles, C. L. (2012) “Similar researcher search in academic environments.” *Proceedings of the 12th ACM/IEEE-CS Joint Conference on Digital Libraries, JCDL '12*. ACM Press, New York, NY: 167-170.

Fang, J., Mitra, P., and, Giles, C. L. (2012) “Table Header Detection and Classification for Digital Documents.” In Jörg Hoffmann, Bart Selman (Eds.), *Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence*. AAAI Press.

Wu, J., Terregowda, P. B., Fernandez Ramirez, J. P., Mitra, P., Zheng, S., and Giles, C. L. (2012) “The evolution of a crawling strategy for an academic document search engine: whitelists and blacklists.” In Noshir S. Contractor, Brian Uzzi, Michael W. Macy, Wolfgang Nejdl (Eds.), *Web Science 2012*. ACM Press, New York, NY: 340-343.

Biyani, P., Bhatia, S., Caragea, C., and, Mitra, P. (2012) “Thread Specific Features are Helpful for Identifying Subjectivity Orientation of Online Forum Threads.” In Martin Kay, Christian Boitet (Eds.): *Proceedings of the 24th International Conference on Computational Linguistics.. The COLING 2012 Organizing Committee*:295-310.

Yuan, D., Mitra, P., Yu, Huiwen, and, Giles, C.L. (2012) “Iterative Graph Feature Mining for Graph Indexing.” *IEEE 28th International Conference on Data Engineering (ICDE)*. IEEE: 198-209.

He, Q., Kifer, D., Pei, J., Mitra, P., and, Giles, C. L. (2011) “Citation Recommendation without Author Supervision.” In Irwin King, Wolfgang Nejdl, Hang Li (Eds.): *Proceedings of the Forth International Conference on Web Search and Web Data Mining, WSDM*. ACM: 755-764.

Caragea, C., Silvescu, A., Kataria, S., Caragea, D., and, Mitra, P. (2011) “Classifying Scientific Publications Using Abstract Features.” In Michael R. Genesereth, Peter Z. Revesz (Eds.): *Proceedings of the Ninth Symposium on Abstraction, Reformulation, and Approximation, SARA, AAAI*.

Caragea, C., McNeese, M. D., Jaiswal, A., Traylor, G., Kim, H., Mitra, P., Wu, D., Tapia, A. H., Giles, C. L., and, Jansen, B. J. (2011) “Classifying Text Messages for the Haiti Earthquake.” *8th International Conference on Information Systems for Crisis Response and Management (ISCRAM)*.

Kataria, S., Mitra, P., Caragea, C., and, Giles, C. L. (2011) “Context Sensitive Topic Models for Author Influence in Document Networks.” In Toby Walsh (Ed.): *22nd International Joint Conference on Artificial Intelligence, (IJCAI 2011)* IJCAI/AAAI:2274-2280.

Pan, C.-C., and, Mitra, P. (2011) “Event Detection with Spatial Latent Dirichlet Allocation.” *ACM/IEEE-CS Joint Conference on Digital Libraries, JCDL 2011*. ACM, New York, NY, USA: 349-358.

Robinson, A. C., Jaiswal, A., Pezanowski, S., Savelyev, A., Blanford, J., and, Mitra, P. (2011) “Geo-Twitter Analytics: Applications in Crisis Management.” In Anne Ruas, *25th International Cartographic Conference*. Paris

Qiu, B., Zhao, K., Mitra, P., Wu, D., Caragea, C., Yen, J., Greer, G. E., and, Portier, K. (2011) “Get Online Support, Feel Better - Sentiment Analysis and Dynamics in an Online Cancer Survivor Community.” *Third IEEE International Conference on Social Computing/SocialCom/PASSAT 2011*. IEEE:274-281.

Zhang, X., Mitra, P., Klippel, A., and, MacEachren, A. M. (2011) "Identifying Destinations Automatically from Human Generated Route Directions." *19th International Conference on Advances in Geographic Information Systems ACM Sigspatial GIS 2011*. ACM, New York, NY, USA:373-376.

Lahiri, S., Mitra, P., and, Lu, X. (2011) "Informality Judgment at Sentence Level and Experiments with Formality Score." In Alexander F. Gelbukh (Ed.): *Computational Linguistics and Intelligent Text Proceedings, Part II. Lecture Notes in Computer Science. 12th International Conference CICLing 2011*. Springer (6609):446-457.

Gollapalli, S. D., Giles, C. L., Mitra, P., and, Caragea, C. (2011) "On Identifying Academic Homepages for Digital Libraries." *ACM/IEEE-CS Joint Conference on Digital Libraries, JCDL 2011*. ACM, New York, NY:123-132.

Bhatia, S., Majumdar, D., and, Mitra, P. (2011) "Query Suggestions in the Absence of Query Logs." In Wei-Ying Ma, Jian-Yun Nie, Ricardo A. Baeza-Yates, Tat-Seng Chua, W. Bruce Croft (Eds.): *The 34th Annual ACM SIGIR Conference, SIGIR 2011*. ACM: 795-804.

Gollapalli, S. D., Mitra, P., and, Giles, C. L. (2011) "Ranking Authors in Digital Libraries." In Glen Newton, Michael J. Wright, Lillian N. Cassel (Eds.): *ACM/IEEE-CS Joint Conference on Digital Libraries, JCDL 2011, ACM:251-254*.

MacEachren, A. M., Jaiswal, A. R., Robinson, A. C., Pezanowski, S., Savelyev, A., Mitra, P., Zhang, X., and, Blanford, J. (2011) "SensePlace2: GeoTwitter Analytics Support for Situational Awareness." *IEEE Conference on Visual Analytics Science and Technology, IEEE VAST 2011*. IEEE Computer Society:181-190.

Caragea, C., Silvescu, A., and, Mitra, P. (2011) "Protein Sequence Classification Using Feature Hashing." In Fang-Xiang Wu, Mohammed Javeed Zaki, Shinichi Morishita, Yi Pan, Stephen Wong, Anastasia Christianson, Xiaohua Hu (Eds.): *IEEE International Conference on Bioinformatics and Biomedicine (BIBM 2011)*, IEEE:538-543.

Lahiri, S., Mitra, P., and, Lu, X. (2011) "Informality Judgment at Sentence Level and Experiments with F-score." In Alexander F. Gelbukh (Ed.): *Proceedings of the Twelfth International Conference on Intelligent Text Processing and Computational Linguistics, Part II. Lecture Notes in Computer Science, Springer-Verlag, Berlin: (6609) 446-457*.

He Q., Pei, J., Kifer, D., Mitra, P., and Giles, C.L. (2010) "Context-aware Citation Recommendation," In Michael Rappa, Paul Jones, Juliana Freire, Soumen Chakrabarti (Eds.): *19th International World Wide Web Conference (WWW 2010)*, ACM Press, 421:430.

He Q., Mitra, P., Giles, C.L., Pei, J., Chen, B., and Qiu, B., (2009) “Detecting Topic Evolution in Scientific Literature: How Can Citations Help?,” *18th ACM Conf. on Information and Knowledge Management (CIKM’09)*, ACM Press, New York, NY, 957-966.

Sun, B., Mitra, P., and Giles C. L., (2009) “Independent Informative Subgraph Mining for Graph Information Retrieval,” *18th ACM Conf. on Information and Knowledge Management (CIKM’09)*, ACM Press, New York, NY, 563-572.

Zhu, L., and Mitra, P. (2009) “Multidimensional Political Spectrum Identification and Analysis,” *18th ACM Conf. on Information and Knowledge Management (CIKM’09)*, ACM Press, New York, NY, 2045-2048.

Bhatia, S., Lahiri, S. and Mitra, P. (2009) “Generating Synopses for Document-Element Search,” *18th ACM Conf. on Information and Knowledge Management (CIKM’09)*, ACM Press, New York, NY, 2003-2006.

Sun, B., Mitra, P., and Giles C. L., (2009) “Learning to Rank Graphs for Online Similar Graph Search,” *18th ACM Conf. on Information and Knowledge Management (CIKM’09)*, ACM Press, New York, NY, 1871-1874.

Afrati, F., Chandrachud, M., Chirkova, R., and Mitra, P. (2009) “Approximate Rewriting of Queries Using Views,” *13th East European Conference on Advances in Databases and Information Systems, (ADBIS’09)*, Riga, Latvia, *Lecture Notes in Computer Science*, Springer Berlin/Heidelberg, 164-178.

Liu, Y., Mitra, P., and Giles, C.L. (2009) “Improving the Table Boundary Detection in PDFs by Fixing the Sequence Error of the Sparse Lines,” *10th International Conference on Document Analysis and Recognition (ICDAR’09)*, Barcelona, Spain, IEEE Computer Society Press, Los Alamitos, 1006-1010.

Tan Q., Mitra, P., and Giles, C.L. (2009) “Effectively Searching for Maps in Web Documents”, *31st European Conference on Information Retrieval, (ECIR’09)*, Toulouse, France, *Advances in Information Retrieval, Lecture Notes in Computer Science*, Springer Berlin/Heidelberg, 5478(2009): 162-176.

Sun, B., Mitra, P., and Giles, C.L. (2008) “Mining, Indexing, and Searching for Textual Chemical Molecule Information on the Web,” *The 17th International World Wide Web Conference, (WWW’08)*, ACM Press, New York, NY, 735-744.

Zhang, X., Lee, W-C., Mitra, P., and Zheng, B. (2008) “Processing Transitive Nearest Neighbor Queries in Multi-Channel Access Environments,” *11th Conf. on Extending Database Technology (EDBT’08)*, Nantes, France, *ACM International Conference Proceeding Series*, ACM Press, New York, 261: 452-463.

Huang, S., Zhao, Q., Mitra, P., and Giles, C.L. (2008) “Hierarchical Location and Topic Based Query Expansion,” *23rd AAAI Conference on Artificial Intelligence*,

(AAAI'08), AAAI Press, 1150-1155.

Kataria, S., Brouwer, W., Mitra, P., and Giles, C.L. (2008) "Automatic Extraction of Data Points and Text Blocks from 2-Dimensional Plots in Digital Documents," *23rd AAAI Conference on Artificial Intelligence, (AAAI'08)*, AAAI Press, 1169-1174.

Brouwer, W., Kataria, S., Das, S., Mitra, P., and Giles, C.L. (2008) "Segregating and extracting overlapping data points in two-dimensional plots," *ACM and IEEE-CS Joint Conference on Digital Libraries, (JCDL'08)*, ACM Press, New York, NY, 276-279.

Liu, Y., Lita, L.V., Niculescu, R.S., Mitra, P., and Giles, C.L. (2008) "Finding a Haystack in Haystacks - Simultaneous Identification of Concepts in Large Bio-Medical Corpora," *SIAM International Conference on Data Mining, (SDM'08)*, SIAM, 668-679.

Liu, Y., Lita, L.V., Niculescu, R.S., Mitra, P., and Giles, C.L. (2008) "Real-Time Data Pre-Processing Technique for Efficient Feature Extraction in Large Scale Datasets," *17th ACM Conf. on Information and Knowledge Management (CIKM'08)*, ACM Press, New York, NY, 981-990.

Liu, Y., Mitra, P., and Giles, C.L. (2008) "Identifying Table Boundaries in Digital Documents via Sparse Line Detection," *17th ACM Conf. on Information and Knowledge Management (CIKM'08)*, ACM Press, New York, NY, 1311-1320.

Tan Q., Mitra, P., and Giles, C.L. (2008) "Metadata Extraction and Indexing for Map Search in Web Documents," *17th ACM Conf. on Information and Knowledge Management (CIKM'08)*, ACM Press, New York, NY, 1367-1368.

Zhuang, Z., Brunk, C., Mitra, P., and Giles, C. L. (2008) "Towards Click-Based Models of Geographic Interests in Web Search," *IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, IEEE Computer Society Press, Los Alamitos, CA, 1: 293-299.

Tan, Q., Mitra, P., and Giles, C.L. (2007) "Designing Clustering-Based Web Crawling Policies for Search Engine Crawlers," *ACM Sixteenth Conference on Information and Knowledge Management (CIKM)*, Lisbon, Portugal, ACM Press, New York, NY, 535-544.

Chen, B., Zhao, Q., Sun, B., and Mitra, P. (2007) "Temporal And Social Network Based Blogging Behavior Prediction In BlogSpace," *Seventh IEEE International Conference on Data Mining (ICDM)*, Omaha, NE, IEEE Computer Society Press, Los Alamitos, CA, 439-444.

Zhao, Q., Mitra, P., Lee, D. and Kang, J. "HICCUP: Hierarchical Clustering Based Value Imputation using Heterogeneous Gene Expression Microarray Datasets," *Seventh IEEE International Conference on Bioinformatics and Bioengineering (BIBE)*, Boston, MA, IEEE Computer Society Press, Los

Alamitos, CA, 71-78.

Lu, X., Wang, J., Mitra, P., and Giles, C.L. (2007) "Automatic Extraction of Data from 2-D Plots in Documents", *9th International Conference on Document Analysis and Recognition (ICDAR)*, Curitiba, Brazil, IEEE Computer Society Press, Los Alamitos, CA, 1:188-192.

Liu, Y., Bai, K., Mitra, P., and Giles, C.L. (2007) "Searching for Tables in Digital Documents" *9th International Conference on Document Analysis and Recognition (ICDAR)*, Curitiba, Brazil, IEEE Computer Society Press, Los Alamitos, CA, 2:934-938.

Zhao, Q., Mitra, P., and Chen, B. (2007) "Temporal and Information Flow Based Event Detection from Social Text Streams," *Twenty-Second Conference on Artificial Intelligence (AAAI)*, Vancouver, Canada, AAAI Press, 1501-1506

Liu, Y., Bai, K., Mitra, P., and Giles, C.L. (2007) "TableRank: A Ranking Algorithm for Table Search and Retrieval," *Twenty-Second Conference on Artificial Intelligence (AAAI)*, Vancouver, Canada, AAAI Press, 317-322.

Sun, B., Mitra, P., Giles, C.L., Yen, J., and Zha, H. (2007) "Topic segmentation with shared topic detection and alignment of multiple documents," *30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, Amsterdam, The Netherlands, ACM Press, New York, NY, 199-206.

Tan, Q., Zhuang, Z., Mitra, P., and Giles, C.L. (2007) "Efficiently Detecting Webpage Updates Using Samples," *Seventh International Conference on Web Engineering (ICWE)*, Como, Italy, *Lecture Notes in Computer Science*, Springer Berlin/Heidelberg, 4607(2007): 285-300.

Liu, Y., Bai, K., Mitra, P., and Giles, C.L. (2007) "TableSeer: automatic table metadata extraction and searching in digital libraries," *ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL)*, Vancouver, Canada, ACM Press, New York, NY, 91-100.

Sun, B., Tan, Q., Mitra, P., and Giles, C.L. (2007) "Extraction and search of chemical formulae in text documents on the web," *16th International World Wide Web Conference, (WWW)*, Banff, Alberta, Canada, 251-260.

Liu, Y., Bai, K., Mitra, P., and Giles, C.L. (2007) "Automatic searching of tables in digital libraries" *16th International World Wide Web Conference, (WWW)*, Banff, Alberta, Canada, 1135-1136.

Tan, Q., Zhuang, Z., Mitra, P., and Giles, C.L. (2007) "Designing efficient sampling techniques to detect webpage updates," *16th International World Wide Web Conference, (WWW)*, Banff, Alberta, Canada, 1147-1148.

[co-author, 45%]

- Lu, X., Wang, J.Z., Mitra, P., and Giles, C.L. (2007) "Deriving knowledge from figures for digital libraries," *16th International World Wide Web Conference, (WWW'07)*, Banff, Alberta, Canada, 1229-1230.
- Zhao, Q., Mitra, P., and Giles, C.L. (2007) "Image annotation by hierarchical mapping of features," *16th International World Wide Web Conference, (WWW'07)*, Banff, Alberta, Canada, 1237-1238.
- Ok, C., Mitra, P., Lee, S. and Kumara, S. (2007) "Maximum Energy Welfare Routing in Wireless Sensor Networks," *Networking 2007, 6th International IFIP-TC6 Networking Conference*, Atlanta, GA, *Lecture Notes in Computer Science*, Springer Verlag GmbH, 203-214.
- Zhao, Q., and Mitra, P. (2007) "Event Detection and Visualization for Social Text Streams," *International Conference on Weblogs and Social Media, (ICWSM'07)*, Boulder, CO, 4 pages.
- Ayyagari, P., Mitra, P., Lee, D., Liu, P., and Lee, W-C. (2007) "Incremental Adaptation of XPath Access Control Views," *2007 ACM Symposium on Information, Computer and, Communications Security, (ASIACCS'07)*, Singapore, 105-116.
- Ok, C.-S., Mitra, P., Lee, S. and Kumara, S. (2007) "Distributed Energy Adaptive Routing for Wireless Sensor Networks," *Third Annual IEEE Conference on Automation Science and Engineering, (CASE 2007)*, Phoenix, AZ, IEEE Press, 905-910.
- Lu, X., Mitra, P., Wang, J.Z., and Giles, C.L. (2006) "Automatic categorization of figures in scientific documents," *ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL 2006)*, Chapel Hill, NC, ACM Press, New York, NY, 129-138.
- Liu, Y., Bai, K., Mitra, P., and Giles, C.L. (2006) "Automatic extraction of table metadata from digital documents," *ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL 2006)*, Chapel Hill, NC, ACM Press, New York, NY, 339-340.
- Pan, C-C., Mitra, P., and Liu, P. (2006) "*Semantic Access Control for Information Interoperation*," *11th ACM Symposium on Access Control Models and Technologies (SACMAT '06)*, Lake Tahoe, CA, ACM Press, New York, NY, 237-246.
- Li, F., Luo, B., Liu, P., Lee, D., Mitra, P., Lee, W-C., and Chu, C-H. (2006) "In-Broker Access Control towards Efficient End-to-End Performance of Information Brokerage Systems," *IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing (SUTC)*, Taichung, Taiwan, 1:252-259.
- Ayyagari, P.S., Mitra, P. and Hurson, A. (2006) "Efficient Object Retrieval from Parallel Air Channels in the Presence of Replicated Objects," *The 7th International Conference on Mobile Data Management (MDM'06)*, Nara, Japan, 16 IEEE Computer Society: (10 pages).

- Mitra, P., Pan, C-C., Liu, P., and Atluri, V. (2006) "Privacy-preserving Semantic Interoperation and Access Control of Heterogeneous Databases," *ACM Symposium on Information, Communcation, and Computer Security (ASIA-CCS)*, Taiwan, 66-77.
- Kang, J., Lee, D., and Mitra, P. (2005) "Identifying Value Mappings for Data Integration: an Unsupervised Approach," *Sixth International Conf. on Web Information Systems Engineering (WISE)*, New York, NY, *Lecture Notes in Computer Science*, Springer Verlag GmbH, 3806(XXI, 771): 544-551.
- Mitra, P., Noy, N.F., and Jaiswal, A. R. (2005) "Ontology Mapping Discovery with Uncertainty," In Gil, Y., Motta, E., Benjamin, V.R., Musen, M.(Eds.) *Fourth International Semantic Web Conference (ISWC)*, Galway, Ireland, *Lecture Notes in Computer Science*, Springer Verlag GmbH, 3729(2005): 537-547.
- Kang, J., Han, T.S., Lee, D., and Mitra, P. (2005) "Establishing Value Mappings Using Statistical Models and User Feedback," *14th ACM Conference on Information and Knowledge Management (CIKM)*, Bremen, Germany, ACM Press, New York, NY, 68-75.
- Mitra, P., Wiederhold, G., and Decker, S. (2005) "Dealing with Semantic Interoperation of Data," *Second IEEE International Symposium on Data Interoperability*, Sardinia, Italy, IEEE Computer Society Press, 134-145.
- On, B-W., Lee, D., Kang, J., and Mitra, P. (2005) "Comparative Study of Name Disambiguation Problem using a Scalable Blocking-based Framework," *ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL 2005)*, Denver, Colorado, ACM Press, New York, NY, 434: 344-353.
- Debnath, S., Mitra, P., and Giles, C. L. (2005) "Identifying Content Blocks from Web Documents," *15th International Symposium on Methodologies for Intelligent Systems (ISMIS 2005)*, In Hacid, M-S., Murray, N.V., Ras, Z.W. and Tsumoto S. (Eds.), *Foundations of Intelligent Systems*, Saratoga Springs, NY, *Lecture Notes in Computer Science*, Springer Verlag GmbH, 3488: 285-293.
- Debnath, S., Mitra, P., and Giles, C.L. (2005) "Finding the Base Time-line of a News Article," In Russell, I. and Markov, Z. (Eds.), *18th International Florida Artificial Intelligence Research Society Conference (FLAIRS 2005)*, Clearwater Beach, FL, AAAI Press, 142-147.
- Debnath, S., Mitra, P., and Giles, C. L. (2005) "Automatic Extraction of Informative Blocks from Webpages," *20th Annual ACM Symposium on Applied Computing (SAC), Special Track on Web Technologies and Applications*, Santa Fe, New Mexico, ACM Press, New York, NY, 1722-1726.
- Afrati, F., Li, C., and Mitra, P. (2004) "On Containment of Conjunctive Queries with Arithmetic Comparisons," In Bertino, E.; Christodoulakis, S.; Plexousakis, D.; Vassilis, C.; Koubarakis, M.; Böhm, K.; Ferrari, E. (Eds.), *Advances in*

Database Technology- (EDBT 2004), 9th International Conference on Extending Database Technology, Heraklion, Crete, Greece, *Lecture Notes in Computer Science*, Springer Verlag GmbH, 2992 (XVIII, 877): 459-476.

Afrati, F., Li, C., and Mitra, P. (2002) "Answering Queries Using Views with Arithmetic Comparisons," *21st ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems (PODS)*, Madison, WI, ACM Press, New York, NY, 209-220.

Mitra, P. and Wiederhold, G. (2001) "An Algebra for Semantic Interoperability of Information Sources," *2nd IEEE Symposium on BioInformatics and Bioengineering (BIBE)*, Bethesda, MD, IEEE Computer Society Press, 174-182.

Mitra, P., Wiederhold, G., and Decker, S. (2001) "A Scalable Framework for Interoperation of Information Sources," In Cruz, I.F., Decker, S., Euzenat, J., McGuinness, D.L., (Eds.), *1st International Semantic Web Working Symposium (SWWS)*, Stanford, CA, 317-329.

Mitra, P. (2001) "An Algorithm for Answering Queries Efficiently Using Views," In Orłowska, M.E., Roddick, J.F., *12th. Australasian Database Conference (ADC 2001)*, *Australian Computer Science Communications*, Gold Coast, Qld., Australia, IEEE Computer Society Press, 23(2): 99-106.

Decker, S., Jannink, J., Melnik, S., Mitra, P., Staab, S., Studer, R., and Wiederhold, G. (2000) "An Information Food Chain for Advanced Applications on the WWW," In Borbinha, J.L.; Baker, T., (Eds.), *Research and Advanced Technology for Digital Libraries, 4th European Conference, ECDL 2000*, Lisbon, Portugal, *Lecture Notes in Computer Science*, Springer Verlag GmbH, 1923: 490-493.

Mitra, P., Kersten, M., and Wiederhold, G. (2000) "A Graph-Oriented Model for Articulation of Ontology Interdependencies," In Zaniolo, C., Lockemann, P.C., Scholl, M.H., Grust, T. (Eds.): *Advances in Database Technology - EDBT 2000, 7th International Conference on Extending Database Technology*, Konstanz, Germany, *Lecture Notes in Computer Science*, Springer Verlag GmbH, 1777.

Mitra, P., Wiederhold, G., and Jannink, J. (1999) "Semi-automatic Integration of Knowledge Sources," *2nd International Conference on Information Fusion (FUSION)*, Sunnyvale, CA, Omnipress, 572-581.

Mitra, P., Payne, D., Shuler, L., van de Geijn, R., and Watts, J. (1995) "Fast Collective Communication Libraries, Please," *Intel Supercomputing Users' Group Meeting*.

Dasgupta, P., Mitra, P., and Chakraborty, P.P. (1994) "Multiobjective Search in VLSI Design," *7th International Symposium on VLSI Design (VLSI)*, Calcutta, India, IEEE Computer Society Press, 395-400.

Peer Reviewed Workshop Publications

Karimzadeh, M., Huang, W., Banerjee, S., Wallgrün, J.O., Hardisty, F., Scott Pezanowski, S., Mitra, P., and, MacEachren, A.M. (2013) “GeoTxt: a web API to leverage place references in text,” In Jones, C.B., Purves, R. (Eds.): *Proceedings of the 7th Workshop on Geographic Information Retrieval, GIR 2013*, ACM, 72-73.

Jiang, Y., Ming, J., Wu, D., Yen, J., Mitra, P., Messner, J.I., and, Leicht, R. (2012) “BIM Server Requirements to Support the Energy Efficient Building Lifecycle,” In *Proceedings of the 2012 ASCE International Workshop on Computing in Civil Engineering held in conjunction with 2012 ASCE International Conference on Computing in Civil Engineering (ASCE 2012)*.

Jiang, Y., Ming, J., Wu, D., Yen, J., Mitra, P., Messner, J.I., and, Leicht, R. (2012) “A cloud-centric building information model data hub via REST interfaces,” In *Proceedings of the 2012 ASCE International Workshop on Computing in Civil Engineering held in conjunction with 2012 ASCE International Conference on Computing in Civil Engineering (ASCE 2012)*.

Wu, J., Teregowda, P.B., Khabsa, M., Carman, S., Jordan, D., Wandelmer, J.S.P., Lu, X., Mitra, P., and, Giles, C.L. (2012) “Web crawler middleware for search engine digital libraries: a case study for citeseerX,” In George H. L. Fletcher, Prasenjit Mitra (Eds.): *Proceedings of the Twelfth International Workshop on Web Information and Data Management, WIDM 2012*, ACM, 57-64.

Yuan, D., and, Mitra, P. (2011) “A Lattice-based Graph Index for Subgraph Search,” In Marian, A., and, Vassalos, V. (Eds.): *Proceedings of the 14th International Workshop on the Web and Databases, WebDB 2011*.

Zhang, X., and, Mitra, P. (2010) “Learning Topical Transition Probabilities in Click Through Data with Regression Models,” In Dong, X.L., and, Naumann, F., (Eds.): *Proceedings of the 13th International Workshop on the Web and Databases 2010, WebDB 2010*.

Zhang, X., Mitra, P., Xu, S., Jaiswal, A.R., Klippel, A., and, MacEachren, A.M. (2009) “Extracting Route Directions from Web Pages,” *12th International Workshop on the Web and Databases, WebDB 2009*.

Tan, Q., Zhuang, Z., Mitra, P., and, Giles, C.L. (2007) “A clustering-based sampling approach for refreshing search engine's database,” *Tenth International Workshop on the Web and Databases, WebDB 2007*.

Jaiswal, A.R., Giles, C.L., Mitra, P., and, Wang, J.Z. (2006) “An architecture for creating collaborative semantically capable scientific data sharing infrastructures,” In Bonifati, A., Fundulaki, I. (Eds.): *Eighth ACM International Workshop on Web Information and Data Management (WIDM 2006)*, ACM 2006, 75-82.

Huang, S., Zhao, Q., Mitra, P., and, Giles, C.L. (2007) "Query Expansion Using Topic and Location," *Workshops Proceedings of the 7th IEEE International Conference on Data Mining (ICDM 2007)*, IEEE Computer Society, 619-624.

Sponsored Projects:

Mitra, P. (Core Faculty), Grant, "IGERT: Big Data Social Science - An Integrative Education and Research Program in Social Data Analytics," National Science Foundation, Federal Agencies. Total awarded: \$1,181,834.00. July 1, 2012 - June 30, 2017.

Mitra, P. (Co-Principal Investigator), Grant, "Collaborative Research: CI-ADDO-EN: Semantic CiteSeer X," National Science Foundation, Federal Agencies. Total awarded: \$198,402.00. July 1, 2010 - June 30, 2014.

Mitra, P. (Co-Principal Investigator), Grant, "Collaborative Research: CI-ADDO-EN: Semantic CiteSeer X," National Science Foundation, Federal Agencies. Total awarded: \$199,622.00. July 1, 2010 - June 30, 2014.

Mitra, P. (Principal Investigator), Grant, "CAREER: An Integrated Framework for Extracting and Utilizing Information from Tables and Figures in Digital Libraries," National Science Foundation, Federal Agencies. Total awarded: \$449,782.00. June 1, 2009 - May 31, 2014.

Mitra, P. (Co-Principal Investigator), Contract, "Chemical Linguistic Analysis with Chem Seer," Dow Chemical Company, Corporations. Total awarded: \$400,000.00. October 26, 2011 - October 25, 2013.

Mitra, P. (Co-Principal Investigator), Contract, "New Methods for Representing and Interacting with Qualitative Geographic Information," U.S. Army Corps of Engineers, Federal Agencies. Total awarded: \$100,000.00. September 30, 2012 - October 16, 2013.

Mitra, P. (Co-Principal Investigator), Grant, "E2E Penn State - Purdue Visual Analytics Integration (GeoText-API)," Purdue University, Universities and Colleges. Total awarded: \$48,988.00. November 15, 2012 - June 30, 2013.

Mitra, P. (Principal Investigator), Wang, J. Z. (Principal Investigator), Snow, D. R. (Co-Principal Investigator), Kalisperis, L. N. (Co-Principal Investigator), Giles, C. L. (Co-Principal Investigator), Grant, "REU: EAGER: Analysis and Intelligent Search for Cypriot Works of Art and Secretariat Corpus," National Science Foundation, Federal Agencies. Total awarded: \$16,000.00. September 1, 2009 - August 31, 2012.

Giles, C. L. (Principal Investigator), Jansen, B. J. (Co-Principal Investigator), Mitra, P. (Co-Principal Investigator), Grant, "REU Collaborative Research: CI-

ADDO-EN: Semantic CiteSeer X," National Science Foundation, Federal Agencies. Total awarded: \$199,843.00. July 1, 2010 - June 30, 2012.

Ritter, F. E. (Principal Investigator), Hall, D. L. (Co-Principal Investigator), Horgan, J. G. (Co-Principal Investigator), Haynes, S. R. (Co-Principal Investigator), Kreager, D. A. (Co-Principal Investigator), Mitra, P. (Co-Principal Investigator), Liu, P. (Co-Principal Investigator), Georgieva, K. I. (Co-Principal Investigator), Shemanski, D. R. (Co-Principal Investigator), Yen, J. (Co-Principal Investigator), Grant, "Adversarial Motivation and Intent to Acquire, Proliferate, and Potentially Use WMD," Defense Threat Reduction Agency, Federal Agencies. Total awarded: \$348,217.00. June 29, 2009 - June 28, 2012.

Ritter, F. E. (Principal Investigator), Hall, D. L. (Co-Principal Investigator), Horgan, J. G. (Co-Principal Investigator), Haynes, S. R. (Co-Principal Investigator), Kreager, D. A. (Co-Principal Investigator), Mitra, P. (Co-Principal Investigator), Liu, P. (Co-Principal Investigator), Georgieva, K. I. (Co-Principal Investigator), Shemanski, D. R. (Co-Principal Investigator), Yen, J. (Co-Principal Investigator), Grant, "Adversarial Motivation and Intent to Acquire, Proliferate, and Potentially Use WMD," Defense Threat Reduction Agency, Federal Agencies. Total awarded: \$349,524.00. June 29, 2009 - June 28, 2012.

Ritter, F. E. (Principal Investigator), Hall, D. L. (Co-Principal Investigator), Horgan, J. G. (Co-Principal Investigator), Haynes, S. R. (Co-Principal Investigator), Kreager, D. A. (Co-Principal Investigator), Mitra, P. (Co-Principal Investigator), Liu, P. (Co-Principal Investigator), Georgieva, K. I. (Co-Principal Investigator), Shemanski, D. R. (Co-Principal Investigator), Yen, J. (Co-Principal Investigator), Grant, "Adversarial Motivation and Intent to Acquire, Proliferate, and Potentially Use WMD," Defense Threat Reduction Agency, Federal Agencies. Total awarded: \$349,734.00. June 29, 2009 - June 28, 2012.

Yen, J. (Principal Investigator), Mitra, P. (Co-Principal Investigator), Tapia, A. H. (Co-Principal Investigator), Giles, C. L. (Co-Principal Investigator), Jansen, B. J. (Co-Principal Investigator), Grant, "RAPID: Text Message-based Infrastructure for Emergency Response," National Science Foundation, Federal Agencies. Total awarded: \$75,000.00. March 15, 2010 - February 29, 2012.

Maceachren, A. (Principal Investigator), Mitra, P. (Co-Principal Investigator), Klippel, A. K. (Co-Principal Investigator), Grant, "Geographic Contextualization for Accounts of Movement (GeoCAM): Representing, Extracting, Mapping, and Interpreting Movement References in Text," National Geospatial-Intelligence Agency, Federal Agencies. Total awarded: \$91,148.00. August 23, 2007 - December 30, 2011.

Giles, C. L. (Principal Investigator), Mitra, P. (Co-Principal Investigator), Contract, "Develop a tool for automatic pdf conversion to ascii and specific metadata extraction," National Science Foundation, Federal Agencies. Total awarded: \$24,864.00. May 27, 2010 - August 31, 2011.

Mitra, P. (Principal Investigator), Wang, J. Z. (Principal Investigator), Snow, D. R. (Principal Investigator), Kalisperis, L. N. (Co-Principal Investigator), Giles, C. L. (Co-Principal Investigator), Grant, "EAGER: Analysis and Intelligent Search for Cypriot Works of Art and Secretariat Corpus," National Science Foundation, Federal Agencies. Total awarded: \$249,915.00. September 1, 2009 - August 31, 2011.

Maceachren, A. (Principal Investigator), Mitra, P. (Co-Principal Investigator), Klippel, A. K. (Co-Principal Investigator), Grant, "Geographic Contextualization for Accounts of Movement (GeoCAM): Representing, Extracting, Mapping, and Interpreting Movement References in Text," National Geospatial-Intelligence Agency, Federal Agencies. Total awarded: \$58,890.00. (August 23, 2007 - August 22, 2011.

Giles, C. L. (Principal Investigator), Mitra, P. (Co-Principal Investigator), Mueller, K. T. (Co-Principal Investigator), Contract, "ORE-Chem (OAI-ORE Exemplar Project for Open Access Chemistry)," Microsoft Corporation, Corporations. Total awarded: \$200,000.00. August 15, 2008 - August 15, 2011.

Giles, C. L. (Principal Investigator), Mitra, P. (Co-Principal Investigator), Jansen, B. J. (Co-Principal Investigator), Grant, "Collaborative Research: CI-ADDO-EN: Semantic CiteSeer X," National Science Foundation, Federal Agencies. Total awarded: \$299,772.00. July 1, 2010 - June 30, 2011.

Giles, C. L. (Principal Investigator), Jansen, B. J. (Co-Principal Investigator), Mitra, P. (Co-Principal Investigator), Grant, "REU Collaborative Research: CI-ADDO-EN: Semantic CiteSeer X," National Science Foundation, Federal Agencies. Total awarded: \$16,000.00. July 1, 2010 - June 30, 2011.

Maceachren, A. (Principal Investigator), Cai, G. (Principal Investigator), Mitra, P. (Principal Investigator), Grant, "Visualization to Support Coordinated Geospatial Decision-Making Interpretation, Representation of Quality, Data Fusion, and Collaboration," Purdue University, Universities and Colleges. Total awarded: \$168,750.00. July 1, 2009 - May 31, 2011.

Mueller, K. T. (Principal Investigator), Garrison, B. J. (Co-Principal Investigator), Mitra, P. (Co-Principal Investigator), Giles, C. L. (Co-Principal Investigator), Kubicki, J. (Co-Principal Investigator), Grant, "Cyberinfrastructure and research facilities: developing collaborative tools to facilitate multi-disciplinary multi scale research in environmental molecular sciences," National Science Foundation, Federal Agencies. Total awarded: \$1,355,875.00. September 1, 2005 - August 31, 2010.

Maceachren, A. (Principal Investigator), Mitra, P. (Co-Principal Investigator), Klippel, A. K. (Co-Principal Investigator), Grant, "Geographic Contextualization for Accounts of Movement (GeoCAM): Representing, Extracting, Mapping, and Interpreting Movement References in Text," National Geospatial-Intelligence Agency, Federal Agencies. Total awarded: \$148,882.00. August 23, 2007 - August 22, 2010.

Simpson, T. W. (Principal Investigator), Lee, D. (Co-Principal Investigator), Mitra, P. (Co-Principal Investigator), Kremer, G. (Co-Principal Investigator), Zhang, X. (Co-Principal Investigator), Grant, "CI-TEAM Implementation Project: Collaborative Research - A National Engineering Dissection Cyber-Collaboratory," National Science Foundation, Federal Agencies. Total awarded: \$151,700.00. January 1, 2007 - December 31, 2009.

Maceachren, A. (Principal Investigator), Mitra, P. (Principal Investigator), Klippel, A. K. (Principal Investigator), Grant, "Geographic Contextualization for Accounts of Movement (GeoCAM): Representing, extracting, Mapping," National Geospatial-Intelligence Agency, Federal Agencies. Total awarded: \$299,998.00. August 23, 2007 - August 23, 2009.

Maceachren, A. (Principal Investigator), Mitra, P. (Principal Investigator), Cai, G. (Principal Investigator), Contract, "Knowledge-Enabled Visual Analytics: Supporting Individuals and Teams from Analysis through Action," Battelle - Pacific Northwest National Laboratory, Federal Laboratories. Total awarded: \$337,945.00. January 3, 2006 - April 24, 2009.

Maceachren, A. (Principal Investigator), Mitra, P. (Co-Principal Investigator), Cai, G. (Co-Principal Investigator), McNeese, M. D. (Co-Principal Investigator), Gahegan, M. N. (Co-Principal Investigator), Contract, "Knowledge-Enabled Visual Analytics: Supporting Individuals and Teams from Analysis through Action," Battelle - Pacific Northwest National Laboratory, Federal Laboratories. Total awarded: \$344,000.00. January 3, 2006 - September 30, 2008.

Snow, D. R. (Principal Investigator), Giles, C. L. (Research Associate), Mitra, P. (Research Associate), Contract, "Digital Antiquity: Planning a Digital Information Infrastructure for Archaeology," University of Arkansas, Universities and Colleges. Total awarded: \$22,024.00. July 1, 2007 - June 30, 2008.

Maceachren, A. (Principal Investigator), Gahegan, M. N. (Principal Investigator), McNeese, M. D. (Principal Investigator), Mitra, P. (Principal Investigator), Contract, "Visualization to Support Coordinated Geospatial Decision-Making: Interpretation, Representation of Quality," Battelle - Pacific Northwest National Laboratory, Federal Laboratories. Total awarded: \$266,000.00. January 3, 2006 - April 1, 2008.

Mueller, K. T. (Principal Investigator), Garrison, B. J. (Co-Principal Investigator), Mitra, P. (Co-Principal Investigator), Giles, C. L. (Co-Principal Investigator), Kubicki, J. (Co-Principal Investigator), Grant, "Cyberinfrastructure and research facilities: developing collaboratory tools to facilitate multi-disciplinary multi scale research in environmental molecular sciences," National Science Foundation, Federal Agencies. Total awarded: \$937,250.00. September 1, 2005 - August 31, 2007.

Maceachren, A. (Principal Investigator), Gahegan, M. N. (Co-Principal Investigator), McNeese, M. D. (Co-Principal Investigator), Mitra, P. (Co-

Principal Investigator), Contract, "Knowledge-Enabled Visual Analytics: Supporting Individuals and Teams from Analysis through Action (Nat'l. Visualization and, Analytics Ctr.)," Battelle - Pacific Northwest National Laboratory, Federal Laboratories. Total awarded: \$749,943.00. January 3, 2006 - July 2, 2007.

Giles, C. L. (Principal Investigator), Mitra, P. (Co-Principal Investigator), Wang, J. Z. (Co-Principal Investigator), Grant, "Unrestricted Gift," Microsoft Corporation, Corporations. Total awarded: \$20,000.00. March 1, 2006 - February 28, 2007.

Mitra, P. (Principal Investigator), Contract, "Identifying Rare Events from Multiple Information Sources," Battelle - Oak Ridge National Laboratory, Federal Laboratories. Total awarded: \$58,510.00. March 15, 2006 - September 30, 2006.

Chu, C.-H. (Principal Investigator), Liu, P. (Co-Principal Investigator), Mitra, P. (Co-Principal Investigator), Zhu, S. (Co-Principal Investigator), Grant, "DOD IA Scholarship Program at Penn State/Institutional Capacity Building at Penn State," U.S. Department of Defense, Federal Agencies. Total awarded: \$247,139.00. August 12, 2005 - August 11, 2006.

Mitra, P. (Principal Investigator), Bagby, J. W. (Co-Principal Investigator), Puroo, S. (Co-Principal Investigator), Grant, "Empirically Grounded Modeling of Standard-Development Processes," Net Institute, Nonprofit Foundations. Total awarded: \$6,000.00. June 1, 2005 - September 30, 2005.

Chu, C.-H. (Principal Investigator), Liu, P. (Co-Principal Investigator), Kesidis, G. (Co-Principal Investigator), Bagby, J. W. (Co-Principal Investigator), Mitra, P. (Co-Principal Investigator), Lee, D. (Co-Principal Investigator), Grant, "DOD IA Scholarship Program at Penn State/Institutional Capacity Building/with Optional Year 2," U.S. Department of Defense, Federal Agencies. Total awarded: \$201,617.000. August 15, 2004 - August 14, 2005.

ACADEMIC SERVICE

University Service

Committee Member, *University Faculty Senate*. (August 2012 - December 2012).

Committee Member, *Senate Sub-committee on Curricular Affairs*. (August 2012 - December 2012).

Member, *Dean Search Committee* for IST. (2010).

Affiliate Member, *University President's Commission on Racial/Ethnic Diversity*. (June 2008- June 2009).

College Service

Committee Chair, *Faculty Recruiting Committee*. (September 2013 - May 2014).

Chairperson, *IST 220 Course Committee*. (September 2010 - 2013).

Representative of the SIG for Computational Informatics, *Faculty Council*. (September 2013 - August 2014).

Member, *Faculty Recruiting Committee*. (2011 - 2012).

Member, *Travel Committee*. (2011-2012).

Member, *Promotion & Tenure Committee*. (September 2010 - May 2012).

Member, *Diversity Committee*. (2010 - 2013).

Member, *Graduate Recruiting Committee*. (September 2009 - May 2010).

Professional Service

Reviewer, Master of Information and Data Science program proposal, University of California, Berkeley, 2013-2014.

Reviewer, *Computational Technology for Effective Health Care*, from the Computer Science and Telecommunications Board (CSTB) of the National Research Council, 2008.

Panelist, Information & Intelligent Systems Panel, National Science Foundation Panel, 2007-2014.

Review Board, Proceedings of the VLDB Endowment, 2015.

Editorial Board, Journal of Data Mining and Digital Humanities, 2013-current.

Journal Reviews

International Journal of Intelligent Information and Database Systems (IJIIDS), 2014.

PLOS One, 2014.

IEEE Transactions on Knowledge and Data Engineering (TKDE), 2008 - 2014.

ACM Transactions on Intelligent Systems and Technology (TOIST), 2012-2013.

ACM Transactions on Information Systems (TOIS), 2008, 2012.
IEEE Intelligent Systems, 2012.
ETRI Journal, 2012.
Journal on Data Semantics (JDS), 2012.
Journal on Web Semantics (JWS), 2006, 2009, 2012.
Data Mining and Knowledge Discovery (DMKD), 2012.
International Journal of Document Analysis and Recognition (IJ DAR), 2010.
Journal of Intelligent Information Systems (JIIS), 2009.

Conferences and Symposia

Member, *Best Paper Award Committee*, 23rd ACM International Conference on Information and Knowledge Management (CIKM'14). (2014).

Member, *Conference Advisory Board*, 2nd International Conference on Advanced Computing, Networking, and Informatics (ICACNI '14). (2014).

Senior Program Chair for the Track on Recommender Systems in the 28th ACM Symposium On Applied Computing (SAC). (April 2012 - March 2013).

Lead Program Co-Chair, International Conference on Social Intelligence and Technology, SOCIETY 2013, University Park, Pennsylvania, United States. (2012 - 2013).

Area Chair, 23rd ACM International Conference on Information and Knowledge Management (CIKM'14). 2014.

Member, *Senior Program Committee*, Twenty-Sixth AAI Conference on Artificial Intelligence AAI, AI & the Web Track, 2014.

Member, *Senior Program Committee*, Twenty-Fifth AAI Conference on Artificial Intelligence AAI, AI & the Web Track, 2013.

Member, *Senior Program Committee*, Twenty-Third International Joint Conference on Artificial Intelligence (IJCAI), 2013.

Member, Program Committee, The IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 2013, 2015.

Member, Program Committee, Multilingual Multi-document Summarization, 2015.

Member, Program Committee, 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD), 2015.

Member, Program Committee, ACM International Conference on Web Search and Data Mining WSDM (WSDM), 2015.

Member, Program Committee, ACM Multimedia (MM), 2014.

Member, Program Committee, International Conference on Data Mining (ICDM), 2009, 2010, 2014, 2015.

Member, Program Committee, Annual International ACM SIGIR Conference (SIGIR), 2012, 2013.

Member, Program Committee, Conference on Emerging Databases - Technologies, Applications, and Theory (EDBT), 2009, 2012.

Member, Program Committee, IEEE/ASE Social Computing conference, (SocialCom) 2012.

Member, Program Committee, International Semantic Web Conference, (ISWC) 2006, 2007, 2012.

Member, Program Committee, The World Wide Web Conference (WWW) 2008, 2011, 2012.

Member, Program Committee, ACM SIGMOD International Conference on Management of Data (SIGMOD) 2010, Demonstrations, 2011.

Member, Program Committee, AAAI Conference on Artificial Intelligence (AAAI), 2007, 2010.

Member, Program Committee, ACM International Conference on Information and Knowledge Management. (CIKM), 2007, 2010, 2012, 2013.

Member, Program Committee, Twenty-first International Joint Conference on Artificial Intelligence (IJCAI), 2009.

Member, Program Committee, IEEE International Conference on Services Computing (SCC) 2005-2007.

Member, Program Committee, AAAI Fall Symposium on Semantic Web for Collaborative Knowledge Acquisition, 2006.

Workshops

Program Chair, WIDM'12: The 12th International Workshop on Web Information and Data Management in CIKM'12: the 21st ACM International Conference on Information and Knowledge Management. (2012).

Program Co-Chair, WIDM'09: The 11th International Workshop on Web Information and Data Management in CIKM'09: the 18th ACM International Conference on Information and Knowledge Management. (2009).

Program Co-Chair, SNAKDD'09: The 2nd International Workshop on Social Network Mining and Analysis in KDD'08: the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. (2008).

Lead Program Co-Chair, CIMS'07: The 1st Workshop on CyberInfrastructure: Information Management in eScience (CIMS) in CIKM'07: the 16th ACM International Conference on Information and Knowledge Management. (2007).

Member, Program Committee, Second International Workshop on Querying Graph Structured Data (GraphQ), 2013.

Member, Program Committee, Databases and Social Networks (DBSocial), 2010, 2011, 2012.

Member, Program Committee, 2008 AAAI Spring Symposium Series Workshop on Semantic Scientific Knowledge Integration (SSKI), 2008.

Member, Program Committee, Semantic Web and Databases, (SWDB), 2007.

Member, Program Committee, First VLDB International Workshop on Clean Databases (CleanDB), 2006.