

Min Song, Ph.D.

Phone: 02-2123-2416

min.song@yonsei.ac.kr

| University Address | Permanent Address |
|--|--|
| Department of Library and Information Science Yonsei University Email) min.song@yonsei.ac.kr | 29 Westbury Dr. Cherry Hill, NJ, 08003 Tel) 856-489-0960 |

RESEARCH INTERESTS

Text Mining, Information Retrieval, Information Visualization, Knowledge Management, Web Mining, Digital Libraries, and Bioinformatics

EDUCATION

PhD, College of Information Science & Technology, Drexel University, Philadelphia, PA, Spring 2005

Dissertation Title: Robust Knowledge Extraction over Large Text Collections

M.S., School of Library and Information Science, Indiana University, Bloomington, Indiana, Spring 1996

B.A., School of Library and Information Science, Yonsei University, Seoul, Korea, Spring 1992

HONORS AND AWARDS

2011 Outstanding Service Award by ER 2011

2010 NJLA CUS Technology Innovation Award Nominee

2009 NJIT Teaching Award Nominee by IS

2009 Outstanding Service Award by CIKM

2008 NYAS Blavatnik Award Nominee by NJIT

2008 Microsoft Research New Faculty Nominee by NJIT

2005 Beta Phi Mu Minority Doctoral Student Scholarship Award

2005 **Poster Competition Honorable Mention Award** in open division, 2005 GBPA Annual Retreat.

Title: *Biomedical Knowledge Extraction*

2005 **Best Dissertation Award**, Drexel University

2005 Honorable Mentions at 2005 Drexel Research Day, Drexel University

Title: *Robust Knowledge Extraction over Large Text Collections*

2004 **Best Paper award** at IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology

Title: *Extracting and Mining Protein-protein interaction Network from Biomedical Literature*

2004 **Nominated for a Best Paper award** at IEEE/ACM Web Intelligence Conference

Title: *Ontology-based Scalable and Portable Information Extraction System to Extract Biological Knowledge from Huge Collection of Biomedical Web Documents*

2004 **Dean's Award** at 2004 Drexel Research Day, Drexel University

Title: *Applying Mixture Hidden Markov Models to extracting protein-protein interactions in large unstructured natural language text collections*

2003 **Dean's Award** at 2003 Drexel Research Day, Drexel University

Title: *CS-Engine: an intelligent cross search engine for multiple heterogeneous databases*

2001 **Current Kudos**, Institute for Scientific Information

1999 **Current Kudos**, Institute for Scientific Information

1998 **Clayton A. Shepherd Scholarship**, Indiana University

1998 **Summer H. W. Wilson Fellowship**, Indiana University

1996 Fall-1997 Spring **Fellowships of SLIS**, Indiana University

PROFESSIONAL EXPERIENCE

| | |
|---------------------------|--|
| March 2012 - Present | Associate Professor, Department of Library and Information Science, Yonsei University |
| June 2011 – February 2012 | Associate Professor, College of Computing Sciences, New Jersey Institute of Technology |
| June 2006 – May 20011 | Assistant Professor, College of Computing Sciences, New Jersey Institute of Technology |

| | |
|------------------------------|--|
| July 2005 – May 2006 | Adjunct Faculty, Department of Computer & Information Sciences, Temple University |
| January 2005 – May 2006 | Senior Software Engineer, Middleware Server Technologies, Thomson Scientific |
| March 2003 – December 2004 | Software Engineer II, Middleware Server Technologies, Thomson Scientific |
| January 2001 – March 2003 | Software Engineer II, Knowledge Management, Thomson Scientific |
| February 1999 – January 2001 | System Analyst, User Interface Design and Documentation, Thomson Scientific (formally called Institute for Scientific Information) |

TEACHING EXPERIENCE

| | |
|-------------|--|
| Spring 2010 | Enterprise Database Management, NJIT |
| Fall 2009 | Web Mining, NJIT |
| Spring 2009 | Advanced Information Systems, NJIT |
| Fall 2008 | Web Mining, NJIT |
| Spring 2008 | Advanced System Design and Analysis, NJIT |
| Fall 2007 | Web Services and Authentication, NJIT |
| Spring 2007 | Database Management System, NJIT |
| Fall 2006 | System Design and Analysis, NJIT |
| Spring 2006 | Information Retrieval Systems, Drexel University |
| Fall 2005 | Advanced Topics in Database Systems, Temple University |
| Summer 2005 | Object-oriented Programming for Information Systems, Drexel University |
| Summer 2002 | Workshop on Web Services: Concepts and Techniques, Drexel University |

| | |
|-------------|---|
| Spring 2010 | Enterprise Database Management, NJIT |
| Fall 2009 | Web Mining, NJIT |
| Spring 2009 | Advanced Information Systems, NJIT |
| Fall 2008 | Web Mining, NJIT |
| Spring 2008 | Advanced System Design and Analysis, NJIT |
| Fall 2007 | Web Services and Authentication, NJIT |
| Spring 2007 | Database Management System, NJIT |
| Fall 2006 | System Design and Analysis, NJIT |
| Spring 2006 | Information Retrieval Systems, Drexel University |
| Spring 1998 | Computer-based Information Tools (L401), Indiana University |
| Fall 1997 | User Needs: Theories and Practices (L503), Indiana University |

NJIT SERVICES

2008 Spring – 2009 Spring, PhD Program Director of IS Department

PUBLICATIONS

Book

Song, M. and Wu, Y-F. (2008). *Handbook of Research on Text and Web Mining Technologies*, IDEA Group Publishing published in September 2008.

Book Chapter:

1. A Comparative Study of an Unsupervised Word Sense Disambiguation Approach, Xiong, W., Song, M., Watrous-deVersterre, L. Applied Natural Language Processing: Identification, Investigation, Resolution, by IGI Global
2. Song, M. and Prat Tanapaisankit (2011). BioKeySpotter: an Unsupervised Keyphrase Extraction Technique in the Biomedical Full-text Collection, DATA MINING: Foundations and Intelligent Paradigms, in the series 'Intelligent Systems Reference Library'.
3. Song, M. and Song, I.-Y. (2007) *Automatic Query Expansion with Keyphrases and POS Phrase Categorization for Effective Biomedical Text Mining*. In Book: Knowledge Discovery in Bioinformatics: Techniques, Methods, and Applications, 197-208, Wiley Publishing.

4. Manchester, M. and Song, M. (2007) Text Categorization, in: Encyclopedia of Data Warehousing and Mining, 2nd Edition.
5. Maceli, M. and Song, M. (2007) Deep Web Mining through Web Services, in: Encyclopedia of Data Warehousing and Mining, 2nd Edition.
6. Song, M. and Hu, X., and Song, I.-Y. (2005) Information Extraction in Biomedical Literature. in *Encyclopedia of Data Warehousing and Data Mining*.

Journal Papers (Published and Accepted)

1. Said, B., Song, M., Wei, X (2012) Enhancing Biomedical Concept Extraction Using Semantic Relationship Weights, Accepted in *International Journal of Data Mining and Bioinformatics*
1. Zhang, L., Bieber, M., Song, M., Oria, V., and Millard, D. Supplementing Just-in-Time Hypermedia Functionality for Virtual Documents, Accepted in *International Journal on Digital Libraries*.
2. Song, M., Yu, H., and Han, W.S. (2011) A Semi-supervised Learning Technique to Extract Protein Interaction Sentences, *BMC Bioinformatics*
3. Lee, M., Lee, K. H., Song, M., Lee, D. (2011) Building the process-drug-side effect network to discover the relationship between biological Processes and side effects, *BMC Bioinformatics* 2011, 12(Suppl 2):S2 (29 March 2011)
4. Qi, T., Song, M., Yoon, S. J., Watrous-deVersterre, L. (2011) Combining Supervised Learning Techniques to Key-phrase Extraction for Biomedical Full-text, accepted in *International Journal of Intelligent Information Technologies (IJIT)*
5. Rudniy, A., Song, M., Geller, J. (2010) Detecting Duplicate Biological Entities Using Shortest Path Edit Distance, *International Journal of Data Mining and Bioinformatics*, 4(4): 395-410.
6. Ijaz, A.Z., Song, M. and Lee, D. (2010) MKEM: a Multi-level Knowledge Emergence Model for Mining Undiscovered Public Knowledge, *BMC Bioinformatics*, 11(Suppl 2):S3 doi:10.1186/1471-2105-11-S2-S3
7. Song, M. and Rudniy, A. (2010) Detecting Duplicate Biological Entities using Markov Random Field-based Edit Distance, *Knowledge and Information Systems*, 25(3): 371-387, 10.1007/s10115-009-0254-7.
8. Song, M. and Liu, H. (2010) AbbrevExtractor: a Robust Extraction Technique for Biomedical Literature, *International Journal of Functional Informatics and Personalised Medicine*, vol. 3, no. 2.
9. Song, M. Hu, X., Koppel, E. Yoo, I.Y. (2009) A Dynamic and Semantically-aware Technique for Document Clustering in Biomedical Literature, (Accepted)

- International Journal of Data Warehousing and Mining (IJDWM), vol. 5, no. 4.
10. Duan, W., Song, M., Yates, A. (2009). Fast Max-margin Clustering for Unsupervised Word Sense Disambiguation in Biomedical Texts. *BMC Bioinformatics* 2009, 10(Suppl 3):S4doi:10.1186/1471-2105-10-S3-S4
 11. Yoo, I. and Song, M. (2008) Biomedical Ontologies and Text Mining for Biomedicine and Healthcare, *Journal of Computing Science and Engineering (JCSE)*, June, 2(2): 109-136.
 12. Torii, M., Hu, Z., Song, M., Wu, C., Hongfang Liu, H., (2007) A Comparison Study on Algorithms of Detecting Long Form for Short Forms in Biomedical Text, *BMC Bioinformatics*. 2007; 8(Suppl 9): S5.
 13. Song, M, Song, I-Y., Hu, X., Allen R. (2007) "Integration of Association Rules and Ontology for Semantic-based Query Expansion", *Data and Knowledge Engineering*, 63(1): 63-75.
 14. Song, M., Song, I-Y., Hu, X., and Allen, R.B. (2005). KXtractor: an Effective Biomedical Information Extraction Technique Based on Mixture Hidden Markov Models, *LNCS Transactions on Computational Systems Biology*, 2: 68-81.
 15. Hu, X., Lin, T.Y., Song, I-Y., Lin, X., Yoo, I., and Song, M. (2005) A Semi-supervised Efficient Learning Approach to Extract Biological Relationships from Web-based Biomedical Digital Library, *International Journal of Web Intelligence and Agent Systems*.
 16. Song, M., Song, I-Y, and Chen, P.P. (2004). Design and Development of a Cross Search Engine for Multiple Heterogeneous Databases Using UML and Design Patterns, *Information Systems Frontiers: A Journal of Research and Innovation*, 6(1): 77-90.
 17. Song, M. (2000). Visualization in Information Retrieval: A Three-Level Analysis, *Journal of Information Science*, 26(1): 3-19.
 18. Dillon, A. and Song, M. (1997). An Empirical Comparison of the Usability for Novice and Expert Searchers of a Textual and a Graphical Interface to an Art-resource Database, *Journal of Digital Information*, 1, 12-20.
 19. Beebe, C., Nedell, E., Song, M., Sullivan, J., Overfelt, K., Schatz, J., Davenport, E. (1994). New Technology and Publishing: Six Case Studies in Search of a Theory, *Aslib Proceedings*, 46(9): 217-224.

Conference Papers (Published and Accepted)

1. Rudniy, A. Geller, J. and Song, M. "Histogram Difference String Distance for Enhancing Ontology Integration in Bioinformatics," accepted in Int. Conf. on Bioinf. and Comp. Biol., March 2012, Las Vegas, NV.
2. Song, M., Bleik, S., Yu, H., Han, W.S. (2011) Extracting Biomedical Concepts from Fulltext by Relative Importance in a Graph Model, *BHI*
3. Oh, J., Park, S., Yu, H., Song, M. and Park, S.T. (2011). Novel Recommendation based on Personal Popularity Tendency, *The IEEE International Conference on Data Mining series (ICDM)*, 2011.
4. Xu, G., Song, M., and Feng, J. (2011). Research on Multi-agent Comprehensive Evaluation Model Based on Attribute Coordinate, *IEEE International Conference on Granular Computing*. 2011.
5. Tanapaisankit, P. Song, M., Fox, E. A. (2011) Developing a Concept Extraction Technique with Ensemble Pathway, *JCDL 2011*
6. Bleik, S., Wei, X., Wang, Y., Song, M. (2010) Biomedical Concept Extraction using Concept Graphs and Ontology-based Mapping, *BIBM*
7. Lee, S., Lee, D. H., Song, M. (2010) Discovering Biological Processes and Side Effects Relationship Using the Process-drug-side effect Network, *DTMBIO*
8. Rudniy, A., Geller, J., and Song, M. (2010) Shortest Path Edit Distance for Enhancing UMLS Integration and Audit, accepted in *American Medical Informatics Association (AMIA) Annual Symposium*
9. Song, M., Yu, W. J., and Han, W. S. (2010) Combining Active Learning and Semi-supervised Learning Techniques to Extract Protein Interaction Sentences, *BIOKDD 10, in conjunction with KDD 2010*, Washington, D.C. July 25.
10. Rudniy, A., Song, M., and Geller, J. (2010) Shortest Path Edit Distance for Detecting Duplicate Biological Entities. *ACM International Conference on Bioinformatics and Computational Biology*
11. Ho, S. M., Song, M., Bieber, M., (2010) IntegraL: The Effectiveness of a Link-based Federated Search Infrastructure, *iConference 2010*.
12. Xiong, W., Song, M., Watrous-deversterre, L. (2010) A Quantitative Assessment of SENSATIONAL with an Exploration of Its Applications, *the 23rd Florida Artificial Intelligence Research Society Conference (FLAIRS 2010)*.
13. Song, M. and Murillo, J. (2009) An Outcome Discovery System to Determine Mortality Factors in Primary Care Facilities, *DTMBIO 09*.
14. Song, M., Ruecker, S., and Youngman, P. (2009) Indications of Emotional Connection: Epistolary Text Mining for Intimate Language, *BookOnline 09*, October 2, Corfu, Greece

15. Song, M., Ho, S.M., Bieber, M., Koppel, E., Hamidullah, V., and Bokota, P. (2009) A Scalable Digital Library Infrastructure Expands Search and Beyond, *BookOnline 09*, October 2, Corfu, Greece
16. Qi, T., Zhang, Y., and Song, M. (2009). Text Mining for Bioinformatics: State of the Art Review, 2009 *the 2nd IEEE International Conference on Computer Science and Information Technology (ICCSIT 2009)*.
17. Qi, T., Yagci, A., and Song, M. (2009). Extraction of Key-phrases from Biomedical Full-text with Supervised Learning Techniques, *15th Americas Conference on Information Systems*.
18. Bleik, S., Smalter, A., Song, M., Huan, J. and Lushington, G. (2009) CGM: A Biomedical Text Categorization Approach Using Concept Graph Mining, *Workshop of Applications of Machine Learning in Bioinformatics in conjunction with BIBM 09*.
19. Song, M. and Bieber, M. (2008). IntegraL: Lightweight Link-based Integration of Heterogeneous Digital Library Collections and Services in the Deep Web, *The Semantic Web meets the Deep Web workshop, in conjunction with IEEE Joint Conference on E-Commerce Technology and Enterprise Computing, E-Commerce and E-Services 2008*. 20-26
20. Song, M., Hu, X., Koppel, E. Yoo, I.Y. Document Clustering by Semantic Smoothing and Dynamic Growing Cell Structure (DynGCS) for Biomedical Literature, *10th International Conference on Data Warehousing and Knowledge Discovery (DaWaK 2008)*, Turin, Italy, 217-226
21. Song, M. and Rudniy, A. (2008) Markov Random Field-based Edit Distance for Entity Matching, Biomedical Literature, *International Conference on Bioinformatics and Biomedicine*, 2008, 457-460.
22. Song, M. and Chang, P. (2008) Automatic Extraction of Abbreviation for Emergency Management Websites, *International Community on Information Systems for Crisis Response and Management*, Washington, DC, USA, May 4-7, 2008.
23. Song, M. and Yoo, I.Y. (2007) A Hybrid Abbreviation Extraction Technique for Biomedical Literature, *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. November 2-4, 2007.
24. Song, M., Song, I-Y, Lee, K.J., and Bieber M. (2006) Automatic Extraction of Creating a Lexical Repository of Abbreviations in the Biomedical Literature, *8th International Conference on Data Warehousing and Knowledge Discovery (DaWaK 2006)*.
25. Song, M., Song, I.-Y., Allen, R., and Obradovic, Z. (2006) Keyphrase Extraction-

- Based Query Expansion in Digital Libraries, *Joint Conference on Digital Libraries*, Chapel Hill, NC, June 11-15.
26. Zhu, W., Song, M., and Allen R. (2005). TREC 2005 Enterprise Track Results from Drexel, *14th Text REtrieval Conference (TREC 2005)*, National Institute of Standards and Technology (NIST). November 15-18, Gaithersburg, Maryland USA, 415-420.
 27. Hu X., Yoo I., Song M., Y. Zhang, I-Y Song, Mining Undiscovered Public Knowledge from Complementary and Non-interactive Biomedical Literature through Semantic Pruning, *ACM Fourteen Conference on Information and Knowledge Management (ACM CIKM 2005)*, 31st October - 5th November, 2005 Bremen, Germany.
 28. Song, M., Song, I-Y., Lee, K.J. (2005) Automatic Unsupervised Keyphrase-based Query Expansion for Biomedical Domain, International Conference on Knowledge Management, Westin Charlotte, North Carolina on October 27-28.
 29. Song, M, Song, I-Y., Hu, X., Allen R. (2005) Semantic Query Expansion Combining Association Rules with Ontologies and Information Retrieval Techniques, *Data Warehousing and Knowledge Discovery (DaWak 2005)*.
 30. Song, M., Song, I.-Y., Hu, X., and Allen, R.B. (2005). Integrating Text Chunking with Mixture Hidden Markov Models for Effective Biomedical Information Extraction, *International Workshop on Bioinformatics Research and Applications*, Emory University, Atlanta, USA, May 22-25.
 31. Song, M., Song, I.-Y., Hu, X., and Allen, R.B. (2005). DocSpotter: an Automatic Unsupervised Querying Algorithm for Efficient Information Extraction in Biomedical Domain, *9th Pacific Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2005)*, Vietnam, May 18-20.
 32. Hu, X., Yoo, I., Song, I.-Y., Song, M, Han, J., and Lechner, M. (2004). Extracting and Mining Protein-protein interaction Network from Biomedical Literature, *2004 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology*. (**Best Paper award**).
 33. Hu X., Lin T.Y, Song I-Y. Lin X. Yoo, I., Lechner M., and Song, M. (2004). *Ontology-based Scalable and Portable Information Extraction System to Extract Biological Knowledge from Huge Collection of Biomedical Web Documents*, 2004 IEEE/ACM Web Intelligence Conference, 77-83. (**Nominated for a Best Paper award**).
 34. Song, M, Song, I.-Y., and Hu, X. (2004). Designing and Developing Flexible Keyphrase Extraction System with Unified Modeling Language (UML), *ASIS&T*

- 2004 Annual Meeting; "Managing and Enhancing Information: Cultures and Conflicts, Providence, Rhode Island November 13 – 18, 367-372.
35. Song, M., Song, I.-Y., and Hu, X. (2004). An Efficient Keyphrase Extraction System Using Data Mining and Natural Language Processing Techniques, *First International Workshop on Semantic Web Mining and Reasoning (SWMR 2004) In conjunction with the 2004 IEEE/WCI/ACM International Conference on Web Intelligence*. Sept. 20-24, 2004, Beijing, China.
 36. Song, M, Song, I.-Y., and Hu, X. (2003). KPSpotter: A Flexible Information Gain-based Keyphrase Extraction System, *Fifth International Workshop on Web Information and Data Management (WIDM'03), In Conjunction with the 12th International Conference on Information and Knowledge Management (CIKM 2003)*, November 7-8, 2003.
 37. Song, M. and Song, I.-Y. (2003). Developing a Cross Search Engine for Multiple Heterogeneous Databases, *2003 Information Resources Management Association (IRMA) International Conference*, May 18-21, 2003, Philadelphia, USA.
 38. Song, M. and Galardi, P. (2001). Semantic Relationships between Highly Cited Articles and Citing Articles in Information Retrieval, *Proceedings of the 64th ASIST Annual Meeting*, 38: 171-181.
 39. Song, M. and Song, I.-Y. (2001). A Case Study for Web-based Data Delivery E-commerce Design, *Informing Science Conference 01*, June 19-22 Crakow, Poland.
 40. Song, M. (1998). BiblioMapper: A Cluster-based Information Visualization Technique. *IEEE Symposium on Information Visualization (InfoVis 1998), Sponsored by the IEEE Computer Society Technical Committee on Computer Graphics*, October 19-20, 1998 Research Triangle Park, NC, 130-136.
 41. Song, M. (1998). Cluster-based and Association-based Visualization Systems as Information Exploration Tools, In: *the 9th ASIS SIG/CR Classification Research Workshop: An interdisciplinary meeting*. Pittsburgh, PA, October 25.
 42. Song, M (1998). Can Visualizing Document Space Improve User's Information Foraging?, *1998 Annual Meeting of American Society for Information Science*, Pittsburgh, PA October 26-29, 118-129.
 43. Song, M and Gillespie, T. (1997). BiblioMapper: Visualizing a Database of Technical Reports. *Proceedings of HCI International '97: 7th International Conference on Human-Computer Interaction jointly with 13th Symposium on Human Interface (Japan)*. August 24-29, San Francisco, California, 87.
 44. Song, M. (1997). Visualization for Complex Large-Scale Systems: Issues and Research Agenda, In: *the proceedings of AIS 97: Training and Use of Emerging*

Technologies for Information System Track, August 15-17, Indianapolis, IN, 543-546.

Presentations

1. Ho, S.M., Song, M., Bieber, M. (2010). IntegraL connects data, users and scientists. Proceedings of 36th Annual Conference of International Association for Social Science Information Services & Technology (IASSIST 2010), June 1-4, 2010.
2. Song, M. (2008) A Markov Random Field-based Edit Distance (MRFED) Algorithm for Gene Synonym Matching, Emerging Information and Technology – Bioinformatics (EITC-Bio) 2008, Princeton, NJ.
3. Song, M, Bieber, M., Kopel, E., Lori. (2008) IntegraL: Lightweight Link-based Integration of Heterogeneous Digital Library Collections, NSDL, Washington, D.C., Oct.
4. Song, M (2005) Robust Knowledge Extraction over Large Biomedical Text collections, *Greater Philadelphia Bioinformatics Alliance*, October 14, 2005, Penn State Great Valley, Malvern, PA. (**Poster Competition Honorable Mention award in Open Division.**)
5. Song, M. and Chen, C. (2003). Investigation of Power Law Distribution of Citation Data in the field of Medicine, 9th International Conference Scientometrics and Informetrics, China
6. Song, M. and Lin, X. (2002), Interactive Visualization System with Multidimensional Scaling, *American Society for Information Science and Technology ASIST 2002 Annual Meeting November 18-21, 2002 * Wyndham Franklin Plaza Hotel * Philadelphia, PA*, 39: 562-563.
7. Song, M. and Song, I.-Y. (2002). Developing a Cross-Search Component for a Web Information Retrieval System Using UML and Design Patterns, *1st International Semantic Web Conference (ISWC2002) June 9-12th, 2002 Sardinia, Italy* .
8. Song, M. (1997). Can Visualizing Document Space Improve User's Information Foraging?: Research in Progress. *Connections 97: The Third Great Lakes Information Science Conference*, October 24-25, pp. 17 1997

Non-refereed journal

1. Song, M. and Obradovic, Z. (2007) Report on ACM Text Mining in Bioinformatics (TMBIO 2006), *SIGIR Forum*, vol 41, issue 1.

Papers Submitted for Publication

Journal

2. Song, M. and Song, I.-Y. TAKES: Two-step Approach for Knowledge Extraction in Biomedical Digital Libraries, *Journal of Information Science*
3. A. Rudniy, J. Geller, M. Song, "Longest Approximately Common Prefix method for duplicate detection". Submitted to *Int. J. of Data Mining and Bioinformatics*.
4. Yang, C., Tang, X., and Song, M. A Content and Co-authorship Network Analysis to Understand the Development of Scientific Literature in Different Domains, submitted to *Information Processing and Management*

Conference

1. Smalter, A., Bleik, S., Song, M., Huan, J. Biomedical Text Categorization with Concept Graph Representations Using A Controlled Vocabulary,

In Preparation for Journal

1. Ho, S.M., Song, M., Bieber, M. and Zhang, X. (2010). Shaping User's Information Seeking Behavior: A Link-based Federated Search Infrastructure. Intended for *Information Sciences*.
2. Choi, J. J., Lee, D. H., and Song, M. (2010) A Topological Pattern Comparison Algorithm of Multi-type Interaction Networks for Undiscovered Public Knowledge Inference, *SDM*

GRANT EXPERIENCE

Awarded NSF Grants

1. PI - ER 2011: Fostering Interdisciplinary Research and Education in Software Engineering, Amount: \$12,000, to run through 08/12
2. PI - Title: QIC: Query in Context for Educational Collections, Amount: \$149,985.00, to run through 09/11.
3. PI - Title: QuEST: Revolutionizing Repository Cross-referencing to Transform Discovery of Teaching Materials, Amount: \$149,974.00, to run through 09/12
4. Co-PI - Title: Integral: Expanding the NSDL Reach to Traditional Libraries, Amount: \$849,997, to run through 12/09.

Awarded Institute of Museum and Library Services (IMLS) Grant

1. Co-PI - Title: "IntLib: Integrating Library Collections and Services," Amount: \$500,000, to run through 12/09.

NSF Panels

1. NSF Panel for CRI – 2007
2. NSF Panel for IIS Core Program – 2008
3. NSF Panel for CISE Career Program – 2009
4. NSF Panel for HCC Core Program - 2010

CONFERENCE PROGRAM COMMITTEE SERVICE

Organizing Conference

1. Track Co-Chairs: Harnessing Social Media for Business Intelligence, Mini-track in BI, AMCIS 2011.
2. Co-travel award chair: The Nineteenth Conference on Information and Knowledge Management, Toronto, Canada, 2010
3. Co-program chair: Fourth International Workshop in Data and Text Mining in Bioinformatics (DTMBIO 2010), in Toronto, Canada, 2010
4. Workshop chair: The Eighteenth Conference on Information and Knowledge Management, in Hong Kong, China, 2009
5. Co-program chair: Third International Workshop in Data and Text Mining in Bioinformatics (DTMBIO 2009), in Hong Kong, China, 2009
6. Co-chair: Second International Workshop in Data and Text Mining in Bioinformatics (DTMBIO 2008), in Napa Valley, CA, USA
7. Co-chair: International Workshop on Biomedical and Health Informatics in conjunction with the IEEE International Conference on Bioinformatics and Biomedicine 2008 (BIBM 2008), Philadelphia, PA, USA, November 5-7, 2008
8. Registration chair: IEEE International Conference on Bioinformatics and Biomedicine 2008 (BIBM 2008), Philadelphia, PA, USA, November 5-7, 2008
9. Co-chair: International Workshop on Knowledge Discovery and Management in Health Informatics in conjunction with the IEEE International Conference on Bioinformatics and Biomedicine, November 2-4 2007, Silicon Valley, USA
10. Local chair of 2007 IEEE/INFORMS International Conference on Service Operations and Logistics, and Informatics (SOLI)

11. Co-chair of International Workshop in Text Mining in Bioinformatics (TMBIO 2006), in Washington D.C.

Conference Program Committee

1. The Nineteenth Conference on Information and Knowledge Management, Toronto, Canada, 2010
2. ICDM International Workshop on Mining Multiple Information Sources (MMIS)
3. Workshop on Language Technology applied to biomedical and health documents (BioSEPLN 2010)
4. The Joint Conference of 23rd International Conference on the Computer Processing of Oriental Languages (ICCPOL 2010)
5. The 2010 IEEE International Conference on Granular Computing
6. The International workshop on Data Mining in Bioinformatics (BIOKDD'10)
7. The 10th IEEE International Conference on Data Mining (ICDM 2010)
8. The 9th IEEE International Conference on Data Mining (ICDM 2009)
9. The 2009 IEEE International Conference on Granular Computing
10. The 7th ACS/IEEE International Conference on Computer Systems and Applications (AICCSA-2009)
11. The 2008 IEEE International Symposium on Data Mining and Information Retrieval (IEEE DMIR-08)
12. 10th International Conference on Data Warehousing and Knowledge Discovery (DaWaK 2008) Turin, Italy, September 2-5, 2008
13. IEEE ICDM Workshop on Mining and Management of Biological Data in conjunction with IEEE International Conference of Data Mining (ICDM 07), October 28, 2007, Omaha, NE, USA.
14. The 2007 IEEE International Symposium on Data Mining and Information Retrieval (IEEE DMIR-07) in conjunction with the IEEE 21st International Conference on Advanced Information Networking and Applications (AINA-07), May 21-23, 2007 Sheraton Fallsview Hotel, Niagara Falls, Ontario, Canada
15. 2007 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), November 2-4, 2007, San Jose, California, USA
16. GrC 2006 IEEE International Conference On Granular Computing, Georgia State University, Atlanta, USA, May 10-12, 2006, Sponsored by the IEEE Computational Intelligence Society
17. GrC 2005 IEEE International Conference On Granular Computing, Beijing, China, July 25-27, 2005.

18. The 2004 International Conference on Computer and Information Technology
Wuhan, China, 14-16 September 2004.
19. First International Workshop on Semantic Web Mining and Reasoning (SWMR
2004), in conjunction with the 2004 IEEE/WCI/ACM International Conference on
Web Intelligence.

**EXPTERNAL REVIEWER FOR JOURNALS, INTERNATIONAL
CONFERENCES AND FUNDING AGENCY**

1. 2010 Science & Engineering Research Council (SERC), part of the Agency for
Science, Research & Technology (A*STAR), Science & Engineering Research
Council (SERC) in Singapore.
2. The 14th World Multi-Conference on Systemics, Cybernetics and Informatics:
WMSCI 2010
3. Journal of Advances in Bioinformatics, 2010
4. The Eighteenth Conference on Information and Knowledge Management, 2009
5. The 13th World Multi-Conference on Systemics, Cybernetics and Informatics:
WMSCI 2009
6. 2007 Science & Engineering Research Council (SERC), part of the Agency for
Science, Research & Technology (A*STAR), Science & Engineering Research
Council (SERC) in Singapore.
7. The 24rd International Conference on Conceptual Modeling (ER2005)
8. The Tenth International Conference on Rough Sets, Fuzzy Sets, Data Mining, and
Granular Computing (RSFDGrC 2005)
9. The 7th Data Warehousing and Knowledge Discovery (DAWAK) 2005
10. The First International Workshop on Digital Government: Systems and
Technologies (DGOV) 2004
11. The 16th International Conference o Advanced Information Systems Engineering
(CaiSE) 2004
12. The 23rd International Conference on Conceptual Modeling (ER2004)
13. The 4th International Conference on Computer and Information Technology (CIT)
2004
14. The First International Workshop on Semantic Web Mining and Reasoning
(SWMR) 2004
15. The 6th Data Warehousing and Knowledge Discovery (DAWAK) 2004
16. 4th International Conference on Web Information Systems Engineering (WISE)
2003

17. 22nd International Conference on Conceptual Modeling (ER 2003)
18. National Science Foundation Information Technology Research (NSF ITR) 2004

INVITED TALK

1. Title: Adaptive Text Mining for Large Unstructured Text Collection, 12/21/2006, at Georgetown University. I was invited by Department of Biostatistics, Bioinformatics, and Biomathematics at Georgetown University Medical Center.
2. Title: Sensational: An Unsupervised Word Sense Disambiguation Technique for Information Retrieval, 05/26/09, at Seoul National University
3. Title: Sensational: An Unsupervised Word Sense Disambiguation Technique for Information Retrieval, 05/29/09, at Yonsei University
4. Title: Sensational: An Unsupervised Word Sense Disambiguation Technique for Information Retrieval, 06/01/09, at KyungPook National University
5. Title: Sensational: An Unsupervised Word Sense Disambiguation Technique for Information Retrieval, 06/18/09, at Korea University
6. Title: Query In Context (QIC): Incorporating Context into Search, 04/01/2010, at Drexel University

PPROFESSIONAL MEMBERSHIPS

1. American Society for Information Science & Technology (ASIST)
2. Association for Computing Machinery (ACM)
3. Institute of Electrical and Electronics Engineers (IEEE)

EDITORIAL ACTIVITIES

1. Editorial board member - International Journal of Data Mining and Bioinformatics
2. Editorial board member - Online Journal of Bioinformatics
3. Editorial board member - International Journal of Computational Models and Algorithms in Medicine
4. Editorial board member - Journal of Computing and Applications
5. Co-guest editor of a special issue on Text Mining in Bioinformatics – BMC Bioinformatics, 2007
6. Co-guest editor of a special issue on BMC Medical Informatics and Decision Making, 2009

RESEARCH SUPERVISION

Postdoctoral

Li Zhang (CS, NJIT) – 2006 Spring to 2007 December

Research Topic: Hypertext System

Current Affiliation: Associate Scientist, GE Healthcare

Shuyuan Mary Ho (Information Studies, Syracuse University) – 2009 May to 2009 December

Research Topic: User Information Seeking Behavior

Current Affiliation: Assistant Professor of Goodwin College, Drexel Univ.

Ph.D. Thesis Advisee

Alex Rudniy (Co-advisor: James Geller, CS) – 2009 Spring to Present

Research Topic: Duplicate Record Detection

Expected Graduation Date: 2011 Spring

Todd Will (Co-advisor: Brook Wu, IS) – 2010 Spring to Present

Research Topic: Recommendation System

Expected Graduation Date: 2011 Summer

Said Bleik (IS) – 2009 Spring to Present

Research Topic: Biomedical Text/Graph Mining

Expected Graduation Date: 2011 Fall

Lori Watrous-Deversterre (IS) – 2009 Spring to Present

Research Topic: Educational Text Mining

Expected Graduation Date: 2012 Spring

Current Affiliation: University Lecturer of IT program, NJIT

Ph.D. Thesis Committee

Chris Leberknight (IS) 2008 Spring – 2009 Fall

Research Topic: Information Security

Current Affiliation: Postdoctoral Research Fellow at Princeton Univ.

Umar Qasim (IS) 2009 Spring – Present

Research Topic: Recommendation System

Expected Graduation Date: 2011 Spring

Nkechi Nnadi (IS) 2008 Spring – Present
Research Topic: Recommendation System
Expected Graduation Date: 2011 Spring

Ritu Khare (iSchool, Drexel University) 2009 Spring - Present
Research Topic: Web Information Extraction
Expected Graduation Date: 2011 Spring

Ph.D. Research Project

Wei Xiong (IS) – 2008 Fall to Present – Project Title: Word Sense Disambiguation
Yiran Wang (IS) – 2009 Fall to Present – Project Title: Social Network Analysis

Master Research Project

Vahid Hamidullah (CS) – 2010 Spring to Present - Project Title: QIC
Shamshuzia Syed (IS) – 2010 Spring to Present - Project Title: QIC

Undergraduate Research Project

Eric Koppel (CS) – 2007 Spring to 2009 Fall - Project Title: IntegraL
Pawel Bokota (CS) – 2008 Fall to 2009 Fall - Project Title: IntegraL