

Min Song, Ph.D.

Phone: 973-596-5291

min.song@njit.edu

University Address	Permanent Address
Department of Information Systems New Jersey Institute of Technology Email) min.song@njit.edu	29 Westbury Dr. Cherry Hill, NJ, 08003 Tel) 856-489-0960

RESEARCH INTERESTS

Text Mining, Information Retrieval, Information Visualization, Knowledge Management, Data Mining, and Digital Libraries

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

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

June 2006 - Present	Assistant professor, College of Computing Sciences, New Jersey Institute of Technology
August 2005 – May 2006	Adjunct Faculty, Department of Computer & Information Sciences, Temple University
May 2005 – Present	Adjunct Faculty, School of Information Science & Technology, Drexel University
January 2005 – Present	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

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 1998	Computer-based Information Tools (L401), Indiana University
Fall 1997	User Needs: Theories and Practices (L503), Indiana University

PUBLICATIONS

Book

Co editor of book, *Handbook of Research on Text and Web Mining Technologies*, IDEA Group Publishing expected to be published in 2008.

Book Chapter:

1. 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.
2. Manchester, M. and Song, M. (2007) Text Categorization, in: *Encyclopedia of Data Warehousing and Mining*, 2nd Edition.
3. Maceli, M. and Song, M. (2007) Deep Web Mining through Web Services, in: *Encyclopedia of Data Warehousing and Mining*, 2nd Edition.
4. Song, M. and Hu, X., and Song, I.-Y. (2005) Information Extraction in Biomedical Literature. To be published in: *Encyclopedia of Data Warehousing and Data Mining*.

Journal Papers (Published and Accepted)

1. 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*, Winter.
2. Song, M. and Obradovic, Z. (2007) Report on ACM Text Mining in Bioinformatics (TMBIO 2006), *SIGIR Forum*, vol 41, issue 1.
3. Song, M, Song, I-Y., Hu, X., Allen R. (2007) “Integration of Association Rules and Ontology for Semantic-based Query Expansion”, Accepted in *Data and Knowledge Engineering*, 63(1): 63-75.
4. 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.
5. 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*.
6. 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.
7. Song, M. (2000). Visualization in Information Retrieval: A Three-Level Analysis, *Journal of Information Science*, 26(1): 3-19.
8. 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.
9. 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. 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.
2. 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).

3. 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.
4. 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.
5. 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.
6. 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.
7. 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)*.
8. 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.
9. 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.
10. 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)*.
11. 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)**.

12. 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.
13. 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.
14. 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.
15. 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.
16. 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.
17. Song, M. and Song, I.-Y. (2001). A Case Study for Web-based Data Delivery E-commerce Design, accepted to *Informing Science Conference 01*, June 19-22 Crakow, Poland.
18. 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.
19. 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.
20. 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.
21. Song, M and Gillespie, T. (1997). BiblioMapper: Visualizing a Database of Technical Reports. *Poster Session: Abridged Proceedings In: 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.

22. 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. 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.)**

2. 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

3. 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.

4. 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* .

5. 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

Papers Submitted for Publication

Journal

1. Song, M. Robust Knowledge Extraction for Biomedical Digital Libraries, *IEEE/ACM Transactions on Computational Biology and Bioinformatics*.

2. Yoo, I. and Song, M. A Comprehensive Comparison Study of Document Clustering for a Biomedical Digital Library MEDLINE, *Information Retrieval*, Springer.

Conference

1. Song, M., Hu, X., Wu, Y. Document Clustering by Semantic Smoothing and Dynamic Growing Cell Structure (DynGCS) for Biomedical Literature, *SIAM International Conference on Data Mining*, Atlanta, George.

2. Quentine, J., Moldovan, M., Galego, B., Song, M. Describing the Social Geography of Synchronous Interaction Environments, *CHI Tech Note*, 2008

GRANT EXPERIENCE

Awarded NSF Grants

Title: Integral: Expanding the NSDL Reach to Traditional Libraries, Amount: \$849,997, to run through 12/08.

Awarded Institute of Museum and Library Services (IMLS) Grant

Title: "IntLib: Integrating Library Collections and Services," Amount: \$500,000, to run through 12/08.

- NSF Panel for CRI – 2007 March 23.

CONFERENCE PROGRAM COMMITTEE SERVICE

Organizing Conference

1. 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
2. Local chair of 2007 IEEE/INFORMS International Conference on Service Operations and Logistics, and Informatics (SOLI)
3. Co-chair of International Workshop in Text Mining in Bioinformatics (TMBIO 2006), in Washington D.C.

Conference Program Committee

1. The 2007 IEEE International Symposium on Data Mining and Information Retrieval (IEEE DMIR-07)
2. 9th International Conference on Data Warehousing and Knowledge Discovery (DaWaK 2007) Regensburg, Germany, 3-7 September 2007
3. 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.
4. 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

5. 2007 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), November 2-4, 2007, San Jose, California, USA

6. GrC 2006 IEEE International Conference On Granular Computing, Georgia State University, Atlanta, USA, May 10-12, 2006, Sponsored by [the IEEE Computational Intelligence Society](#)

7. GrC 2005 IEEE International Conference On Granular Computing, Beijing, China, July 25-27, 2005.

8. The 2004 International Conference on Computer and Information Technology Wuhan, China, 14-16 September 2004.

9. 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 INTERNATIONAL CONFERENCES AND FUNDING AGENCY

1. 2007 Science & Engineering Research Council (SERC), part of the Agency for Science, Research & Technology (A*STAR), Science & Engineering Research Council (SERC) in Singapore.

2. The 24rd International Conference on Conceptual Modeling (ER2005)

3. The Tenth International Conference on Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing (RSFDGrC 2005)

4. The 7th Data Warehousing and Knowledge Discovery (DAWAK) 2005

5. The First International Workshop on Digital Government: Systems and Technologies (DGOV) 2004

6. The 16th International Conference o Advanced Information Systems Engineering (CaiSE) 2004

7. The 23rd International Conference on Conceptual Modeling (ER2004)

8. The 4th International Conference on Computer and Information Technology (CIT) 2004

9. The First International Workshop on Semantic Web Mining and Reasoning (SWMR) 2004

10. The 6th Data Warehousing and Knowledge Discovery (DAWAK) 2004

11. 4th International Conference on Web Information Systems Engineering (WISE) 2003
12. 22nd International Conference on Conceptual Modeling (ER 2003)
13. National Science Foundation Information Technology Research (NSF ITR) 2004

INVITED TALK

Adaptive Text Mining for Large Unstructured Text Collection on Dec. 21 (Thursday), 2006 in PIR library at George Town University. I was invited by Department of Biostatistics, Bioinformatics, and Biomathematics at Georgetown University Medical Center.

PROFESSIONAL MEMBERSHIPS

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

EDITORIAL ACTIVITIES

1. Editorial board member -- Online Journal of Bioinformatics
2. Co guest editor of a special issue on Text Mining in Bioinformatics – BMC Bioinformatics, 2007