

Yi Chen

Department of Computer Science and Engineering, School of Computing and Informatics
Ira A. Fulton School of Engineering, Arizona State University, Tempe, AZ 85281
E-mail: yi@asu.edu WWW: <http://www.public.asu.edu/~ychen127>

EDUCATION

- 2005 Ph.D. Computer Science, University of Pennsylvania, PA.
Advisor: Susan B. Davidson
Thesis: Efficient XPath Query Processing on Stored and Streaming XML Data
- 2000 M.S. Computer Science, University of Pennsylvania, PA.
- 1999 B.S. Computer Science, Central South University, China

EMPLOYMENT

- 2005- Assistant Professor, Department of Computer Science and Engineering,
Arizona State University, AZ
- Summer 2003 Research Intern. IBM Research, T.J. Watson Center, NY.
- Summer 2002 Research Intern. IBM Research, T.J. Watson Center, NY.
- 2001 - 2005 Research Assistant, University of Pennsylvania, PA.
- 2000 - 2001 Teaching Assistant, University of Pennsylvania, PA.
- 1999 - 2000 Research Assistant (Dean's Fellow), University of Pennsylvania, PA.

REFEREED PUBLICATIONS

- 1) Qihong Shao, Peng Sun, and Yi Chen: "WISE: a Workflow Information Search Engine." Demo Description. In Proceedings of the 25th International Conference on Data Engineering (ICDE), 2009 (to appear).
- 2) Ziyang Liu and Yi Chen: "Reasoning and Identifying Relevant Matches for XML Keyword Search." In *PVLDB Journal (Proceedings of the VLDB Endowment)*, Volume 1, Number 1, 2008
- 3) Qihong Shao, Yi Chen, Shu Tao, Xifeng Yan, and Nikos Anerousis: "EasyTicket: A Ticket Routing Recommendation Engine for Enterprise Problem Resolution". In *PVLDB Journal (Proceedings of the VLDB Endowment)*, Volume 1, Number 2, 2008
- 4) Yu Huang, Ziyang Liu and Yi Chen: "eXtract: Generating Query Biased Result Snippet for XML Search." In *PVLDB Journal (Proceedings of the VLDB Endowment)*, Volume 1, Number 2, 2008
- 5) Qihong Shao, Yi Chen, Shu Tao, Xifeng Yan, and Nikos Anerousis: "Efficient Ticket Routing by Resolution Sequence Mining." (long presentation). In *Proceedings of 14th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2008.
- 6) Qihong Shao, Peng Sun, and Yi Chen. "Efficiently discovering critical workflows in scientific explorations." *Journal of Future Generation Computer Systems (FGCS)*, 2008 (to appear).

- 7) Yu Huang, Ziyang Liu and Yi Chen. "Query Biased Snippet Generation in XML Search." In *Proceedings of 27th ACM SIGMOD International Conference on Management of Data*, 2008
- 8) Ziyang Liu and Yi Chen. "Answering Keyword Queries on XML Using Materialized Views." In *Proceedings of The 24th International Conference on Data Engineering*, 2008
- 9) Ziyang Liu and Yi Chen. "Identifying Meaningful Return Information for XML Keyword Search." In *Proceedings of 26th ACM SIGMOD International Conference on Management of Data*, 2007
- 10) Garrett Wolf, Hemal Khatri, Bhaumik Chokshi, Jianchun Fan, Yi Chen and Subbarao Kambhampati: "Query Processing over Incomplete Autonomous Databases." In *Proceedings of 33rd International Conference on Very Large Data Bases (VLDB)*, 2007
- 11) Ziyang Liu, Jeffrey Walker and Yi Chen: "XSeek: A Semantic XML Search Engine Using Keywords." Demo description. In *Proceedings of 33rd International Conference on Very Large Data Bases (VLDB)*, 2007
- 12) Hemal Khatri, Jianchun Fan, Yi Chen, and Subbarao Kambhampati. "QPIAD: Query Processing over Incomplete Autonomous Databases." In *Proceedings of 23rd International Conference on Data Engineering (ICDE)*, 2007
- 13) Nitin Agarwal, Magdiel Galan Oliveras, and Yi Chen: "Approximate Structural Matching over Ordered XML Documents." In *Proceedings of Eleventh International Database Engineering & Applications Symposium (IDEAS)*, 2007.
- 14) Qihong Shao, Michel Kinsy and Yi Chen: "Storing and Discovering Critical Workflows from Log in Scientific Exploration." In *Proceedings of the 2007 IEEE International Workshop on Scientific Workflows (SWF)*, 2007.
- 15) Subbarao Kambhampati, Garrett Wolf, Yi Chen, Hemal Khatri, Bhaumik Chokshi, Jianchun Fan, and Ullas Nambiar. "QUIC: A System for Handling Imprecision & Incompleteness in Autonomous Databases." In *Proceedings of 3rd Biennial Conference on Innovative Data Systems Research (CIDR)*, 2007
- 16) Te Li, Qihong Shao, and Yi Chen. "PEPX: A Query-Friendly Probabilistic XML Database." In *Proceedings of 15th ACM Conference on Information and Knowledge Management (CIKM)*, 2006.
- 17) Natalia Kwasnikowska, Yi Chen, and Zoe Lacroix. "Modeling and Storing Scientific Protocols." In *Proceedings of International Workshop on Knowledge Systems in Bioinformatics (KSinBIT)*, 2006.
- 18) Yi Chen, Susan B. Davidson, and Yifeng Zheng. "An Efficient XPath Query Processor for XML Streams." In *Proceedings of 22nd International Conference on Data Engineering (ICDE)*, 2006.
- 19) Steven Bird, Yi Chen, Susan B. Davidson, Haejoong Lee, and Yifeng Zheng. "Designing and Evaluating an XPath Dialect for Linguistic Queries." In *Proceedings of 22nd International Conference on Data Engineering (ICDE)*, 2006.

- 20) Yi Chen, Susan B. Davidson, and Yifeng Zheng. "ViteX: A Streaming XPath Processing System." Demonstration description. In *Proceedings of 21st International Conference on Data Engineering (ICDE)*, 2005.
- 21) Steven Bird, Yi Chen, Susan B. Davidson, Haejoong Lee, and Yifeng Zheng. "Extending XPath to Support Linguistic Queries." In *Proceedings of Programming Language Technologies for XML (PLAN-X)*, 2005.
- 22) Yi Chen, Susan B. Davidson, and Yifeng Zheng. "BLAS: An Efficient XPath Processing System." In *Proceedings of 23rd ACM SIGMOD International Conference on Management of Data*, pp. 47-58, 2004.
- 23) Yi Chen, George A. Mihaila, Susan B. Davidson, and Sriram Padmanabhan. "EXPedite: A System for Encoded XML Processing." In *Proceedings of 13rd ACM Conference on Information and Knowledge Management (CIKM)*, pp. 108-117, 2004.
- 24) Yi Chen, George A. Mihaila, Sriram Padmanabhan, and Rajesh Bordawekar. "L-Tree: A Dynamic Labeling Structure for Ordered XML Data." In *Proceedings of International Workshop on Database Technologies for Handling XML Information on the Web (dataX)*, in conjunction with *EDBT*, 2004. Springer Lecture Notes in Computer Science 3268, pp. 209-218, 2004.
- 25) Yi Chen, George A. Mihaila, Susan B. Davidson, and Sriram Padmanabhan. "Efficient Path Query Processing on Encoded XML." In *Proceedings of International Workshop on High Performance XML Processing*, in conjunction with *WWW*, 2004.
- 26) Yi Chen, Susan B. Davidson, Carmem Hara, and Yifeng Zheng. "RRXS: Redundancy Reducing XML Storage in Relations." In *Proceedings of 29th International Conference on Very Large Data Bases (VLDB)*, pp. 189-200, 2003.
- 27) Yi Chen, Susan B. Davidson, and Yifeng Zheng. "XKvalidator: A Constraint Validator for XML." In *Proceedings of 11th ACM Conference on Information and Knowledge Management (CIKM)*, pp. 446-452, 2002.
- 28) Yi Chen, Susan B. Davidson, and Yifeng Zheng. "Constraint Preserving XML Storage in Relations." In *Proceedings of 5th International Workshop on the Web and Databases (WebDB)*, in conjunction with *SIGMOD*, pp. 7-12, 2002.

OTHER PUBLICATIONS

- 29) Susan D. Urban, Suzanne W. Dietrich, and Yi Chen. "An XML Framework for Integrating Continuous Queries, Composite Event Detection, and Database Condition Monitoring for Multiple Data Streams." In *Dagstuhl Online Proceedings of the Event Processing Seminar*, 2007.
- 30) Yi Chen. "Data Quality and Provenance in Information Integration." Position paper. In *workshop on Information Integration*, 2006.

TEACHING EXPERIENCE

- CSE 340 "Principles of Programming Languages": Spring 2006, Fall 2006, Fall 2007

- CSE412/598, CBS598 “Database Management”: Spring 2008, Fall 2008
- CSE 591, “Data on the Web”: Fall 2005, Spring 2007, Fall 2007

INVITED TALKS

- “Keyword Search on XML Data and Beyond”, *University of California, Santa Barbara*, Nov 10, 2008;
- “What Users Want in XML Keyword Search”, *Hong Kong University of Science and Technology*, May 14, 2007;
- “Efficiently Querying Tree-structured Data”, *University of Arizona*, September 21, 2006;
- “Efficient Query Processing on XML Streams”, *University of California, Los Angeles*, May 19, 2006;
- “Efficient Query Processing on XML Streams”, *University of California, San Diego*, April 21 2006.
- “Enhancing Traditional Databases to Support Scientific Data Management Applications.” *ASU Math Dept*, January 23, 2006.
- “Managing XML Data Effectively in Relational Databases” *Bell Labs*, October 28, 2005.
- “Constraint Preserving XML Storage in Relations” *IBM T.J. Watson Research Center*, 2002.

RESEARCH GRANTS

2009 National Science Foundation (NSF). IIS-0845647

“CAREER: Analyzing and Exploiting Meta-information for Keyword Search on Semi-structured Data”
\$446,659. 3/2009 - 2/2014 (sole PI)

2007 National Science Foundation (NSF). IIS-0740129.

“Enabling Effective Access to Scientific Workflows”
\$87,370, 9/2007 – 8/2009. (sole PI)

2007 Science Foundation Arizona (SFAZ)

“Generalized Text Extraction from Life Science and Biomedicine Abstracts: empowering the
CBioC Mass Collaborative Curation and Reasoning Systems”
\$138,749, 3/2007 – 2/2008 (one of four PIs)

2006 National Science Foundation (NSF). IIS-0612273.

“ProtocolDB: Archiving and Querying Scientific Protocols, Data and Provenance”
\$479,996, 8/2006 – 7/2009 (one of two PIs)

PROFESSIONAL SERVICE

- Program Committee co-Chair, 1st workshop on Keyword Search on Structured Data (KEYS), collocated with SIGMOD 2009.
- Publicity and Proceeding Chair, 28th ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems (PODS), 2009
- Program Committee Member:
 - 2009 International Conference on Very Large Data Bases (VLDB)
International Conference on Data Engineering (ICDE)
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)
 - 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 (associated with VLDB)
10th 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, associated 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, associated with VLDB)
- Reviewer
 - Journals: ACM Transactions on Database Systems, IEEE Transactions on Knowledge and Data Engineering, Data & Knowledge Engineering, Information Systems, VLDBJ
 - Conferences: SIGMOD 2005, PODS 2005, ICDE 2005, SIGMOD 2004, PODS 2004; ICDE 2004, CIKM 2004, XSym 2004; WebDB 2004, SIGMOD 2003;
- Others:
 - Grace Hopper Scholarship Reviewer, 2007, 2008
 - NSF panel, 2007, 2009