

Yi Chen

Curriculum Vitae

School of Computing, Informatics and Decision Systems Engineering
Arizona State University, Tempe, AZ 85287
Email: yi@asu.edu, Tel: (480) 292-1873
<http://www.public.asu.edu/~ychen127/>

Education

- **Ph.D. Computer Science**
University of Pennsylvania, USA, August 2005.
Dissertation: *Efficient XPath Query Processing on Stored and Streaming XML Data*
Advisor: Professor Susan B. Davidson
- **M.S. Computer Science**
University of Pennsylvania, USA, December 2000.
- **B.S. Computer Science**
Central South University, China, July 1999.

Employment

- Arizona State University: Associate Professor (with tenure) in Computer Science, 2011-present.
- Arizona State University: Affiliate Faculty in Biomedical Informatics Department, 2010-present.
- Arizona State University: Assistant Professor in Computer Science, 2005-present.
- IBM Research T.J. Watson Center: Research Intern, summer 2003.
- IBM Research T.J. Watson Center: Research Intern, summer 2002.

Awards and Honors:

- General Co-Chair, International Conference on the Management of Data (SIGMOD), 2012.
- Google Research Award, 2011.
- Outstanding Researcher in Computer Science and Engineering, Arizona State University, 2011.
- IBM Faculty Award, 2010.
- National Science Foundation CAREER Award, 2009.

- Faculty Recognition, ASU Ira A. Fulton Schools of Engineering, 2009.
- Steering Committee Member, KEYS (Keyword Search on Structured Data) workshop, 2009-present
- Dean’s Fellowship, University of Pennsylvania, 1999.

Research Interests

- Handling imprecise user queries on uncertain data;
- Text mining and natural language processing;
- Intelligent social computing and social media analysis,
- Scientific and business workflow management, including archiving, search, provenance analysis and optimization, etc.;
- Performance optimizations and scalability;
- Biomedical and health informatics.

Publications

Projects that are led by me are indicated by a # symbol at the end. To promote students’ visibility I put student participants ahead in the author lists.

Book Chapters

1. “Keyword Search on XML Data.” Z. Liu and Y. Chen. In *Advanced Applications and Structures in XML Processing: Label Streams, Semantics Utilization and Data Query Technologies*, co-edited by Changqing Li and Tok Wang Ling, IGI Global (2010). #

Refereed Journal Articles

2. “Differentiating Search Results on Structured Data.” Z. Liu, and Y. Chen. *ACM Transactions on Database Systems (TODS)* (2012) (to appear). #
3. “Incremental Information Extraction Using Relational Databases.” L. Tari, P. H. Tu, J. Hakenberg, Y. Chen, T. C. Son, G. Gonzalez, and C. Baral. *IEEE Transactions on Knowledge and Data Engineering (TKDE)*: 86 -99, (2012) (to appear).
4. “Query Expansion Based on Clustered Results.” Z. Liu, S. Natarajan, and Y. Chen. *Proceedings of the VLDB Endowment (PVLDB)* 4(1) (2011). Presented at the 37th International Conference on Very Large Data Bases (**VLDB 2011**). #
5. “Generating Sound Workflow Views for Correct Provenance Analysis.” Z. Liu, S. B. Davidson, and Y. Chen. *ACM Transactions on Database Systems (TODS)* 36(1):6, (2011). #

6. “Efficient Iceberg Query Evaluation using Compressed Bitmap Index.” B. He, H. Hsiao, Z. Liu, and Y. Chen. *IEEE Transactions on Knowledge and Data Engineering (TKDE)*: 14 pages, (2011).
7. “Processing Keyword Search on XML: a Survey.” Z. Liu, and Y. Chen. *World Wide Web Journal (WWWJ)* 14(5-6): 671-707 (2011). #
8. “Improving XML Search by Generating and Utilizing Informative Result Snippets.” Z. Liu, Y. Huang, and Y. Chen. *ACM Transactions on Database Systems (TODS)* 35(3): 43 pages, (2010). #
9. “Return Specification Inference and Result Clustering for Keyword Search on XML.” Z. Liu and Y. Chen. *ACM Transactions on Database Systems (TODS)* 35(2): 47 pages, (2010). #
10. “Searching Workflows with Hierarchical Views.” Z. Liu, Q. Shao, and Y. Chen. *Proceedings of the VLDB Endowment (PVLDB)* 3(1): 918-927 (2010). Presented at the 36th International Conference on Very Large Data Bases (**VLDB 2010**). #
11. “A Bi-labeling Based XPath Processing System.” Y. Chen, S. B. Davidson, and Y. Zheng. *Information Systems Journal* 35(2): 170-185 (2010).
12. “Query Processing over Incomplete Autonomous Databases: Query Rewriting Using Learned Data Dependencies.” G. Wolf, A. Kalavagattu, H. Khatri, R. Balakrishnan, B. Chokshi, J. Fan, Y. Chen, and S. Kambhampati. *International Journal on Very Large Data Base (VLDBJ)*, Special Issue on Uncertain and Probabilistic Databases, 18(5): 1167-1190 (2009).
13. “Structured Search Result Differentiation.” Z. Liu, P. Sun, and Y. Chen. *Proceedings of the VLDB Endowment (PVLDB)* 2(1): 313-324 (2009). Presented at the 35th International Conference on Very Large Data Bases (**VLDB 2009**). #
14. “Efficiently Discovering Critical Workflows in Scientific Explorations.” Q. Shao, P. Sun, and Y. Chen. *Future Generation Computer Systems* 25(5): 577-585 (2009). #
15. “Reasoning and Identifying Relevant Matches for XML Keyword Search.” Z. Liu and Y. Chen. *Proceedings of the VLDB Endowment (PVLDB)* 1(1): 921-932 (2008). Presented at the 34th International Conference on Very Large Data Bases (**VLDB 2008**). #

Refereed Conference Articles

16. “Enabling Privacy in Provenance-Aware Workflow Systems (vision paper).” S. B. Davidson, S. Khanna, S. Roy, J. Stoyanovich, V. Tannen, Y. Chen, and T. Milo. In *Proceedings of the 5th Biennial Conference on Innovative Data Systems Research (CIDR 2011)*: 215-218.
17. “Efficient and Scalable Data Evolution with Column Oriented Databases.” Z. Liu, B. He, H. Hsiao, and Y. Chen. In *Proceedings of the 14th International Conference on Extending Database Technology (EDBT 2011)*: 105-116. #
18. “Content-Aware Resolution Sequence Mining for Ticket Routing.” P. Sun, S. Tao, X. Yan, N. Anerousis, and Y. Chen. In *Proceedings of the 8th International Conference on Business Process Management (BPM 2010)*: 243-259. #

19. “Generative Models for Ticket Resolution in Expert Networks.” G. Miao, L. Moser, X. Yan, S. Tao, Y. Chen, and N. Anerousis. In Proceedings of the *16th International Conference on Knowledge Discovery and Data Mining (SIGKDD 2010)*: 733-742.
20. “XSACT: A Comparison Tool for Structured Search Results.” Z. Liu, S. Natarajan, P. Sun, S. Booher, T. Meehan, R. Winkler, and Y. Chen. Demonstration description. In Proceedings of the *VLDB Endowment (PVLDB)* 3(2): 1581-1584. Presented at the 36th International Conference on Very Large Data Bases (**VLDB 2010**). #
21. “CODS: Evolving Data Efficiently and Scalably in Column Oriented Databases.” Z. Liu, S. Natarajan, B. He, H. Hsiao, and Y. Chen. Demonstration description. In Proceedings of the *VLDB Endowment (PVLDB)* 3(2): 1521-1524. Presented at the 36th International Conference on Very Large Data Bases (**VLDB 2010**). #
22. “Assessing Expertise Awareness in Resolution Networks.” Y. Chen, S. Tao, X. Yan, N. Anerousis, and Q. Shao. In Proceedings of the *International Conference on Social Networks Analysis and Mining (ASONAM 2010)*: 128-135. #
23. “MASS: a Multi-fAcet domain-Specific influential blogger mining System.” Y. Cai, and Y. Chen. Demonstration description. In Proceedings of the *26th IEEE International Conference on Data Engineering (ICDE 2010)*: 1109-1112. #
24. “GenerIE: Information Extraction Using Database Queries.” L. Tari, P. H. Tu, J. Hakenberg, Y. Chen, T. C. Son, G. Gonzalez, and C. Baral. Demonstration description. In Proceedings of the *26th IEEE International Conference on Data Engineering (ICDE 2010)*: 1121-1124.
25. “TargetSearch: A Ranking Friendly XML Keyword Search Engine.” Z. Liu, Y. Cai, and Y. Chen. Demonstration description. In Proceedings of the *26th IEEE International Conference on Data Engineering (ICDE 2010)*: 1101-1104. #
26. “Detecting and Resolving Unsound Workflow Views for Correct Provenance Analysis.” P. Sun, Z. Liu, S. B. Davidson, and Y. Chen. In Proceedings of the *28th ACM SIGMOD International Conference on Management of Data (SIGMOD 2009)*: 549-562. #
27. “WOLVES: Achieving Correct Provenance Analysis by Detecting and Resolving Unsound Workflow Views.” P. Sun, Z. Liu, S. Natarajan, S. Davidson, and Y. Chen. Demonstration description. In Proceedings of the *VLDB Endowment (PVLDB)* 2(2): 1614-1617. Presented at the 35th International Conference on Very Large Data Bases (**VLDB 2009**). #
28. “Mining Influential Bloggers: From General to Domain Specific.” Y. Cai and Y. Chen. In Proceedings of the *13th International Conference on Knowledge-Based and Intelligent Information and Engineering Systems (KES 2009)*: 447-454. #
29. “WISE: a Workflow Information Search Engine.” Q. Shao, P. Sun, and Y. Chen. Demonstration description. In Proceedings of the *25th International Conference on Data Engineering (ICDE 2009)*: 1491-1494. #
30. “Query Biased Snippet Generation in XML Search.” Y. Huang, Z. Liu, and Y. Chen. In Proceedings of the *27th ACM SIGMOD International Conference on Management of Data (SIGMOD 2008)*: 315-326. Provisional patent application filed in 2008. #

31. "Efficient Ticket Routing by Resolution Sequence Mining." Q. Shao, Y. Chen, S. Tao, X. Yan, and N. Anerousis. In Proceedings of the *14th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (SIGKDD 2008)*: 605-613. #
32. "EasyTicket: A Ticket Routing Recommendation Engine for Enterprise Problem Resolution." Q. Shao, Y. Chen, S. Tao, X. Yan, and N. Anerousis. Demonstration description. In Proceedings of the *VLDB Endowment (PVLDB)* 1(2): 1436-1439. Presented at the 34th International Conference on Very Large Data Bases (**VLDB 2008**). #
33. "eXtract: Generating Query Biased Result Snippet for XML Search." Y. Huang, Z. Liu, and Y. Chen. Demonstration description. In Proceedings of the *VLDB Endowment (PVLDB)* 1(2): 1392-1395. Presented at the 34th International Conference on Very Large Data Bases (**VLDB 2008**). #
34. "Answering Keyword Queries on XML Using Materialized Views." Z. Liu and Y. Chen. In Proceedings of the *24th International Conference on Data Engineering (ICDE 2008)*: 1501-1503. #
35. "Identifying Meaningful Return Information for XML Keyword Search." Z. Liu and Y. Chen. In Proceedings of the *26th ACM SIGMOD International Conference on Management of Data (SIGMOD 2007)*: 329-340. #
36. "Query Processing over Incomplete Autonomous Databases." G. Wolf, H. Khatri, B. Chokshi, J. Fan, Y. Chen, and S. Kambhampati. In Proceedings of the *33rd International Conference on Very Large Data Bases (VLDB 2007)*: 651-662.
37. "XSeek: A Semantic XML Search Engine Using Keywords." Z. Liu, J. Walker, and Y. Chen. Demonstration description. In Proceedings of the *33rd International Conference on Very Large Data Bases (VLDB 2007)*: 1330-1333. #
38. "QPIAD: Query Processing over Incomplete Autonomous Databases." H. Khatri, J. Fan, Y. Chen, and S. Kambhampati. In Proceedings of the *23rd International Conference on Data Engineering (ICDE 2007)*: 1430-1432.
39. "Approximate Structural Matching over Ordered XML Documents." N. Agarwal, M. G. Oliveras, and Y. Chen. In Proceedings of the *11th International Database Engineering & Applications Symposium (IDEAS 2007)*: 54-62. #
40. "QUIC: A System for Handling Imprecision & Incompleteness in Autonomous Databases." G. Wolf, H. Khatri, Y. Chen, S. Kambhampati. Demonstration description. In Proceedings of the *3rd Biennial Conference on Innovative Data Systems Research (CIDR 2007)*: 263-268.
41. "PEPX: A Query-Friendly Probabilistic XML Database." T. Li, Q. Shao, and Y. Chen. In Proceedings of the *15th ACM Conference on Information and Knowledge Management (CIKM 2006)*: 848-849. #
42. "An Efficient XPath Query Processor for XML Streams." Y. Chen, S. B. Davidson, and Y. Zheng. In Proceedings of the *22nd International Conference on Data Engineering (ICDE 2006)*: 79-90.

43. “Designing and Evaluating an XPath Dialect for Linguistic Queries.” S. Bird, Y. Chen, S. B. Davidson, H. Lee, and Y. Zheng. In Proceedings of the *22nd International Conference on Data Engineering (ICDE 2006)*: 52-61.
44. “ViteX: A Streaming XPath Processing System.” Y. Chen, S. B. Davidson, and Y. Zheng. Demonstration description. In Proceedings of the *21st International Conference on Data Engineering (ICDE 2005)*: 1118-1119.
45. “BLAS: An Efficient XPath Processing System.” Y. Chen, S. B. Davidson, and Y. Zheng. In Proceedings of the *23rd ACM SIGMOD International Conference on Management of Data (SIGMOD 2004)*: 47-58.
46. “EXPedite: A System for Encoded XML Processing.” Y. Chen, G. A. Mihaila, S. B. Davidson, and S. Padmanabhan. In Proceedings of the *13rd ACM Conference on Information and Knowledge Management (CIKM 2004)*: 108-117.
47. “RRXS: Redundancy Reducing XML Storage in Relations.” Y. Chen, S. B. Davidson, C. Hara, and Y. Zheng. In Proceedings of the *29th International Conference on Very Large Data Bases (VLDB 2003)*: 189-200.
48. “XKvalidator: A Constraint Validator for XML.” Y. Chen, S. B. Davidson, and Y. Zheng. In Proceedings of the *11th ACM Conference on Information and Knowledge Management (CIKM 2002)*: 446-452.

Refereed Tutorials

49. “Keyword-based Search and Exploration on Databases.” Y. Chen, W. Wang, and Z. Liu. Tutorial Description. In Proceedings of the *27th International Conference on Data Engineering (ICDE 2011)*. #
50. “Keyword Search on Structured and Semi-structured Data.” Y. Chen, W. Wang, Z. Liu, and X. Lin. Tutorial Description. In Proceedings of the *28th ACM SIGMOD International Conference on Management of Data (SIGMOD 2009)*: 1005-1010. #

Referred Workshop Articles

51. “Evaluating Rank Accuracy based on Incomplete Pairwise Preferences.” B. Ackerman and Y. Chen. In Proceedings of the *User-Centric Evaluation of Recommender Systems and Their Interfaces (UCERSTI 2011)* #
52. “Storing and Discovering Critical Workflows from Log in Scientific Exploration.” Q. Shao, M. Kinsy, and Y. Chen. In Proceedings of the *IEEE International Workshop on Scientific Workflows (IEEE SCW 2007)*: 209-212. #
53. “Modeling and Storing Scientific Protocols.” N. Kwasnikowska, Y. Chen, and Z. Lacroix. In Proceedings of the *International Workshop on Knowledge Systems in Bioinformatics (KSin-BIT 2006)*: 730-739.
54. “Extending XPath to Support Linguistic Queries.” S. Bird, Y. Chen, S. B. Davidson, H. Lee, and Y. Zheng. In Proceedings the of *Programming Language Technologies for XML (PLAN-X 2005)*.

55. "Efficient Path Query Processing on Encoded XML." Y. Chen, G. A. Mihaila, S. B. Davidson, and S. Padmanabhan. In Proceedings of the *International Workshop on High Performance XML Processing*, in conjunction with WWW, 2004.
56. "L-Tree: A Dynamic Labeling Structure for Ordered XML Data." Y. Chen, G. A. Mihaila, S. Padmanabhan, and R. Bordawekar. In Proceedings of the *International Workshop on Database Technologies for Handling XML Information on the Web (DataX 2004)*: 209-218, in conjunction with EDBT, 2004. Springer Lecture Notes in Computer Science 3268, 2004. US Patent: US 7,313,563 B2.
57. "Constraint Preserving XML Storage in Relations." Y. Chen, S. B. Davidson, and Y. Zheng. In Proceedings of the *5th International Workshop on the Web and Databases (WebDB 2002)*: 7-12, in conjunction with SIGMOD.

Invited Papers & Talks

58. "Searching, Analyzing and Exploring Databases." Y. Chen, W. Wang, and Z. Liu. *Invited Tutorial*. In Proceedings of the *16th International Conference on Database Systems for Advanced Applications (DASFAA 2011)*. #
59. "On Provenance and Privacy." S. B. Davidson, S. Khanna, S. Roy, J. Stoyanovichi, V. Tannen, and Y. Chen. *Keynote Talk*. In Proceedings of the *14th International Conference on Database Theory (ICDT 2011)*.
60. "Problem-Solving in Collaborative Social Networks." Y. Chen, S. Tao, and X. Yan. *Invited Paper*. In Proceedings of the Workshop on *Personalized Access, Profile Management, and Context Awareness in Databases (PersDB 2011)*. #
61. "Query Results Ready, Now What?" Z. Liu, and Y. Chen. *IEEE Data Engineering Bulletin* 33(1): 46-53 (2010). #
62. "On User Views in Scientific Workflow Systems." S. Davidson, Y. Chen, P. Sun and S. Cohen-Boulakia. *The First International Workshop on the role of Semantic Web in Provenance Management (SWPM 2009)*, in conjunction with ISWC.
63. "Challenges, Techniques and Directions in Building XSeek: an XML Search Engine." Z. Liu, P. Sun, Y. Huang, Y. Cai, and Y. Chen. *IEEE Data Engineering Bulletin* 32(2): 36-43 (2009). #

Other Publications

64. "An XML Framework for Integrating Continuous Queries, Composite Event Detection, and Database Condition Monitoring for Multiple Data Streams." S. D. Urban, S. W. Dietrich, and Y. Chen. *Event Processing Seminar*, 2007
65. "Data Quality and Provenance in Information Integration." Y. Chen. Position paper. NSF workshop on *Information Integration*, 2006. #
66. "Efficient XPath Query Processing on Stored and Streaming XML Data." Y. Chen. Ph.D. Dissertation, Computer and Information Science Department, University of Pennsylvania, August, 2005. #

Research Experience

Research Grants

- Google Research Award, “Supporting Keyword Search on Web Databases”, \$49,561. 2011. PI: Yi Chen
- NSF IIS “Student Research and Educational Activities at ACM SIGMOD 2012”, \$20,00. 8/1/2011 - 7/31/2012. PI: K. Selcuk Candan, co-PI: Yi Chen
- IBM Faculty Award, “Effectively Managing Collaborative Processes and Associated Social Networks”, \$10,000. 2010. PI: Yi Chen
- IBM Award, “Designing Efficient Query Evaluation Algorithms by Exploiting Storage Class Memory”, \$10,000. 2010. PI: Yi Chen
- NSF III “CAREER: Analyzing and Exploiting Meta-information for Keyword Search on Semi-structured Data” with REU supplements, \$478,659. 3/1/2009 - 2/28/2014. PI: Yi Chen
- NSF III Collaborative Research “Mining and Optimizing Ad Hoc Workflows,” \$249,817. 9/1/2009 - 8/31/2012. PI: Yi Chen
- NSF III “Enabling Effective Access to Scientific Workflows,” \$87,370, 9/1/2007 - 8/31/2009. PI: Yi Chen
- NSF SEIII Collaborative Research “ProtocolDB: Archiving and Querying Scientific Protocols, Data and Provenance,” \$499,996. 8/1/2006 - 7/31/2009. PI: Zoe Lacroix; co-PI: Yi Chen
- Science Foundation Arizona (SFAZ) “Generalized Text Extraction from Life Science and Biomedicine Abstracts: empowering the CBioC Mass Collaborative Curation and Reasoning Systems,” \$138,749, 3/1/2007 - 2/28/2008. PI: Chitta Baral; co-PIs Yi Chen, Graciela Gonzalez, Lokesh Joshi.

Patents and Applications

- “Method, System and Recording Medium for Maintaining the Order of Nodes in a Hierarchical Document”. Rajesh Bordawekar, Yi Chen, George A. Mihaila, and Sriram Padmanabhan (The inventors are listed alphabetically), United States Patent: US 7,313,563 B2, December 2007.
- “Method of Creating One or More Snippets and System for the Same”. Yi Chen, Ziyang Liu and Yu Huang. Provisional Application for United States Patent, submitted by Arizona Science and Technology Enterprises LLC, Docket No. M8-103/0224549, 2008.

Invited Talks

- “Keyword-based Search and Exploration on Databases.” with Wei Wang, Ziyang Liu. Tutorial, ICDE, Hannover, Germany, April 15, 2011.

- “Information Discovery through Learning, Search and Exploration”, Computer Science Department Colloquium, University of California, Riverside, Feb 22, 2011.
- “Keyword Search on Databases and Beyond”, Computer Science Department Colloquium, Purdue University, IN, Nov 15, 2010.
- “Keyword Search on Structured and Semi-Structured Data” with Wei Wang, Ziyang Liu, and Xuemin Lin. Tutorial, SIGMOD, Providence, RI, July 1, 2009.
- “Keyword Search on XML Data and Beyond”, Computer Science Department Colloquium, University of California, Santa Barbara, CA, Nov 10, 2008.
- “Expediting the Search of Information, Experts, and Beyond”, Computer Science and Engineering Department Faculty Research Seminar, Arizona State University, Oct 31, 2008.
- “What Users Want in XML Keyword Search”, Computer Science Department Colloquium, Hong Kong University of Science and Technology, HK, May 14, 2007.
- “Efficiently Querying Tree-structured Data”, Computer Science Department Colloquium, University of Arizona, AZ, September 21, 2006.
- “Efficient Query Processing on XML Streams”, Database Group Seminar, University of California, Los Angeles, May 19, 2006;
- “Efficient Query Processing on XML Streams”, Database Group Seminar, University of California, San Diego, April 21 2006.
- “Enhancing Traditional Databases to Support Scientific Data Management Applications.” ASU Mathematics Department Seminar, AZ, January 23, 2006.
- “Managing XML Data Effectively in Relational Databases” Bell Labs, NJ, October 28, 2005.
- “Constraint Preserving XML Storage in Relations” IBM T.J. Watson Research Center, NY, 2002.

University Experience

Ph.D. Dissertation Supervision

- Yi Shan, Thesis Topic: *Scalable Processing of Keyword Queries on Large-Scale Databases.*
- Yunzhong Liu, Thesis Topic: *Modeling and Analyzing Mass Collaboration*
- Foruhar Shiva (co-advised with Prof. Susan Urban), Thesis Topic: *Application of a Temporal Database Framework for Processing Event Queries*
- Jian Yang (co-advised with Prof. Graciela Gonzalez), Thesis Topic: *Data Warehousing and Analytics for Health-Care*

- Ziyang Liu (August 2011), *Enhancing the Usability of Complex Structured Data by Supporting Keyword Search*. Initial appointment at NEC research labs.

Master's Thesis Supervision

- Brian Ackerman, Thesis Topic: *Context-Aware Recommendation By Exploiting Implicit Feedback*.
- Doug Stoeckmann, Thesis Topic: *Workflow Search*
- Sivaramakrishna Natarajan (May 2011), *Query Expansion for Keyword Search on Databases*. Employment: eBay.
- Yichuan Cai (December 2009), *A Framework for Identifying Influential Bloggers in the Blogosphere*. Employment: Amazon.com
- Karan Popli (May 2008), *TwigMF: Extending the XPATH Streaming Processor to Support the Following and Following-Sibling axes*. Employment: Bloomberg L.P.
- Ziyang Liu (May 2008), *Semantics Inference and Efficiency Improvement for XML Keyword Search*. Continuing PhD study in ASU.

Undergraduate Research Project Supervision

- *2011*: Brian Ackerman - NSF REU grant,
Project: *Context-aware Recommendation*;
- *2010*: Brian Ackerman, Tim Johnsen, Brian Smith - NSF REU grant,
Project: *Search and Recommendation in Entity Databases*;
Stephen Booher, Brad Bootz, Michael Sedillo - NSF REU grant,
Project: *Keyword Search on Heterogeneous Relational Databases*;
- *2009*: Stephen Booher, Tim Meehan, Robert Winkler - NSF REU grant,
Project: *Result Differentiation for Structured Search*;
Todd MacLeod - NSF REU grant,
Project: *Searching Expertise in Social Networks*;
- *2008*: Temitope Akinwande - CSE Undergraduate Research Scholarship,
Project: *Searching Workflows with Hierarchical Views*;
- *2007*: Michel Kinsy - CSE Undergraduate Research Scholarship,
Project: *Workflow Optimization by Log Analysis*;
Jeffrey Walker - Fulton Undergraduate Research Initiative,
Project: *Keyword Search on XML Data*.

Awards of My Students

- Outstanding Graduate, Computer Science and Engineering (CSE), ASU: Ziyang Liu, 2011, in recognition of his outstanding research during his PhD studies.

- Ford Graduate Engineering Fellowship: Brian Ackerman, 2011-2012, \$5,000 based on his research achievements advised by me and outstanding academic studies during his undergraduate studies.
- Science Foundation Arizona (SFAz) Graduate Fellowship: Ziyang Liu, 2008-2010. Approx. \$30K/ year stipend with tuition waived, in recognition of his outstanding research during his M.S. studies.

Courses Taught

- CSE 591- Data on the Web
(Arizona State University, Fall 2005, Spring 2007, Fall 2007, Spring 2009)
This newly developed graduate course addresses the data management issues for semi-structured data, which are commonly found in web and scientific applications. Techniques on modeling, storing, querying and searching semi-structured data are covered. Enrollment: 15-40.
- CSE 511- Semi-Structured Data Management
(Arizona State University, Fall 2009, Spring 2011, Fall 2011)
This course is evolved from CSE 591 - Data on the Web, after its successful offerings as “special topics” course. Enrollment: 15-40.
- CSE 412/598, CBS 598- Database Management
(Arizona State University, Spring 2008, Fall 2008, Spring 2009, Fall 2009, Fall 2010, Fall 2011)
This course is crossed linked for undergraduate and graduate students in computer science and graduate students in computational biosciences. It gives introduction to DBMS concepts, including data models, query languages, relational database theory, database security, transactions and concurrency control. Enrollment: 40-80.
- CSE 340 Principles of Programming Languages
(Arizona State University, Spring 2006, Fall 2006, Fall 2007, Spring 2012)
This is a core undergraduate course for introduction to formal syntactic and semantic descriptions, compilation and implementation issues, and theoretical foundations for several programming paradigms. Enrollment: 35-80.
- ASU101- The ASU Experience
(Arizona State University, Fall 2009, Fall 2010)
This is a required course for first-year students that introduces the culture and requirements of ASU, and Computer Systems Engineering and Computer Science as a profession. Enrollment: 19.

University Service

- **School Committees:**
Fulton Undergraduate Research Initiative (FURI) (2011-2012)
Fulton Schools of Engineering Student Affairs (2007-2008)
- **Departmental/Program Committees:**
Graduate Admission Committee (2011-present),
Undergraduate Program Committee (2007-2009),
Computing Resources Committee (2006-2007);

- **Course Coordinator:**
CSE 340 Principles of Programming Languages (2008-present)
CSE 414 Advanced Database Concepts (2008-present).

Professional Experience

- General Co-Chair: ACM SIGMOD International Conference on Management of Data (SIGMOD)'2012.
- Guest Editor
 - IEEE Transactions on Knowledge and Data Engineering (TKDE), Special Issue on “Keyword Search on Structured Data” (2011),
 - Proceedings of the VLDB Endowment (PVLDB) (2010),
- Steering Committee Member: KEYS (Keyword Search on Structured Data) workshop, 2009-present
- Program Committee Co-Chair: First International Workshop on Keyword Search on Structured Data (KEYS), co-located with SIGMOD, 2009.
- Publicity and Proceedings Chair: ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems (PODS), 2009.
- ACM SIGSPATIAL NSF Travel Award Committee member, 2010, 2011
- Program Committee Member:
 - 2012:*
 - ACM SIGMOD International Conference on Management of Data (SIGMOD)
 - International Conference on Data Engineering (ICDE)
 - International Conference on Extending Database Technology (EDBT)
 - International Conference on Database Systems for Advanced Applications (DASFAA)
 - International Workshop on Keyword Search on Structured Data (KEYS) (co-located with SIGMOD).
 - 2011:*
 - ACM SIGMOD International Conference on Management of Data (SIGMOD)
 - International World Wide Web Conference (WWW)
 - The Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD)
 - International Conference on Advanced Data Mining and Applications (ADMA)
 - International Workshop on Information Integration on the Web (IIWeb) (co-located with WWW)

- International Workshop on Ranking in Databases (DBRank) (co-located with VLDB)

2010:

- ACM SIGMOD International Conference on Management of Data (SIGMOD)
- International World Wide Web Conference (WWW)
- ACM Conference on Information and Knowledge Management (CIKM)
- Web and Databases (WebDB) (co-located with SIGMOD)
- International Workshop on Massive Data Analytics over the Cloud (co-located with WWW)
- Sixth International XML Database Symposium (XSym) (co-located with VLDB)
- International Workshop on New Trends In Information Integration (co-located with ICDE)

2009:

- International Conference on Very Large Data Bases (VLDB)
- International Conference on Data Engineering (ICDE)
- ACM Conference on Information and Knowledge Management (CIKM)
- International Conference on Database Systems for Advanced Applications (DASFAA)
- The Joint International Conferences on Asia-Pacific Web Conference (APWeb) and Web-Age Information Management (WAIM)
- ACM International Workshop on Web Information and Data Management (WIDM)

2008:

- International Conference on Very Large Data Bases (VLDB),
- AAAI Conference on Artificial Intelligence (AAAI)
- ACM Conference on Information and Knowledge Management (CIKM)
- International Conference on Database and Expert Systems Applications (DEXA)
- International Workshop on New Trends in Information Integration (co-located with VLDB)
- ACM International Workshop on Web Information and Data Management (WIDM)

2007:

- International Conference on Database and Expert Systems Applications (DEXA)
- Workshop on Database Management and Application over Networks (DBMAN).

2006:

- International XML Database Symposium (XSym, co-located with VLDB)
- International Conference on Image and Video Retrieval (CIVR)

- International Workshop on XML, Web, and Internet Contents Technologies (XWICT)
- 2005: International XML Database Symposium (XSym, co-located with VLDB)
- Journal Reviewer
 - *ACM Transactions on Database Systems* (TODS).
 - *International Journal on Very Large Data Bases* (VLDBJ).
 - *IEEE Transactions on Knowledge and Data Engineering* (TKDE).
 - *World Wide Web Journal* (WWWJ)
 - *ACM Transactions on Knowledge Discovery from Data* (TKDD).
 - *ACM Transactions on Internet Technology* (TOIT).
 - *Information Systems*
 - *Data & Knowledge Engineering* (DKE).
 - *Journal of Computer and System Sciences* (JCSS).
 - *Journal of Computer Science and Technology* (JCST).
- Grant Reviewer & Panelist
 - National Science Foundation, Grant Panelist (NSF): 2007, 2009, 2010, 2011.
 - The Air Force Office of Scientific Research (AFOSR), Grant Reviewer (USA): 2010.
 - Research Grant Council (RGC) of Hong Kong, Grant Reviewer: 2011.
 - Natural Sciences and Engineering Research Council of Canada (NSERC), Grant Reviewer: 2011.
- Grace Hopper Scholarship Reviewer: 2007, 2008, 2009.