

CiteSeer^X Related Publications

Links to papers are currently available on Professors Giles [homepage](#).

Journal Papers:

1. Madian Khabsa, C. Lee Giles, "Chemical entity extraction using CRF and an ensemble of extractors. *J. Cheminformatics* 7(S-1): S12, 2015.
2. Madian Khabsa, C. Lee Giles, "The Number of Scholarly Documents on the Public Web," *PLoS ONE*, May, 2014.
3. Dayu Yuan, Prasenjit Mitra, C. Lee Giles, "Mining and Indexing Graphs for Supergraph Search," *Proceedings of the VLDB Endowment (PVLDB)* 6(10): 829-840, 2013.
4. Sumit Bhatia, Cornelia Caragea, Hung-Hsuan Chen, Jian Wu, Pucktada Treeratpituk, Zhaohui Wu, Madian Khabsa, Prasenjit Mitra, C. Lee Giles, "Specialized Research Datasets in the CiteSeerx Digital Library," *D-Lib Magazine* 18(7/8), 2012.
5. Bingjun Sun, Prasenjit Mitra, C. Lee Giles, Karl T. Mueller: Identifying, Indexing, and Ranking Chemical Formulae and Chemical Names in Digital Documents. *ACM Trans. Inf. Syst.* 29(2): 12, 2011.
6. Xiaonan Lu, Saurabh Kataria, William J. Brouwer, James Ze Wang, Prasenjit Mitra, C. Lee Giles, "Automated analysis of images in documents for intelligent document search," *IJDAR* 12(2): 65-81, 2009.
7. Umer Farooq, Craig H. Ganoe, John M. Carroll, C. Lee Giles, "Designing for e-science: Requirements gathering for collaboration in CiteSeer," *Int. J. Human Computer Studies* 67(4): 297-312, 2009.
8. Umer Farooq, Craig H. Ganoe, John M. Carroll, Isaac G. Council, C. Lee Giles, "Design and evaluation of awareness mechanisms in CiteSeer," *Inf. Process. Manage.* 44(2): 596-612, 2008.
9. Huajing Li, Wang-Chien Lee, Anand Sivasubramaniam, C. Lee Giles, "Workload analysis for scientific literature digital libraries," *Int. J. on Digital Libraries* 9(2): 139-149, 2008.
10. Umer Farooq, Yang Song, John M. Carroll, C. Lee Giles, "Social Bookmarking for Scholarly Digital Libraries." *IEEE Internet Computing* 11(6): 29-35, 2007.
11. C.L. Giles, I.G. Council, "Who gets acknowledged: Measuring scientific contributions through automatic acknowledgement indexing," *Proceedings of the National Academy of Sciences*, 101, 17599-17604, 2004.
12. S. Lawrence, D.M. Pennock, G.W. Flake, R. Krovetz, F. Coetzee, E.J. Glover, F. Nielsen, Andries Kruger, C.L. Giles, "Persistence of Web References in Scientific Research," *IEEE Computer*, 34(2), 26-31, 2001.
13. A. Goodrum, K.W. McCain, S. Lawrence, C.L. Giles, "Scholarly publishing in the Internet age: a citation analysis of computer science literature," *Information Processing and Management*, 37(5), 661-675, 2001.
14. D. Pennock, S. Lawrence, C.L. Giles, F. Nielsen, "The Real Power of Artificial Markets", *Science*, 291(5506), 987-988, 2001.
15. K. Bollacker, S. Lawrence, C. L. Giles, "Discovering Personally Relevant Scientific Literature on the Web," *IEEE Intelligent Systems*, 15(2), p. 42, 2000.
16. S. Lawrence, C.L. Giles, "Accessibility of Information on the Web," *Nature*, 400, 107-109, 1999.
17. S. Lawrence, C.L. Giles, "Searching the Web: General and Scientific Information Access," *IEEE Communications*, Special issue on "Digital Library Using Next Generation Internet," 37(1), p. 116, 1999.
18. S. Lawrence, C.L. Giles, K. Bollacker, "Digital Libraries and Autonomous Citation Indexing," *IEEE Computer*, 32(6), 67-71, 1999.

Conference Papers:

19. Jian Wu, Kyle Williams, Hung-Hsuan Chen, Madian Khabsa, Cornelia Caragea, Alexander Ororbia, Douglas Jordan, C. Lee Giles, "CiteSeerX: AI in a Digital Library Search Engine,"

- Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence, Innovative Applications of Artificial Intelligence, 2930-2937, 2014.
20. Kyle Williams, Hung-Hsuan Chen, C. Lee Giles, "Supervised Ranking for Plagiarism Source Retrieval," Working Notes for CLEF 2014 Conference, 1021-1026, 2014.
 21. Kyle Williams, Hung-Hsuan Chen, C. Lee Giles, "Classifying and ranking search engine results as potential sources of plagiarism," ACM Symposium on Document Engineering (DocEng), 97-106, 2014.
 22. Cornelia Caragea, Jian Wu, Alina Maria Ciobanu, Kyle Williams, Juan Pablo Fernández Ramírez, Hung-Hsuan Chen, Zhaohui Wu, C. Lee Giles, "CiteSeer x : A Scholarly Big Dataset," Advances in Information Retrieval - 36th European Conference on IR Research (ECIR), 311-322, 2014.
 23. Jian Wu, Alexander Ororbia, Kyle Williams, Madian Khabsa, Zhaohui Wu, C. Lee Giles, "Utility-Based Control Feedback in a Digital Library Search Engine: Cases in CiteSeerX," 9th USENIX International Workshop on Feedback Computing, 2014.
 24. Jian Wu, Pradeep B. Teregowda, Kyle Williams, Madian Khabsa, Douglas Jordan, Eric Treece, Zhaohui Wu, C. Lee Giles, "Migrating a Digital Library to a Private Cloud," 2014 IEEE International Conference on Cloud Engineering (IC2E) 97-106, 2014.
 25. Kyle Williams, Jian Wu, Sagnik Ray Choudhury, Madian Khabsa, C. Lee Giles, "Scholarly big data information extraction and integration in the CiteSeerX digital library," Proceedings of the 30th International Conference on Data Engineering (ICDE) Workshops: IIWeb 2014 — 10th International Workshop on Information Integration on the Web, 68-73, 2014.
 26. Kyle Williams, Lichi Li, Madian Khabsa, Jian Wu, Patrick C. Shih, C. Lee Giles, "A Web Service for Scholarly Big Data Information Extraction," 2014 IEEE International Conference on Web Services (ICWS 2014), 105-112, 2014.
 27. Zhaohui Wu, Jian Wu, Madian Khabsa, Kyle Williams, Hung-Hsuan Chen, Wenyi Huang, Suppawong Tuarob, Sagnik Ray Choudhury, Alexander Ororbia, Prasenjit Mitra, C. Lee Giles, "Towards building a scholarly big data platform: Challenges, lessons and opportunities," IEEE/ACM Joint Conference on Digital Libraries (JCDL 2014), 117-126, 2014.
 28. Zhaohui Wu, Wenyi Huang, Liang Chen, C. Lee Giles, "Crowd-sourcing Web knowledge for metadata extraction," IEEE/ACM Joint Conference on Digital Libraries (JCDL 2014), 141-144, 2014
 29. Hung-Hsuan Chen, Madian Khabsa, C. Lee Giles, "The feasibility of investing in manual correction of metadata for a large-scale digital library," IEEE/ACM Joint Conference on Digital Libraries (JCDL 2014), 225-228, 2014.
 30. Wenyi Huang, Zhaohui Wu, Prasenjit Mitra, C. Lee Giles, "RefSeer: A citation recommendation system," IEEE/ACM Joint Conference on Digital Libraries (JCDL 2014), 371-374, 2014.
 31. Kyle Williams, Jian Wu, Sagnik Ray Choudhury, Madian Khabsa, C. Lee Giles: Scholarly big data information extraction and integration in the CiteSeerX digital library. ICDE Workshops 2014: 68-73
 32. Sujatha Das Gollapalli, Yanjun Qi, Prasenjit Mitra, C. Lee Giles, "Extracting Researcher Metadata with Labeled Features," Proceedings of the 2014 SIAM International Conference on Data Mining (SDM 2014), 740-748, 2014.
 33. Suppawong Tuarob, Sumit Bhatia, Prasenjit Mitra, C. Lee Giles, "Automatic Detection of Pseudocodes in Scholarly Documents Using Machine Learning," 12th International Conference on Document Analysis and Recognition (ICDAR 2013), 738-742, 2013.
 34. Sujatha Das Gollapalli, Prasenjit Mitra, C. Lee Giles, "Ranking experts using author-document-topic graphs. JCDL 2013: 87-96
 35. Cornelia Caragea, Adrian Silvescu, Prasenjit Mitra, C. Lee Giles, "Can't see the forest for the trees?: a citation recommendation system," 13th ACM/IEEE-CS Joint Conference on Digital Libraries, (JCDL '13), 111-114, 2013.
 36. Suppawong Tuarob, Line C. Pouchard, C. Lee Giles: Automatic tag recommendation for metadata annotation using probabilistic topic modeling. 13th ACM/IEEE-CS Joint Conference on Digital Libraries, (JCDL '13), 239-248, 2013.
 37. Suppawong Tuarob, Prasenjit Mitra, C. Lee Giles: "A classification scheme for algorithm citation function in scholarly works," 13th ACM/IEEE-CS Joint Conference on Digital Libraries, (JCDL '13), 367-368, 2013.

38. Sagnik Ray Choudhury, Suppawong Tuarob, Prasenjit Mitra, Lior Rokach, Andi Kirk, Silvia Szep, Donald Pellegrino, Sue Jones, Clyde Lee Giles, "A figure search engine architecture for a chemistry digital library," 13th ACM/IEEE-CS Joint Conference on Digital Libraries, (JCDL '13), 369-370, 2013.
39. Hung-Hsuan Chen, Pucktada Treeratpituk, Prasenjit Mitra, C. Lee Giles, "CSSeer: an expert recommendation system based on CiteseerX," 13th ACM/IEEE-CS Joint Conference on Digital Libraries, (JCDL '13), 381-382, 2013.
40. Sujatha Das Gollapalli, Cornelia Caragea, Prasenjit Mitra, C. Lee Giles: Researcher homepage classification using unlabeled data. WWW 2013: 471-482
41. Zhaohui Wu, Sujatha Das, Zhenhui Li, Prasenjit Mitra, C. Lee Giles, "Searching online book documents and analyzing book citations," ACM Symposium on Document Engineering (DocEng 2013), 81-90, 2013.
42. Kyle Williams, C. Lee Giles, "Near duplicate detection in an academic digital library," ACM Symposium on Document Engineering (DocEng 2013), 91-94, 2013.
43. Pradeep B. Teregowda, Clyde Lee Giles, "Scaling SeerSuite in the Cloud," IEEE International Conference on Cloud Engineering (IC2E 2013), 146-155, 2013.
44. Sagnik Ray Choudhury, Prasenjit Mitra, Andi Kirk, Silvia Szep, Donald Pellegrino, Sue Jones, C. Lee Giles: Figure Metadata Extraction from Digital Documents. ICDAR 2013: 135-139
45. Sujatha Das, Prasenjit Mitra, C. Lee Giles, "Phrase Pair Classification for Identifying Subtopics," European Conference on IR Research (ECIR 2012), 489-493, 2012.
46. Jing Fang, Prasenjit Mitra, Zhi Tang, C. Lee Giles, "Table Header Detection and Classification," Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence (AAAI 2012), 599-605, 2012.
47. Dayu Yuan, Prasenjit Mitra, Huiwen Yu, C. Lee Giles, "Iterative Graph Feature Mining for Graph Indexing," IEEE 28th International Conference on Data Engineering (ICDE 2012), 198-209, 2012.
48. Sujatha Das Gollapalli, Prasenjit Mitra, C. Lee Giles, "Similar researcher search in academic environments," Proceedings of the 12th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL 2012), 167-170, 2012.
49. Madian Khabsa, Pucktada Treeratpituk, C. Lee Giles, "AckSeer: a repository and search engine for automatically extracted acknowledgments from digital libraries," Proceedings of the 12th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL 2012), 185-194, 2012.
50. Suppawong Tuarob, Prasenjit Mitra, C. Lee Giles, "Improving algorithm search using the algorithm co-citation network," Proceedings of the 12th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL 2012), 277-280, 2012.
51. Pradeep B. Teregowda, Madian Khabsa, Clyde Lee Giles, "A system for indexing tables, algorithms and figures," Proceedings of the 12th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL 2012), 343-344, 2012.
52. Jian Wu, Pradeep B. Teregowda, Juan Pablo Fernández Ramírez, Prasenjit Mitra, Shuyi Zheng, C. Lee Giles, "The evolution of a crawling strategy for an academic document search engine: whitelists and blacklists," Web Science 2012 (WebSci 2012), 340-343, 2012.
53. Madian Khabsa, Sharon Koppman, C. Lee Giles, "Towards Building and Analyzing a Social Network of Acknowledgments in Scientific and Academic Documents," Social Computing, Behavioral - Cultural Modeling and Prediction - 5th International Conference (SBP 2012), 357-364, 2012.
54. Bingjun Sun, Prasenjit Mitra, C. Lee Giles, Karl T. Mueller, "Identifying, Indexing, and Ranking Chemical Formulae and Chemical Names in Digital Documents," ACM Transactions on Information Systems, 29(2), 12, 2011.
55. Saurabh Kataria, Prasenjit Mitra, Cornelia Caragea, C. Lee Giles, "Context Sensitive Topic Models for Author Influence in Document Networks," International Joint Conferences on Artificial Intelligence (IJCAI 2011), 2274-2280, 2011.
56. Sujatha Das Gollapalli, C. Lee Giles, Prasenjit Mitra, Cornelia Caragea, "On identifying academic homepages for digital libraries," Joint Conference on Digital Libraries (JCDL 2011), 123-132, 2011.
57. Sujatha Das Gollapalli, Prasenjit Mitra, C. Lee Giles, "Ranking authors in digital libraries," Joint Conference on Digital Libraries (JCDL 2011), 251-254, 2011.

58. Qi He, Daniel Kifer, Jian Pei, Prasenjit Mitra, C. Lee Giles, " Citation recommendation without author supervision," ACM Web Search and Data Mining (WSDM 2011), 744-764, 2011.
59. Sujatha Das Gollapalli, Prasenjit Mitra, C. Lee Giles: Ranking authors in digital libraries. JCDL 2011: 251-254, 2011.
60. Pucktada Treeratpituk; Pradeep Teregowda; Jian Huang; C. Lee Giles. SEERLAB: A System for Extracting Keyphrases from Scholarly Documents, *Proceedings of the 5th International ACL Workshop on Semantic Evaluation (SemEval 2010)*, 182-185, 2010.
61. P.B. Teregowda, B. Urgaonkar, C.L. Giles, "Cloud Computing: A Digital Libraries Perspective", *3rd IEEE International Conference on Cloud Computing*, 2010.
62. P.B Teregowda, B. Urgaonkar, C.L. Giles, "Cost Implications Of Moving To The Cloud: A Digital Libraries Perspective", *2nd USENIX Workshop on Hot Topics in Cloud Computing (HotCloud '10)*, 2010.
63. P.B. Teregowda, I.G. Councill, J.P Fernández R., M. Kasbha, S. Zheng, C.L. Giles, "SeerSuite: Developing a Scalable and Reliable Application Framework for Building Digital Libraries by Crawling the Web", 1st USENIX Conference on Web Application Development, 2010.
64. Na Li, Leilei Zhu, Prasenjit Mitra, Karl Mueller, Eric Poweleit, C. Lee Giles, "oreChem ChemXSeer: a semantic digital library for chemistry," Proceedings of the 10th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL 2010), 245-254, 2010.
65. Liang Gou, Xiaolong Zhang, Hung-Hsuan Chen, Jung-Hyun Kim, C. Lee Giles: Social network document ranking. JCDL, 313-322, 2010.
66. Sumit Bhatia, Prasenjit Mitra, C. Lee Giles: Finding algorithms in scientific articles. WWW, 1061-1062, 2010.
67. Qi He, Jian Pei, Daniel Kifer, Prasenjit Mitra, C. Lee Giles: Context-aware citation recommendation. WWW, 421-430, 2010.
68. Qingzhao Tan, Prasenjit Mitra, C. Lee Giles: Effectively Searching Maps in Web Documents. ECIR 2009: 162-176, 2009.
69. Ying Liu, Kun Bai, Prasenjit Mitra, C. Lee Giles: Improving the Table Boundary Detection in PDFs by Fixing the Sequence Error of the Sparse Lines. ICDAR 2009: 1006-1010, 2009.
70. Pucktada Treeratpituk, C. Lee Giles: Disambiguating authors in academic publications using random forests. JCDL 2009: 39-48, 2009.
71. Levent Bolelli, Seyda Ertekin, Ding Zhou, C. Lee Giles: Finding topic trends in digital libraries. JCDL 2009: 69-72, 2009.
72. Shuyi Zheng, Pavel Dmitriev, C. Lee Giles: Graph-based seed selection for web-scale crawlers. CIKM 2009: 1967-1970, 2009.
73. Shuyi Zheng, Pavel Dmitriev, C. Lee Giles: Graph based crawler seed selection. WWW 2009: 1089-1090, 2009.
74. Ying Liu, Kun Bai, Prasenjit Mitra, C. Lee Giles: Improving the Table Boundary Detection in PDFs by Fixing the Sequence Error of the Sparse Lines. ICDAR 2009: 1006-1010.
75. Shu Huang, Qiankun Zhao, Prasenjit Mitra, C. Lee Giles, "Hierarchical Location and Topic Based Query Expansion," Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence, (AAAI 2008), 1150-1155, 2008.
76. Saurabh Kataria, William Browner, Prasenjit Mitra, C. Lee Giles, "Automatic Extraction of Data Points and Text Blocks from 2-Dimensional Plots in Digital Documents," Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence (AAAI 2008), 1169-1174, 2008.
77. Xiaolong Zhang, Yan Qu, C. Lee Giles, Piyou Song, "CiteSense: supporting sensemaking of research literature," CHI 2008: 677-680, 2008.
78. Huajing Li, Zaiqing Nie, Wang-Chien Lee, C. Lee Giles, Ji-Rong Wen, "Scalable community discovery on textual data with relations," Proceedings of the 17th ACM Conference on Information and Knowledge Management (CIKM 2008), 1203-1212, 2008.
79. Ying Liu, Prasenjit Mitra, C. Lee Giles, "Identifying table boundaries in digital documents via sparse line detection," Proceedings of the 17th ACM Conference on Information and Knowledge Management (CIKM 2008), 1311-1320, 2008
80. Qingzhao Tan, Prasenjit Mitra, C. Lee Giles, "Metadata extraction and indexing for map search in web documents," Proceedings of the 17th ACM Conference on Information and Knowledge Management (CIKM 2008), 1367-1368, 2008.

81. Yang Sun, Huajing Li, Isaac G. Councill, Wang-Chien Lee, C. Lee Giles, "Measuring user preference changes in digital libraries," Proceedings of the 17th ACM Conference on Information and Knowledge Management (CIKM 2008), 1497-1498, 2008.
82. Jian Huang, Omid Madani, C. Lee Giles, "Error-driven generalist+experts (edge): a multi-stage ensemble framework for text categorization," Proceedings of the 17th ACM Conference on Information and Knowledge Management (CIKM 2008), 83-92, 2008.
83. Xiaonan Lu, Brewster Kahle, James Ze Wang, C. Lee Giles: A metadata generation system for scanned scientific volumes," ACM/IEEE Joint Conference on Digital Libraries (JCDL 2008), 167-176, 2008.
84. William Browner, Saurabh Kataria, Sujatha Das, Prasenjit Mitra, C. Lee Giles, "Segregating and extracting overlapping data points in two-dimensional plots," ACM/IEEE Joint Conference on Digital Libraries (JCDL 2008), 276-279, 2008.
85. Yang Sun, Huajing Li, Isaac G. Councill, Jian Huang, Wang-Chien Lee, C. Lee Giles, "Personalized ranking for digital libraries based on log analysis," 10th ACM International Workshop on Web Information and Data Management (WIDM 2008), 133-140, 2008.
86. Jian Huang, Ziming Zhuang, Jia Li, C. Lee Giles, "Collaboration over time: characterizing and modeling network evolution," Proceedings of the International Conference on Web Search and Web Data Mining (WSDM2008), 107-116, 2008.
87. Ding Zhou, Shenghuo Zhu, Kai Yu, Xiaodan Song, Belle L. Tseng, Hongyuan Zha, C. Lee Giles, "Learning multiple graphs for document recommendations," Proceedings of the 17th International Conference on World Wide Web (WWW 2008), 141-150, 2008.
88. Isaac G. Councill, C. Lee Giles, Min-Yen Kan, "ParsCit: an Open-source CRF Reference String Parsing Package," Proceedings of the International Conference on Language Resources and Evaluation (LREC 2008), 2008.
89. Bingjun Sun, Prasenjit Mitra, C. Lee Giles, "Mining, indexing, and searching for textual chemical molecule information on the web," Proceedings of the 17th International Conference on World Wide Web (WWW 2008), 735-744, 2008.
90. Ying Liu, Kun Bai, Prasenjit Mitra, C. Lee Giles, "Automatic searching of tables in digital libraries," Proceedings of the 16th International Conference on World Wide Web (WWW 2007), 1135-1136, 2007.
91. Qingzhao Tan, Ziming Zhuang, Prasenjit Mitra, C. Lee Giles, "Designing efficient sampling techniques to detect webpage updates," Proceedings of the 16th International Conference on World Wide Web (WWW 2007), 1147-1148, 2007.
92. Yang Song, Jian Huang, Isaac G. Councill, Jia Li, C. Lee Giles, "Generative models for name disambiguation," Proceedings of the 16th International Conference on World Wide Web (WWW 2007), 1163-1164, 2007.
93. Xiaonan Lu, James Ze Wang, Prasenjit Mitra, C. Lee Giles, "Deriving knowledge from figures for digital libraries," Proceedings of the 16th International Conference on World Wide Web (WWW 2007), 1229-1230, 2007.
94. Bingjun Sun, Qingzhao Tan, Prasenjit Mitra, C. Lee Giles, "Extraction and search of chemical formulae in text documents on the web," Proceedings of the 16th International Conference on World Wide Web (WWW 2007), 251-260, 2007.
95. Qingzhao Tan, Ziming Zhuang, Prasenjit Mitra, C. Lee Giles, "A clustering-based sampling approach for refreshing search engine's database," Tenth International Workshop on the Web and Databases (WebDB 2007), 2007.
96. Bingjun Sun, Prasenjit Mitra, C. Lee Giles, John Yen, Hongyuan Zha, "Topic segmentation with shared topic detection and alignment of multiple documents," Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2007), 199-206, 2007.
97. Yookyung Jo, Carl Lagoze, C. Lee Giles, "Detecting research topics via the correlation between graphs and texts," Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2007), 370-379, 2007.
98. Huajing Li, Wang-Chien Lee, Anand Sivasubramaniam, C. Lee Giles, "SearchGen: a synthetic workload generator for scientific literature digital libraries and search engines," ACM/IEEE Joint Conference on Digital Libraries, (JCDL 2007), 137-146, 2007.

99. Su Yan, Dongwon Lee, Min-Yen Kan, C. Lee Giles, "Adaptive sorted neighborhood methods for efficient record linkage," ACM/IEEE Joint Conference on Digital Libraries, (JCDL 2007), 185-194, 2007.
100. Ziming Zhuang, Ergin Elmacioglu, Dongwon Lee, C. Lee Giles, "Measuring conference quality by mining program committee characteristics," ACM/IEEE Joint Conference on Digital Libraries, (JCDL 2007), 225-234, 2007.
101. Yang Song, Jian Huang, Isaac G. Councill, Jia Li, C. Lee Giles, "Efficient topic-based unsupervised name disambiguation," ACM/IEEE Joint Conference on Digital Libraries, (JCDL 2007), 342-351, 2007.
102. Ying Liu, Kun Bai, Prasenjit Mitra, C. Lee Giles, "TableSeer: automatic table metadata extraction and searching in digital libraries," ACM/IEEE Joint Conference on Digital Libraries, (JCDL 2007), 91-100, 2007.
103. Ding Zhou, Sergey A. Orshanskiy, Hongyuan Zha, C. Lee Giles, "Co-ranking Authors and Documents in a Heterogeneous Network," Proceedings of the 7th IEEE International Conference on Data Mining (ICDM 2007) 739-744, 2007.
104. Shuyi Zheng, Ding Zhou, Jia Li, C. Lee Giles, "Extracting Author Meta-Data from Web Using Visual Features." Proceedings of the 7th IEEE International Conference on Data Mining (ICDM 2007) Workshop on Data Mining in Web 2.0 Environments, 33-40, 2007.
105. Xiaonan Lu, James Ze Wang, Prasenjit Mitra, C. Lee Giles, "Automatic Extraction of Data from 2-D Plots in Documents," 9th International Conference on Document Analysis and Recognition (ICDAR 2007), 188-192, 2007.
106. Huajing Li, Wang-Chien Lee, Anand Sivasubramaniam, C. Lee Giles, "A Hybrid Cache and Prefetch Mechanism for Scientific Literature Search Engines," Web Engineering, 7th International Conference, (ICWE 2007), 121-136, 2007.
107. Qingzhao Tan, Ziming Zhuang, Prasenjit Mitra, C. Lee Giles, "Efficiently Detecting Webpage Updates Using Samples," Web Engineering, 7th International Conference, (ICWE 2007), 285-300, 2007.
108. Ying Liu, Kun Bai, Prasenjit Mitra, C. Lee Giles, "TableRank: A Ranking Algorithm for Table Search and Retrieval," Proceedings of the Twenty-Second AAAI Conference on Artificial Intelligence (AAAI 2007), 317-322, 2007.
109. Prasenjit Mitra, C. Lee Giles, Bingjun Sun, Ying Liu, "ChemXSeer: a digital library and data repository for chemical kinetics, Proceedings of the First Workshop on CyberInfrastructure: Information Management in eScience (CIMS 2007), 7-10, 2007.
110. Yang Sun, C. Lee Giles, "Popularity Weighted Ranking for Academic Digital Libraries," ECIR 2007 Advances in Information Retrieval, 29th European Conference on IR Research (ECIR 2007), 605-612, 2007.
111. Umer Farooq, Craig H. Ganoë, John M. Carroll, C. Lee Giles, "Supporting distributed scientific collaboration: Implications for designing the CiteSeer collaboratory," Hawaii International Conference on System Sciences (HICSS 2007), 26, 2007.
112. Yang Song, Ding Zhou, Jian Huang, Isaac G. Councill, Hongyuan Zha, C. Lee Giles, "Boosting the Feature Space: Text Classification for Unstructured Data on the Web," IEEE International Conference on Data Mining (ICDM 2006), 1064-1069, 2006.
113. Ding Zhou, Xiang Ji, Hongyuan Zha, C. Lee Giles, "Topic evolution and social interactions: how authors effect research," Proceedings of the 13th ACM Conference on Information and Knowledge Management (CIKM 2006), 248-257, 2006.
114. C. Lee Giles, "The Future of CiteSeer: CiteSeer^x," 10th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD 2006), 2, 2006. (Keynote Talk).
115. Levent Bolelli, Seyda Ertekin, C. Lee Giles, "Clustering Scientific Literature Using Sparse Citation Graph Analysis." 10th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD 2006), 30-41, 2006.
116. Jian Huang, Seyda Ertekin, C. Lee Giles, "Efficient Name Disambiguation for Large-Scale Databases," 10th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD 2006), 536-544, 2006.
117. Anuj R. Jaiswal, C. Lee Giles, Prasenjit Mitra, James Ze Wang, "An architecture for creating collaborative semantically capable scientific data sharing infrastructures," 8th ACM International Workshop on Web Information and Data Management (WIDM 2006), 75-82, 2006.

118. Isaac G. Councill, C. Lee Giles, Ernesto Di Iorio, Marco Gori, Marco Maggini, Augusto Pucci, "Towards Next Generation CiteSeer: A Flexible Architecture for Digital Library Deployment," European Conference on Digital Libraries (ECDL 2006), 111-122, 2006.
119. Huajing Li, Isaac Councill, Wang-Chien Lee, C. Lee Giles, "CiteSeer^x: an architecture and web service design for an academic document search engine," 15th International Conference on World Wide Web (WWW 2006), 883-884, 2006.
120. Huajing Li, Isaac Councill, Levent Bolelli, Ding Zhou, Yang Song, Wang-Chien Lee, Anand Sivasubramaniam, C. Lee Giles, "CiteSeerX - A Scalable Autonomous Scientific Digital Library" Proceedings of the 1st International Conference on Scalable Information Systems 2006 (Infoscale 2006), 2006.
121. Isaac G. Councill, Huajing Li, Ziming Zhuang, Sandip Debnath, Levent Bolelli, Wang-Chien Lee, Anand Sivasubramaniam, C. Lee Giles, "Learning metadata from the evidence in an on-line citation matching scheme." Proceedings of the ACM and IEEE Joint Conference on Digital Libraries (JCDL 2006), 276-285, 2006. (Nominated for Best Paper Award).
122. Ying Liu, Prasenjit Mitra, C. Lee Giles, Kun Bai, "Automatic extraction of table metadata from digital documents," Proceedings of the ACM and IEEE Joint Conference on Digital Libraries (JCDL 2006), 339-340, 2006.
123. X. Lu, P. Mitra, J.Z. Wang, C.L. Giles, "Automatic Categorization of Figures in Scientific Documents," Proceedings of the ACM and IEEE Joint Conference on Digital Libraries (JCDL 2006), 129-138, 2006.
124. H. Han, H. Zha, C.L. Giles, "Name disambiguation in author citations using a K-way spectral clustering method," ACM/IEEE Joint Conference on Digital Libraries (JCDL 2005), 334-343, 2005.
125. Z. Zhuang, R. Wagle, C.L. Giles, "What's there and what's not?: focused crawling for missing documents in digital libraries," ACM/IEEE Joint Conference on Digital Libraries (JCDL 2005), 301-310, 2005. (Nominated for best student paper).
126. Y. Petinot, C.L. Giles, V. Bhatnagar, P.B. Teregowda, H. Han, I. Councill, "A Service-Oriented Architecture for Digital Libraries", Proceedings of the 2nd International Conference on Service Oriented Computing (ICSOC 2004), 263-268, 2004.
127. Y. Petinot, C.L. Giles, V. Bhatnagar, P.B. Teregowda, H. Han, I. Councill, "CiteSeer-API: Towards Seamless Resource Location and Interlinking for Digital Libraries", Proceedings of the 13th Conference on Information and Knowledge Management (CIKM 2004), Washington D.C., 2004.
128. Y. Petinot, C.L. Giles, V. Bhatnagar, P.B. Teregowda, H. Han, "Enabling interoperability for autonomous digital libraries: an API to CiteSeer services," ACM/IEEE Joint Conference on Digital Libraries (JCDL 2004), 372-373, 2004.
129. H. Han, C.L. Giles, H. Zha, C. Li, K. Tsioutsoulklis, "Two supervised learning approaches for name disambiguation in author citations," ACM/IEEE Joint Conference on Digital Libraries (JCDL 2004), 296-305, 2004.
130. G. Pant, K. Tsioutsoulklis, J. Johnson, C.L. Giles, "Panorama: extending digital libraries with topical crawlers," ACM/IEEE Joint Conference on Digital Libraries (JCDL 2004), 142-150, 2004.
131. J. Johnson, K. Tsioutsoulklis, C.L. Giles, "Evolving Strategies for Focused Web Crawling," International Conference for Machine Learning (ICML 2003), 298-305, 2003.
132. C.L. Giles, Y. Petinot, P.B. Teregowda, H. Han, S. Lawrence, A. Rangaswamy, N. Pal, "eBizSearch: a niche search engine for e-business," 26th Annual International ACM Conference on Research and Development in Information Retrieval (SIGIR), (poster), 413-414, 2003.
133. H. Han, C.L. Giles, E. Manavoglu, H. Zha, Z. Zhang, E. Fox, "Automatic Document Metadata Extraction using Support Vector Machines," ACM/IEEE Joint Conference on Digital Libraries (JCDL 2003), 37-48, 2003. (Nominated for Best Paper Award).
134. Petinot, P.B. Teregowda, H. Han, C.L. Giles, S. Lawrence, A. Rangaswamy, N. Pal, "eBizSearch: An OAI-Compliant Digital Library for eBusiness," ACM/IEEE Joint Conference on Digital Libraries (JCDL 2003), 199-209, 2003.
135. S. Debnath, D.M. Pennock, S. Lawrence, E. Glover, C.L. Giles, "Characterizing efficiency and information incorporation in sports betting markets," Ninth Research Symposium on Emerging Electronic Markets (RSEEM 2002), 39-52, 2002.

136. S.-T. Park, D.M. Pennock, C.L. Giles, R. Krovetz. "Analysis of lexical signatures for finding lost or related documents," 25th Annual International ACM Conference on Research and Development in Information Retrieval (SIGIR), 11-18, 2002.
137. M. Diligenti, F.M. Coetzee, S. Lawrence, C.L. Giles, M. Gori, "Focused Crawling Using Context Graphs," 26th International Conference on Very Large Databases (VLDB 2000), pp.527-534, 2000.
138. A. Popsecul, G.W. Flake, S. Lawrence, L. Ungar, C.L. Giles, "Clustering and Identifying Temporal Trends in Document Databases," IEEE Advances in Digital Libraries, 2000.
139. S. Lawrence, K. Bollacker, C.L. Giles, "Indexing and Retrieval of Scientific Literature on the Web," Eighth International Conference on Information and Knowledge Management (CIKM'99), 1999.
140. K. Bollacker, S. Lawrence, C.L. Giles, "A System For Automatic Personalized Tracking of Scientific Literature on the Web," DL'99 Digital Libraries, The 4th ACM Conference on Digital Libraries, p. 105, 1999.
141. K. Bollacker, S. Lawrence, C.L. Giles, "Distributed Error Correction," DL'99 Digital Libraries, The 4th ACM Conference on Digital Libraries, p. 232 (poster), 1999.
142. S. Lawrence, C.L. Giles, K. Bollacker, "Scientific Literature Indexing on the Web," Poster Proceedings of the 8th International World Wide Conference, 120-121, 1999.
143. S. Lawrence, C.L. Giles, K. Bollacker, "Autonomous Citation Matching." Proceedings of the 3rd ACM Annual Conference on Autonomous Agents, (poster) 392-393, 1999.
144. C.L. Giles, K. Bollacker, S. Lawrence, "CiteSeer: An Automatic Citation Indexing System," DL'98 Digital Libraries, The 3rd ACM Conference on Digital Libraries, 89-98, 1998 [one of eight papers short listed for best paper award].
145. K. Bollacker, S. Lawrence, C.L. Giles, "CiterSeer: An Autonomous Web Agent for Automatic Retrieval and Identification of Interesting Publications," Agents'98, 2nd International ACM Conference on Autonomous Agents, 116-123, 1998.