

# Susan B. Davidson

## *Curriculum Vitae*

susan@cis.upenn.edu  
<http://www.cis.upenn.edu/~susan/home.html>  
Tel: (215) 898-3490, Fax: (215) 898-0587

Department of Computer and Information Science  
University of Pennsylvania  
Philadelphia, PA 19104-6389

### Education

- **Ph.D. Computer Science**  
Princeton University, October, 1982.  
Dissertation title: *An Optimistic Protocol for Partitioned Distributed Database Systems*
- **M.S.E., M.A. Computer Science**  
Princeton University, October, 1982.
- **B.A. Mathematics**  
Cornell University, College of Arts and Sciences, June, 1978.  
Graduated *Cum Laude* in Mathematics with distinction in all subjects.

### Employment

- University of Pennsylvania: Professor of Computer and Information Science, 1998-present; Secondary, Genetics Department (School of Medicine) 2000-2003.
- University of Pennsylvania: Associate Professor 1989-98.
- University of Pennsylvania: Assistant Professor 1983-1989.
- University of Pennsylvania: Visiting Assistant Professor 1982.

### Recent Leadership Positions and Awards:

- Chair, Board of Directors, CRA (2015-2017)
- Department Chair (2008-2013).
- Deputy Dean, SEAS (2005-2007).
- Founder and Chair, Advancing Women in Engineering (SEAS) (2007-present).
- Co-Director, Center for Bioinformatics, University of Pennsylvania (1997-2003).
- Co-director, Greater Philadelphia Bioinformatics Alliance (2003-2006).
- ACM Fellow (November 2001).
- Corresponding Fellowship of the Royal Society of Edinburgh (March 2015).
- Member, Computing Research Association Board of Directors (2011-2017)

- Member, Computing Community Consortium (2012-2015)
- NSF CISE Advisory Committee (2012-2016).
- External Advisory Board, Department of Computer Science, Princeton University (June 2012-2015).
- External Advisory Board, Department of Computer Science, Lehigh University (June 2010-present).
- External Advisory Board, Department of Computer Science, Johns Hopkins University (September 2016).
- Fulbright Scholar, Hitachi Chair (INRIA, France) (January - June 2004).
- Steering Committee, VLDB Association (2008-2014).
- Program Chair, 2015 International Conference on the Management of Data (SIGMOD'15).
- General Chair, 1999 International Conference on the Management of Data (SIGMOD'99).
- Trustees' Council of Penn Women/Provost Award (April 2015)
- ACM Service Award for performance as General Chair of 3rd International Conference on Parallel and Distributed Information Systems (1994), and for General Chair of SIGMOD99.
- Lenore Rowe Williams Award (AWFA) (April 2002).
- Member, National Research Council Committee on Engaging the Computer Science Research Community in Health Care Informatics (2007-2008).
- Member, President's Council of Advisors on Science and Technology (PCAST) Networking and Information Technology (NIT) Technical Advisory Group (TAG) (2006- 2007).
- Invited by the British Research Council to be a member of an international team to review the status of Computer Science Research in the UK (June 2001, December 2006).

### **Research Interests**

- *Databases*: data models and languages, complex object systems, semistructured data and XML, database transformations, data integration, data warehousing, transactions and updates, constraints, data cleaning, data reconciliation, data citation.
- *Bioinformatics*: biomedical databases, workflow systems, data provenance.

## University Experience

### Postdoc Supervision

- Sarah Cohen-Boulakia (2005-2007), Associate Professor (Maitre de conférences), Laboratoire de Recherche en Informatique (LRI), Univ. Paris-Sud.
- Anat Eyal (2007), Program Director, XIV SW Development at IBM
- Julia Stoyanovich (2009-2011 Computing Innovation Fellow), Assistant Professor of Computer Science, Drexel University.

### Ph.D. Dissertation Supervision

- Magdi Kamel, *Efficient Enforcement of Integrity Constraints Using Redundant Data* (Ph.D., August 1988). Full professor at the Naval Postgraduate School in Monterey.
- Victor Wolfe, *Language Constructs for Distributed Real-Time Programming* (Ph.D., August 1991). Professor, Department of Computer Science and Statistics, University of Rhode Island (co-advised with Insup Lee).
- Aaron Watters, *Constructive Deductive Databases and Partial Information* (Ph.D. August 1992). First Computer Science recipient of an AT&T fellowship. Initial appointment at AT&T (Susan Davidson, primary advisor; Peter Buneman, co-advisor).
- Anthony Kosky, *Transforming Databases with Recursive Data Structures* (December 1995). Initial appointment as research scientist at Lawrence Berkeley Labs (Susan Davidson, primary advisor; Peter Buneman, co-advisor).
- Hanene Ben-Abdellah, *GCSR: A Graphical Language for the Specification, Refinement and Analysis of Real-Time Systems* (August 1996). Professor, Département d'Informatique Faculté des Sciences Economiques et de Gestion, Université de Sfax, Tunisia (Co-advised with Insup Lee).
- Carmem Hara, *Reasoning about Functional and Key Dependencies in Hierarchically Structured Data* (December 2004). Professor, Departamento de Informatica, Universidade Federal do Parana (Susan Davidson, primary advisor; Wenfei Fan, co-advisor).
- Vanessa Braganhola, *From XML to Relational View Updates: Applying Old Solutions to Solve a New Problem* (December 2004). Currently an Associate Professor at Universidade Federal Fluminense (Co-advised with Carlos Heuser at Universidade Federal do Rio Grande do Sul).
- Yi Chen *Efficient XPath Query Processing on Stored and Streaming XML Data* (August 2005). Associate Professor, Henry J. Leir Chair in Healthcare, New Jersey Institute of Technology.
- Yifeng Zheng, *Efficient Scientific Data Management over Trees* (August 2006). Currently at Google (NYC).

- Zhuowei Bao, *Efficiently Tracking Provenance in Scientific Workflows* (August 2012). Currently at Facebook.
- Sudeepa Roy, *Data Provenance and Uncertainty* (August 2012). Assistant Professor, Department of Computer Science, Duke University (Co-advised with Sanjeev Khanna.)
- Xiaocheng Huang, (Advisor, Xiaojie Yuan, Nankai University; spent 2 years at Penn.) Currently at the Genome Institute of Singapore.

### Master's Thesis Supervision

- Mark Winkler, *Performance Analysis of Optimistic and Pessimistic Partition Failure Protocols* (July 1985).
- Ronald Reuss, *Byzantine Agreement and Transaction Commit: A Comparison* (August 1987).
- Bob Greenawalt, *Real-Time Databases: Case Studies and Analysis*, (August 1988). Currently on the faculty at West Point Academy.
- Hartmut Liefke, *Extending CPL/Kleisli With Updates* (May 1997).
- Deborah Pinney, *Integration of RNA and Protein in the GUS Database* (May 2002).

### University Service

- **University Committees:** Senate Executive Committee (1987-89), Information Center Committee (1983), University Minor in Computer Science Committee (Spring 1994), Master's of Biotechnology executive committee (1997), Genomics Institute executive committee (2001-present), Gender Equity Committee (2000), Planning Committee for Life Sciences (2001), SEPA Bioinformatics Greenhouse Committee(2002), Genomics Institute Review Committee (2005-6), Vice Provost for Research Search Committee (2006), Genomics Institute Oversight Committee(2005-7), Academic Planning and Budget Committee (2010-2013).
- **School Committees:** Advancing Women in Engineering (2007-present), Personnel Committee (2004), Undergraduate Affairs Committee (1990-1992), Academic Performance Committee (1985-1990), "Forty years of computing at Penn" Committee (1985-86), Pender Award Committee (1985), Library Committee (1985), Faculty Council (Chair, 1998-2000).
- CIS Undergraduate Group Chair (1990-92): Major accomplishments: revision of CSE-BSE program, significant rejuvenation of the Dining Philosophers, establishment of the X-lab through numerous grants (internal as well as external).
- SEAS Deputy Dean (2004-7): Major accomplishments include creating a program for recruiting and retaining women undergraduates in engineering (Advancing Women in Engineering Oversight Committee, AWE); creating a postdoc association (ePOD); creating a mentoring policy and holding mentoring meetings; creating a bridge funding policy; and forming a Computing Committee to evaluate the state and needs of research computing within SEAS.

- PhD and Postdoc Program in Computational Biology: participated in NSF Research Training Grant for establishing a multi-disciplinary research training program in computational biology. Much of the success in obtaining this grant was based on existing active collaborations between the database research group (P. Buneman, S.B. Davidson and V. Tannen) and the informatics group in support of the Human Genome Project Center for Chromosome 22 (C. Overton and D. Searls). More recently, have helped established a PhD program in the Biomedical Graduate Studies program for Genomics and Computational Biology (R. Spielman, D. Roos, W. Ewens).
- Masters of Biotechnology, track in Computational Biology/Bioinformatics: administered this track for three years in conjunction with Dan Hammer and then Scott Diamond, directors of the Masters of Biotechnology program. Responsibilities included admissions as well as advising all students.
- BAS in Computational Biology: designed and am administering an SEAS undergraduate degree in computational biology. Also participated with creating a sister degree in SAS within Biology.
- Genomics and Computational Biology degree program (track within Biomedical Graduate Studies): founding member.
- Center for Bioinformatics: C. Overton, S. Davidson ( Interim-Director), P. Buneman and V. Tannen, established a center for the organization, analysis, and visualization of biomedical information through the Genomics Institute with co-sponsorship by the Institute for Research in Cognitive Science (IRCS). It is a major initiative which has successfully recruited two faculty members with primary appointments in Genetics, SOM (Artemis Hatzegeorgiou and Sridhar Hannenhall) and one faculty member with a primary appointment in Biology, SAS (Juhnjong Kim). W. Ewens (Biology- SAS) led up the educational component of the center.
- Greater Philadelphia Bioinformatics Alliance (GPBA): Charged by Provost Barchi to head a collaborative proposal to BioAdvance involving nine area universities and research institutions (U. Penn., Drexel, Temple University, CHOP, WISTAR, Fox Chase Cancer Institute, Thomas Jefferson University, Penn State Great Valley and University of the Sciences in Philadelphia. The proposal was funded for \$2.4M over three years, and was used to support a yearly retreat and regional seminar series, graduate student internships, and short courses. The group also successfully attracted an additional \$500,000 of state money from Harrisburg, funding from NSF Partnerships for Innovation, as well as the Department of Labor for \$3.4M over 3 years with another \$3.4 of matching funds.

## Professional Experience

### Professional Activities

See also leadership positions and awards.

- Associate Editor, ACM Transactions on Database Systems (2009-2012).
- Associate Editor, Journal of Computational Biology (1997-present).
- Steering Committees: Testing and Evaluation for ONR's Real-Time System Initiative (Summer 1988), Conference on Parallel and Distributed Information Systems (1995), Bioinformatics Workshop (National Research Council, February 2000).
- General Chair: Real-Time Systems Symposium (with I. Lee, December 1993), 3<sup>rd</sup> International Symposium on Parallel and Distributed Information Systems (September 1994), Data Integration in the Life Sciences (June 2007).
- Organizer: ARO Workshop on Integrating Heterogeneous Information Systems (March 8-9, 1993); First Greater Philadelphia Bioinformatics Alliance Research Retreat (Oct. 24, 2003); Workshop on Information Integration (Oct. 26-27, 2006).
- Program Chair: Real-Time Systems Symposium (with I. Lee, December 1992).
- Area Chair: International Conference on Data Engineering 2010, SIGMOD 2010.
- Tutorials Chair: International Conference on Very Large Database 2000, 2014.
- Workshop co-Chair: International Conference on Very Large Database 2008.
- Awards Committee, Chair: International Conference on Very Large Database 2009, 2014.
- Program Committees: IEEE Symposium on Reliable Distributed Systems (October 1989), Real-Time Systems Symposium (December 1989), 10<sup>th</sup> International Conference on Distributed Computing Systems (May 28-June 1, 1990), Conference on Parallel and Distributed Information Systems (December 1991) Conference on Reliable Distributed Systems (October 1994), 3<sup>rd</sup> International Conference on Information and Knowledge Management (November 1994), 15<sup>th</sup> International Conference on Distributed Computing Systems (June 1995), Meeting on Interconnecting Molecular Biology Databases (August 1994, July 1995), International Conference on Very Large Databases (September 1995, 1997, 1999, 2001, 2008), International Semantics Web Conference (2003), SIGMOD (May 1996, 2003, 2005, 2007), ISMB (2005), Database Programming Languages (August 1997), Data Engineering (1997, 1998), Brazilian Symposium on Databases (1998, 2003), PODS (2000), International Provenance and Annotation Workshop (2014, 2016), Theory and Principles of Provenance (2014).
- Publicity Chair: Joint Conference ACM SIGMOD International Conference on Management of Data and Principles of Database Systems '93 (May 1993).
- Grant Panels: NSF (1988, 1993, 1995-7, 1999, 2000, 2012), NIH (1999), Sloan/DOE Bioinformatics Fellowships (2002-3).

- Panels: “Scientific Data Management: An Orphan in the Database Community?” ICDE (April 2008). “Scaling the Tower of Babel: Data Integration and Warehousing with the WWW,” Conference on Information Technology Applications in Biomedicine (November 2000), “Time Constrained Communication,” 6’t Workshop on Real-Time Operating Systems (May 1989), “Practicality of Distributed System Design Tools,” 9’t International Conference on Distributed Computing Systems (June 1989), “Issues in Application Development,” Workshop on Operating Systems for Mission Critical Computing, College Park MD (Sept. 19-21, 1989).
- Professional Consulting: General Electric Aerospace (1993), Smith-Kline Beecham (1995-96).

### Journal and Book Reviewing

- ACM: *Communications of the ACM*, *Transactions on Database Systems*, *Transactions on Programming Languages and Systems*.
- IEEE: *Transactions on Computers*, *Transactions on Software Engineering*, *Transactions on Knowledge and Data Engineering*.
- *Journal of Parallel and Distributed Computing*, *Information Processing Letters*
- NIH, NSF and ARO Grant Proposals.
- Texts in algorithms and data structures and in database systems for Reston Publishing Company, and Allyn and Bacon Publishing Company.

## Research Experience

### Current Research Grants

- NSF CICI: Data Provenance: Provenance-Based Trust Management for Collaborative Data Curation (Zack Ives, PI; Sampath Kannan, Val Tannen, Susan B. Davidson, co-PI) \$500,000.00 (Sept. 1, 2015- Aug. 31, 2018).
- NSF III Medium: Data Citation (Susan B. Davidson, PI; Peter Buneman, Val Tannen, co-PI; Jim Frew, PI UCSB; Wenfei Fan, Senior Researcher) \$850,000 (Aug. 1, 2013- July 31, 2016).
- NIH 1 U01 EB020954-01: Approximating and Reasoning about Data Provenance (Zack Ives, PI; S. Kannan, J. Kim, S.B. Davidson co-PI) \$1,474,964 (June 1, 2015- May 31, 2018)

### Completed Grants

- NSF IIS-0803524: III-COR-Medium: Providing Provenance through Workflows and Database Transformations (Susan B. Davidson, PI; Sanjeev Khanna, co-PI; Val Tannen, co-PI), \$816,600.
- NSF BPC: Penn COMP-ACT: A College Service Learning Course to Promote and Increase COMPUTational Thinking and ACTivities in Afterschool and Summer Programs (Susan Davidson, PI; Jean Griffin, Yasmin Kafai, Rita Powell and Joe Sun, co-PI), \$600,000.

- NSF IIS RI: Medium: New Tools and Methods for Very-Large-Scale Phonetics Research (Jiahong Yuan, PI; Andreas Stolcke, Mark Liberman, Susan Davidson, co-PI), \$450,000.
- Google Research Award: Identifying Ranked Agreement among Raters (Susan B. Davidson and Julia Stoyanovich, PI) \$10,000 (January 1, 2012- December 31, 2012).
- NSF ATOL “Core Database Technologies to Enable the Integration of AToL Information” (Val Tannen, PI; Susan Davidson, co-PI, Zachary Ives, co-PI) \$1,750,000.00.
- NSF SEIII “ProtocolDB: Archiving and Querying Scientific Protocols, Data and Provenance” (Zoe Lacroix, PI; Yi Chen, co-PI; Susan Davidson, co-PI) \$323,323.00.
- NSF IDM, “Preserving Constraints in XML Data Exchange,” (Susan Davidson, PI). \$305,000 (June 1, 2005- May 31, 2008).
- NSF SEII “II: Data Cooperatives: Rapid and Incremental Data Sharing with Applications to Bioinformatics” (Susan Davidson, PI; Zachary Ives, Christian J. Stoeckert, Val Tannen and Pete White, co-PI’s). \$1,295,278 (\$22,943 subcontract to CHOP), July 1, 2005 - June 30, 2008.
- BioAdvance/FoxChase, “Database Fusion and Advanced Mining Tools in Biomedical Research” (subcontract) \$116,836.
- AHRQ “Workshop on Information Integration” (Susan Davidson, PI) 05/05/2006-05/04/2007, \$10,000.00.
- ONR “Workshop on Information Integration” (Susan Davidson, PI) 05/05/2006-05/04/2007, \$10,000.00.
- NSF SEIII “Workshop on Information Integration” (Susan Davidson, PI; Zachary Ives, co-PI) 05/05/2006-05/04/2007, \$76,276.00.
- NSF (Cognitive, Psychological & Language Science), “Querying Linguistic Databases,” (Mark Liberman, PI, Michael Maxwell, Steven Bird, Beatrice Santorini, and Susan Davidson, co-PI) \$148,293 (August 1, 2003- July 31, 2006).
- NIH pre-NPEBC, “Comparative Approaches to Bio-Knowledge Discovery” (Junhyong Kim, PI; S. Davidson, co-PI). \$1,000,000 (Sept. 1, 2003 to Aug. 31, 2006).
- BioAdvance, “Philadelphia Greenhouse Consortium for Bioinformatics: Collaborative Research, Education and Training” (Collaboration with the University of Pennsylvania, Drexel University, Temple University, University of the Sciences, Children’s Hospital, WISTAR Institute, Fox Chase Cancer Research Center, Thomas Jefferson University). \$2,319,400 (June 1, 2003- April 30, 2006).
- NSF Partnership for Innovation “Philadelphia Partnership for Innovations in Biotechnology,” (Aydin Tozerin, PI; S. Davidson, co-PI; Gary KGurtzman, co-PI). \$320,000 (\$150,000 subcontract to UPenn) (November 1, 2003- October 31, 2005).

- NSF ITR, “Mining the Bibliome – Information Extraction from the Biomedical Literature.” (Aravind Joshi, PI., Mark Liberman, Susan Davidson co-PI) \$4,976,335 (Sept. 1, 2002 - Aug. 31, 2005).
- NSF (Database Activities in the Biological Sciences), “Refreshing Curated Warehouses,” (Susan Davidson, PI; Peter Buneman, Chris Overton, co-PI) \$561,769 (10/1/1999-9/30/2003).
- NSF (Digital Libraries Initiative), “Data Provenance,” (Peter Buneman, PI; Susan Davidson, Val Tannen, co-PI) \$504,990 (6/1/1999-5/31/2003).
- NIH NCCR (Facilities Renovation), “Center for Bioinformatics” (William Kelley, PI; Susan Davidson and Chris Overton, ghostwriters) \$692,000.
- ARO “Engineering Robust Distributed Databases,” (Val Tannen, PI; Susan Davidson, Peter Buneman, co-PI) \$270,000 (6/1/98-6/30/01).
- DARPA, “Ultra-High Speed Optical Network Architecture,” (Davidson, PI on a subcontract from Jonathan Smith) \$300,000 (1/1/2000-12/31/2001).
- SmithKline Beecham, “Tools for Integrating Biological Data Sources,” (V. Tannen, PI; S. Davidson, C. Overton, P. Buneman, co-PI) \$542,588 (7/1/1996-12/31/2000).
- DOE, “Schema Transformations with Biological Databases” (with Chris Overton and Peter Buneman), \$1,225,799 (September 1, 1995 - August 31, 1998).
- ARPA, “Transformation Languages for Heterogeneous Data Sources” (with Peter Buneman, Val Tannen and Larry Alexander), \$400,000 (August 1, 1994 – Feb 28, 1998).
- ARO, “Extending Database Integration Technology” (with Peter Buneman, Val Tannen, and Chris Overton), \$495,000 (May 1, 1995 – April 30, 1998).
- NSF-CISE (Planning and Equipment Grant), “Telementoring: A Novel Approach to Teaching Undergraduate Computer Scientists” (with Jonathan Smith, David Farber, Insup Lee and Ira Winston), \$400,000 (August 1, 1992- July 31, 1995).
- Sun Academic Equipment Grant, “Telementoring: A Novel Approach to Teaching Undergraduate Computer Scientists” (with Jonathan Smith, David Farber, Insup Lee and Ira Winston), \$85,000 (June 1, 1993 - May 31, 1996).
- NSF-ILI (Equipment Grant), “Teleconferenced Workstations: Improving Experimentation in Undergraduate Education” (with Insup Lee and Jonathan Smith), \$76,000 (July 1, 1994- June 30, 1996).
- NSF, “Mediated Access to Biological Databases and Applications” (PI, with Chris Overton and Peter Buneman) \$200,000, (July 1, 1994- June 30, 1996).
- AT&T, “Study of Protocol Verification Issues for DMI” (with Insup Lee), \$175,000 (Sept. 1986- August 1989).
- AT&T, “A Programming System for Heterogeneous Distributed Databases” (with Peter Buneman), \$175,000 (Sept. 1986- August 1988).

- ARO, “Distributed Expert Systems” (with Tim Finin, Insup Lee, Lokendra Shastri, and Peter Buneman), \$75,000 (Sept. 1986 - June 1987).
- NSF Equipment Grant, “A Distributed Testbed for Real-time Active Sensing” (with Insup Lee and K. Wohn), \$29,305 (May 1988- Oct. 1989).
- ONR, “Formally Integrating Real-Time Specification and Implementation” (with Insup Lee), \$408,000 (Oct. 1988- Sept. 1991).
- ARO, “New Techniques for Merging Database Schemas” (with Peter Buneman), \$65,988 (May 1, 1991- April 30, 1993).
- ONR, “Formally Integrating Real-Time Specification and Implementation” (with Insup Lee), \$258,928 (Oct. 1991- Sept. 1993).
- NSF, “A Formal Approach to Real-Time Specification and Analysis” (with Insup Lee), \$407,961 (June 1991- Nov. 1994).
- University of Pennsylvania Undergraduate Initiatives Fund (Equipment Grant), “A Laboratory For Visualizing Computing Paradigms,” (Project Director, with Peter Buneman and Insup Lee), \$41,282 (September 1991- June 1992).
- University of Pennsylvania Undergraduate Initiatives Fund (Equipment Grant), “A Meta Language Approach to Introductory Programming,” (with Peter Buneman, Val Tannen, Carl Gunter, Jonathan Smith and Ira Winston), \$ 41,143 (September 1992- June 1993).

**Invited Talks and Short Courses (2005-present):**

- “Data Citation: a Computational Challenge”: CRA Headquarters (October 2015); NYU Tandon School of Engineering (November 2015); Keynote IPAW2016 (June 2016).
- “Query Driven Crowd Mining”: IC Colloquium, EPFL, Switzerland (September 15, 2014); Laboratoire d’Informatique de Grenoble, Centre National de Recherche Scientifique, Grenoble, France (November 3, 2014).
- “Clouds, Crowds and Provenance,” Keynote IEEE ICWS, Cloud, BigData 2014 (June 2014).
- “Using the Crowd for Group-by and Top-k Queries,” Tel Aviv University (October 2014).
- “Querying Workflow Provenance,” WebDAM Workshop (Sept. 2014).
- “Crowd Data Mining,” University of Buffalo, Distinguished Lecture (April 2014).
- “Hiding Data and Structure in Workflow Provenance,” DNIS (December 2011).
- “On provenance and privacy,” Keynote ICDT (2011); Keynote IPAW (2010).
- “Provenance and Scientific Workflows: Challenges and Opportunities,” with Juliana Freire. Tutorial, SIGMOD (June 3, 2008).
- “Provenance and Scientific Workflows,” Spring ’08 DB/IR Day, Columbia University (April 17, 2008).

- “On Provenance and User Views in Scientific Workflows,” University of Utah (September 17, 2007).
- “Querying and Managing Provenance through User Views in Scientific Workflows,” University of the Sciences in Philadelphia (March 27, 2007).
- “Querying and Managing Provenance through User Views in Scientific Workflows,” New England Database Society (May 18, 2007).
- “Biological/scientific data errors”, MS Research/U Washington Summer Institute on Managing Imprecise Information, August 2005.
- “Data Cooperatives: Rapid and Incremental Data Sharing in Bioinformatics”. OGSB’05 ”Ontology, Grid and Semantic Integration for Biology” (workshop of JOBIM’05, the French national conference on Bioinformatics), July 2005.

## Publications

**NOTE:** With very few exceptions, I have had a policy of publishing papers with author names listed in alphabetical order.

### Refereed Journals

1. “Why Data Citation is a Computational Problem,” by Peter Buneman, Susan Davidson, and James Frew. *Communications of the ACM* (accepted).
2. “Effective and efficient similarity search in scientific workflow repositories,” by Johannes Starlinger, Sarah Cohen Boulakia, Sanjeev Khanna, Susan B. Davidson, and Ulf Leser. *Future Generation Comp. Syst.* 56: 584-594 (2016).
3. “Top-k and Clustering with Noisy Comparisons,” by Susan Davidson, Sanjeev Khanna, Tova Milo and Sudeepa Roy. *ACM Trans. Database Syst.* 39:4: 35:1-35:39 (2014).
4. “Generating sound workflow views for correct provenance analysis,” by Ziyang Liu, Susan B. Davidson, and Yi Chen. *ACM Trans. Database Syst.* 36(1): 6 (2011).
5. “A Bi-labeling Based XPath Processing System,” by Yi Chen and Susan B. Davidson. *Information Systems Journal* 35(2): 170-185 (2010).
6. “Addressing the Provenance Challenge using ZOOM,” by Sarah Cohen-Boulakia, Shirley Cohen, and Susan B. Davidson. *Concurrency and Computation: Practice and Experience* 20(5):497 - 506, Special Issue : The First Provenance Challenge (April 2008), pp. 497-506.
7. “Provenance in Scientific Workflow Systems,” by Susan B. Davidson, Sarah Cohen-Boulakia, Anat Eyal, Bertram Ludäescher, Timothy McPhillips, Shawn Bowers, Manish Kumar Anand, and Juliana Freire. *IEEE Data Engineering Bulletin* 30(4): 44-50, December 2007.
8. “BioGuideSRS: Querying Multiple Sources with a user-centric perspective,” by Sarah Cohen-Boulakia, Susan B. Davidson and Christine Froidevaux. *Bioinformatics* 2007 23(10):1301-1303.
9. “Path-based systems to guide life scientists in the maze of biological data sources”, by Sarah Cohen-Boulakia, Susan B. Davidson, Christine Froidevaux, Zoe Lacroix and Maria-Esther Vidal. *Journal of Bioinformatics and Computational Biology* 4:5 (October 2006).
10. “PATA XO: a framework to allow updates through XML views”, by Vanessa P. Braganholo, Susan B. Davidson and Carlos A. Heuser. *ACM Transactions on Database Systems* 13:3 (Sept. 2006), pp. 839 - 886.
11. “Propagating XML Constraints to Relations”, by C. Hara, W. Fan and S. Davidson. *Journal of Computer and System Sciences* 73:3 (May 2007), pp. 316-361.
12. “Digital library information-technology infrastructures”, by Yannis E. Ioannidis, David Maier, Serge Abiteboul, Peter Buneman, Susan B. Davidson, Edward A. Fox, Alon Y. Halevy, Craig A. Knoblock, Fausto Rabitti, Hans-Joerg Schek, and Gerhard Weikum. *International Journal on Digital Libraries* 5(4): 266-274 (2005).

13. "Reasoning about Keys for XML," by P. Buneman, S. Davidson, W. Fan, C. Hara and W.C. Tan. *Information Systems* 28:8, 1037-1063 (2003).
14. "Keys for XML," by P. Buneman, S. Davidson, W. Fan, C. Hara, and W.C. Tan. *Computer Networks* 39:5 (August 2002, Elsevier Publishers), pp. 473-487. N.B. Best Author Award.
15. "K2/Kleisli and GUS: Experiments in Integrated Access to Genomic Data Sources," by S. Davidson, J. Crabtree, B. Brunk, J. Schug, V. Tannen, C. Overton and C. Stoeckert. *IBM Systems Journal* 40:2, pp. 512-531 (2001).
16. "Semantics of Database Transformations," A. Kosky, S.B. Davidson and P. Buneman. In *Semantics of Databases*, edited by L. Libkin and B. Thalheim, Springer LNCS 1358 (February 1998).
17. "BioKleisli: A Digital Library for Biomedical Researchers," S.B. Davidson, C. Overton, V. Tannen and L. Wong. *Journal of Digital Libraries* 1:1 (November 1996), pp. 36-53.
18. "Challenges in Integrating Biological Data Sources," S.B. Davidson, C. Overton and P. Buneman. *Journal of Computational Biology* 2:4 (Winter 1995), pp 557-572.
19. "Deadlock Prevention in Concurrent Real-time Systems," S.B. Davidson, I. Lee and V. Wolfe. *Real-Time Systems Journal*, 5:3 (October 1993), pp. 305-318.
20. "RTC: Language Support for Real-Time Concurrency," V. Wolfe, S.B. Davidson and I. Lee. *Real-Time Systems Journal*, 5:1 (March, 1993), pp. 63-87.
21. "A Semantics for Complex Objects and Approximate Queries", P. Buneman, S.B. Davidson and A. Watters. Invited paper to *Journal of Computer and System Sciences* 43:1 (August 1991), pp. 170-218. (See below under refereed conference papers.)
22. "Timed Atomic Commitment", S.B. Davidson, I. Lee and V. Wolfe. *IEEE Transactions on Computers* 40:5 (May 1991), pp. 573-583.
23. "A Performance Analysis of Timed Synchronous Communication Primitives," I. Lee and S.B. Davidson. *IEEE Transactions on Computers* 39:9 (Sept. 1990), pp. 1117-1131.
24. "Semi-Materialization: A Technique for Optimizing Frequently Executed Queries," S.B. Davidson and M. Kamel. *Data & Knowledge Engineering* 5 (1990), pp. 93-103. "Querying Independent Databases," P. Buneman, S.B. Davidson and A. Watters. *Information Sciences- an International Journal* 49:3 (November 1989).
25. "Federated Approximations in Heterogeneous Databases," P. Buneman, S.B. Davidson and A. Watters. Invited paper to *IEEE Office Knowledge Engineering*, 3:2 (August 1989), pp. 27-34.
26. "Adding Time to Synchronous Process Communications," I. Lee. and S.B. Davidson. *IEEE Trans. on Computers* C-36:8 (Special issue on Real Time Systems, Aug. 1987), pp. 941-948. (Reprinted in *IEEE Tutorial: Hard Real-Time Systems*, edited by John A. Stankovic and Krithi Ramamritham 1988).

27. "A Performance Comparison of Optimistic versus Conservative Strategies during Partition Failures in Distributed Databases," S.B. Davidson and M. Winkler. *Journal of Management and Information Systems* 4:2 (Fall 1987), pp. 113-133.
28. "Is Byzantine Agreement Useful in a Distributed Database?," S.B. Davidson, H. Garcia-Molina and F. Pittelli. *ACM Trans. on Database Sys.* 11:1 (March 1986), pp. 27-47.
29. "Consistency in a Partitioned Network: a Survey," S.B. Davidson, H. Garcia-Molina and D. Skeen. *ACM Computing Surveys* 17:3 (Sept. 1985), pp. 341-370. Also published in the Japanese journal *Bit*.
30. "Optimism and Consistency in Partitioned Distributed Database Systems," S.B. Davidson. *ACM Trans. on Database Sys.* 9:3 (Sept. 1984), pp. 456-481.

### Refereed Conference Proceedings

31. "Answering Regular Path Queries on Workflow Provenance," by Xiaocheng Huang, Zhuowei Bao, Susan B. Davidson, Tova Milo and Xiaojie Yuan. ICDE 2015: 375-386.
32. "Managing General and Individual Knowledge in Crowd Mining Applications," by Yael Amsterdamer, Susan B. Davidson, Anna Itin, Tova Milo, Slava Novgorodov, and Amit Somech. CIDR 2015.
33. "Approximated Summarization of Data Provenance," by Eleanor Ainy, Pierre Bourhis, Susan B. Davidson, Daniel Deutch, and Tova Milo. CIKM 2015: 483-492
34. Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data, by Timos K. Sellis, Susan B. Davidson, and Zachary G. Ives. May 31 - June 4, 2015, , Melbourne, Victoria, Australia. ACM 2015, ISBN 978-1-4503-2758-9.
35. "Layer Decomposition: An Effective Structure-based Approach for Scientific Workflow Similarity," by Johannes Starlinger, Sarah Cohen-Boulakia, Sanjeev Khanna, Susan Davidson, and Ulf Leser. IEEE e-Science Conference, Oct 2014.
36. "OASSIS: Query Driven Crowd Mining ," by Yael Amsterdamer, Susan B. Davidson, Tova Milo, Slava Novgorodov, and Amit Somech. SIGMOD 2015.
37. "Learning to Explore Scientific Workflow Repositories," by Julia Stoyanovich, Paramveer Dhillon, Brian Lyons and Susan B. Davidson. SSDBM 2013 (July 2013).
38. "Search and Result Presentation in Scientific Workflow Repositories," by Susan B. Davidson, Xiaocheng Huang, Julia Stoyanovich and Xiaojie Yuan. SSDBM 2013 (July 2013).
39. "Understanding Local Structure in Ranked Datasets," by Julia Stoyanovich, Sihem Amer-Yahia, Susan B. Davidson, Marie Jacob, and Tova Milo. CIDR 2013.
40. "A propagation model for provenance views of public/private workflows," by Susan B. Davidson, Tova Milo, and Sudeepa Roy. ICDT 2013: 165-176.
41. "Using the crowd for top-k and group-by queries," by Susan B. Davidson, Sanjeev Khanna, Tova Milo, and Sudeepa Roy. ICDT 2013: 225-236.

42. "Learning to explore scientific workflow repositories," by Julia Stoyanovich, Paramveer S. Dhillon, Susan B. Davidson, and Brian Lyons. SSDBM 2013: 31.
43. "A cascading mentoring pedagogy in a CS service learning course to broaden participation and perceptions," by Yasmin B. Kafai, Jean Griffin, Quinn Burke, Michelle Slattery, Deborah A. Fields, Rita Manco Powell, Michele Grab, Susan B. Davidson, and Joseph S. Sun. SIGCSE 2013: 101-106.
44. "Labeling Workflow Views with Fine-Grained Dependencies," by Zhuowei Bao, Susan B. Davidson, and Tova Milo: . PVLDB 5(11): 1208-1219 (2012).
45. "The reflective mentor: charting undergraduates' responses to computer science service learning (abstract only)," by Quinn Burke, Yasmin B. Kafai, Jean Griffin, Rita Manco Powell, Michele Grab, Susan B. Davidson, and Joseph S. Sun. SIGCSE 2012: 673.
46. "Putting Lipstick on Pig: Enabling Database-style Workflow Provenance," by Yael Amsterdamer, Susan B. Davidson, Daniel Deutch, Tova Milo, Julia Stoyanovich, and Val Tannen. PVLDB 5(4): 346-357 (2011).
47. "Enabling Privacy in Provenance-Aware Workflow Systems," by Susan B. Davidson, Sanjeev Khanna, Val Tannen, Sudeepa Roy, Yi Chen, Tova Milo, and Julia Stoyanovich. CIDR 2011: 215-218.
48. "Deriving probabilistic databases with inference ensembles," by Julia Stoyanovich, Susan B. Davidson, Tova Milo, and Val Tannen. ICDE 2011: 303-314.
49. "On provenance and privacy," by Susan B. Davidson, Sanjeev Khanna, Sudeepa Roy, Julia Stoyanovich, Val Tannen, and Yi Chen. ICDT 2011: 3-10.
50. "Provenance views for module privacy," by Susan B. Davidson, Sanjeev Khanna, Tova Milo, Debmalya Panigrahi, and Sudeepa Roy. PODS 2011: 175-186.
51. "Labeling recursive workflow executions on-the-fly," by Zhuowei Bao, Susan B. Davidson, and Tova Milo. SIGMOD Conference 2011: 493-504.
52. "An optimal labeling scheme for workflow provenance using skeleton labels," by Zhuowei Bao, Susan B. Davidson, Sanjeev Khanna, and Sudeepa Roy: . SIGMOD Conference 2010: 711-722.
53. "Optimizing user views for workflows," by Olivier Biton, Susan B. Davidson, Sanjeev Khanna, and Sudeepa Roy. ICDT 2009: 310-323
54. "Detecting and resolving unsound workflow views for correct provenance analysis," by Peng Sun, Ziyang Liu, Susan B. Davidson, and Yi Chen. SIGMOD Conference 2009: 549-562.
55. "PDiffView: Viewing the Difference in Provenance of Workflow Results," by Zhuowei Bao, Sarah Cohen-Boulakia, Susan B. Davidson, Anat Eyal, and Sanjeev Khanna. International Conference on Data Engineering (Shanghai, April 7-9, 2009).

56. "Querying and Managing Provenance through User Views in Scientific workflows," by Olivier Biton, Sarah Cohen-Boulakia, Susan B. Davidson and Carmem Hara. ICDE 2008, pp. 1072-1081.
57. "Designing and Evaluating an XPath Dialect for Linguistic Queries," by Steven Bird, Yi Chen, Susan B. Davidson, Haejoonyg Lee and Yifeng Zheng. International Conference on Data Engineering (Atlanta, April 4-6, 2006), p. 52.
58. "An Efficient XPath Query Processor for XML Streams," by Yi Chen, Susan B. Davidson and Yifeng Zheng. International Conference on Data Engineering (Atlanta, April 4-6, 2006), p. 79.
59. "Efficiently supporting structure queries on phylogenetic trees", by Susan B. Davidson, Junhyong Kim and Yifeng Zheng. International Conference on Scientific and Statistical Database Management (Santa Barbara, June 27-29, 2005), pp. 93-102.
60. "ViteX: a Streaming XPath Processing System", by Yi Chen, Susan B. Davidson and Yifeng Zheng. Demo paper in International Conference on Data Engineering (Japan, April 5-8, 2005), pp. 1118 - 1119.
61. "From XML View Updates to Relational View Updates: old solutions to a new problem", by V. Braganholo, S.B. Davidson and C. Heuser. Proc. International Conference on Very Large Databases (Sept. 2004), 276-287.
62. "BLAS: An Efficient XPath Processing System", by Yi Chen, Susan Davidson and Yifeng Zheng. Proceedings of SIGMOD2004 (June 2004), pp. 47-58.
63. "UXQuery: building updatable XML views over relational databases," by V. Braganholo, S.B. Davidson and C. Heuser. Proceedings of the Brazilian Symposium on Databases (October, 2003), pp. 26-40. Voted one of the top five papers in the conference.
64. "RRXS: Redundancy Reducing XML Storage in Relations," by Y. Chen, S.B. Davidson and Y. Zheng. VLDB2003 (September 2003), pp. 189-200.
65. "Propagating XML Constraints to Relations," by S.B. Davidson, W. Fan, C. Hara and J. Qin. ICDE2003, pp. 543-556.
66. XKvalidator: a constraint validator for XML," by Y. Chen, S.B. Davidson and Y. Zheng. CIKM2002 (November 2002), pp. 446-452.
67. "Reasoning about keys for XML," by P. Buneman, S. Davidson, W. Fan, C. Hara, and W.C. Tan. DBPL (Rome, September 2001), pp. 133-148.
68. "Keys for XML," by P. Buneman, S. Davidson, W. Fan, C. Hara, and W.C. Tan. WWW10 (Hong Kong, May 2001), pp. 201-210.
69. "View Maintenance for Hierarchical Semistructured Data," by H. Liefke and S.B. Davidson. DaWaK'00 (London, England, September 6-8, 2000).

70. "Specifying Updates in Biomedical Databases," by H. Liefke and S.B. Davidson. 11th International Conference on Scientific and Statistical Database Management (Cleveland, July 28-30 1999), pp. 44-53.
71. "Processing Updates on Complex Value Databases," by H. Liefke and S.B. Davidson. *International Resources Management Association International Conference* (Hershey, May 16-19 1999).
72. "Reasoning about Nested Functional Dependencies," C. Hara and S.B. Davidson. *Proceedings of the Symposium on Principles of Database Systems* (Philadelphia, May 31-June 2 1999), pp. 91-100.
73. "Querying an Object-Oriented Database Using CPL," S.B. Davidson, C. Hara and L. Popa. *Proceedings of the Brazilian Symposium on Databases* (Oct. 1997), pp. 137-153.
74. "WOL: A Language for Database Transformations and Constraints," S.B. Davidson and A. Kosky. *Proceedings of the International Conference of Data Engineering* (April 1997), pp. 55-65.
75. "Adding Structure to Unstructured Data," P. Buneman, S. Davidson, M. Fernandez and D. Suciu. *International Conference on Database Theory, Springer LNCS 1* (January 1997), pp. 336-351.
76. "A Query Language and Optimization Techniques for Unstructured Data," P. Buneman, S.B. Davidson, G. Hillebrand and D. Suciu. *Proceedings of SIGMOD'96* (May 1996), pp. 505-516.
77. "Programming Constructs for Unstructured Data," P. Buneman, S.B. Davidson and D. Suciu. *Proceedings of the Workshop on Database Programming Languages* (Sept. 1995).
78. "A Data Transformation System for Biological Data Sources," P. Buneman, S.B. Davidson, K. Hart, C. Overton and L. Wong. *Proceedings of the 21'st International Conference on Very Large Data Bases* (September 1995), pp. 158-169.
79. "Facilitating Transformations in a Human Genome Project Database," S.B. Davidson, A. Kosky and B. Eckman. *3rd International Conference on Information and Knowledge Management* (November 1994), pp. 423-432.
80. "Deadlock Prevention in the RTC Programming System for Distributed Real-Time Applications," V. Wolfe, S.B. Davidson and I. Lee. *13th International Conference on Distributed Computing Systems* (May 25-28, 1993).
81. "Theoretical Aspects of Schema Merging," P. Buneman, S.B. Davidson and A. Kosky. *Extending Database Technology* (March 23-27, 1992), pp. 152-167.
82. "A Basis for Interactive Schema Merging," P. Buneman, S.B. Davidson and A. Kosky. *Proceedings of the 25'th Hawaii International Conference on System Sciences* (Jan. 7-10, 1992), pp. 311-322.

83. "RTC: Language Support for Distributed Real-Time Concurrency," V. Wolfe, S.B. Davidson and I. Lee. *Real-Time Systems Symposium* (Dec. 2-6, 1991), pp. 43-52.
84. "Semi-Materialization: A Performance Analysis," M. Kamel and S.B. Davidson. *Proceedings of the 24<sup>th</sup> Hawaii International Conference on System Sciences*, (Jan. 8-11, 1991), pages pp. 125-135.
85. "Deadlock Prevention in Real-time Object-Oriented Systems," with S.B. Davidson, I. Lee and V. Wolfe. *Proceeding of the Objects and Transactions Workshop for the OOPSLA/ECOOP Conference*, (October 1990).
86. "Redundancy: An Approach to the Efficient Implementation of Semantic Integrity Assertions", M. Kamel and S.B. Davidson. *Proceedings of the 23<sup>rd</sup> Hawaii International Conference on System Sciences* (Jan. 2-5, 1990), pp. 393-399.
87. "Language Constructs for Timed Atomic Commitment", S.B. Davidson , V. Wolfe and I. Lee. *Proceeding of the 19th International Symposium on Fault-Tolerant Computing* (June 1989), pp. 470-477.
88. "Protocols for Timed Atomic Commitment", S.B. Davidson, I. Lee and V. Wolfe. *Proceedings of the 9th International Conference on Distributed Computing Systems* (June 1989), pp. 199-206.
89. "A Semantics for Complex Objects and Approximate Queries: Extended Abstract," P. Buneman, S.B. Davidson and A. Watters. *Proceedings of the 7th Symposium on Principles of Database Systems* (March 1988), pp. 305-314.
90. "Generalized I/O With Timing Constraints," I. Lee and S.B. Davidson. *7th International Conference on Distributed Computing Systems* (Sept. 1987), pp. 316-323.
91. "Protocols for Timed Synchronous Process Communications", I. Lee and S.B. Davidson. *Real-Time Systems Symposium* (December 1986), pp. 128-135.
92. "A Performance Comparison of Optimistic versus Conservative Conflict Graph Analysis," S.B. Davidson and M. Winkler. *Proceedings of the Conference on Current Issues in Database Systems*, Rutgers University (May 27, 1986).
93. "Is Byzantine Agreement Useful in a Distributed Database?", S.B. Davidson, H. Garcia-Molina and F. Pittelli. *Proceedings of the 3rd Symposium on Principles of Database Systems* (April 1984), pp. 61-69.
94. "Protocols for Partitioned Distributed Database Systems," S.B. Davidson and H. Garcia-Molina. *Proceedings IEEE/ACM Symposium on Reliability in Distributed Software and Database Systems* (July 1981) pp. 145-149.

#### **Refereed Workshop Papers**

95. "Approximated Provenance for Complex Applications," by Eleanor Ainy, Susan B. Davidson, Daniel Deutch and Tova Milo. *Workshop on Theory and Practice of Provenance*, June 2014.

96. "Keyword Search in Workflow Repositories with Access Control," by Susan B. Davidson, Soohyun Lee, and Julia Stoyanovich. *Alberto Mendelzon Workshop* 2011.
97. "A Fine-Grained Workflow Model with Provenance-Aware Security Views," by Zhuowei Bao, Susan B. Davidson, and Tova Milo. *Workshop on Theory and Practice of Provenance*, June 2011.
98. "ZOOM," by Sarah Cohen-Boulakia, Shirley Cohen and Susan B. Davidson. *Provenance Challenge Workshop* (Sept. 13, 2006).
99. "Towards a Model of Provenance in Scientific Workflows," by Shirley Cohen, Sarah Cohen-Boulakia and Susan B. Davidson. *3rd International Workshop on Data Integration in the Life Sciences* (July 2006), pp. 264-279.
100. "A Model for User-Oriented Data Provenance in Pipelined Scientific Workflows," by Shawn Bowers, Timothy McPhillips, Bertram Lucascher, Shirley Cohen and Susan B. Davidson. *International Provenance and Annotation Workshop IPAW'06* (Chicago, May 3-5, 2006).
101. "A User-centric Framework for Accessing Biological Sources and Tools," by Sarah Cohen-Boulakia, Susan Davidson and Christine Froidevaux. *2nd International Workshop on Data Integration in the Life Sciences* (San Diego, July 20-22, 2005), pp. 3-18.
102. "Extending XPath to Support Linguistic Queries", by Steven Bird, Yi Chen, Susan B. Davidson, Haejoong Lee and Yifeng Zheng. *Plan-X: Programming Language Technologies for XML* (Long Beach, Jan. 11, 2005).
103. "EXPedite: a system for Encoded XML Processing," Yi Chen, Susan Davidson, Yifeng Zheng. *First International Workshop on High Performance XML Processing* (May 18, 2004).
104. "On the updatability of XML views over relational databases", Vanessa P. Braganholo, Susan B. Davidson, and Carlos A. Heuser. *WebDB Workshop* (June 2003), pp. 31-36.
105. "Constraint Preserving XML Storage in Relations," Y. Chen, S.B. Davidson and Y. Zheng. *WebDB Workshop* (June 6-7, 2002).
106. "Nested Functional Dependencies," C. Hara and S.B. Davidson. *Argentina Workshop on Database Systems* (Sept. 3-4, 1998).
107. "A Logic Grammar Based Approach to Transforming Data Between Files and Biological Databases," W. Fan, S.B. Davidson and C. Overton. *Meeting on Interconnecting Molecular Biology Databases* (July 20-22, 1995).
108. "Morphing Sparsely Populated Data," A. Kosky, S.B. Davidson and C. Hara. *Meeting on Interconnecting Molecular Biology Databases* (July 20-22, 1995).
109. "TSL: A Transformation Language for Integration and Evolution of Biological Databases," A. Kosky, S.B. Davidson and B. Eckman. *Meeting on Interconnecting Molecular Biology Databases* (August 1994).

110. "Facilitating Transformations in a Human Genome Project Database," A. Kosky, S.B. Davidson and B. Eckman. *NSF Meeting on Interconnecting Molecular Biology Databases* (August 9-12, 1994).
111. "CCSR 92: Calculus for Communicating Shared Resources, vintage 1992," I. Lee, P. Bremond and S.B. Davidson. *First North American Process Algebra Workshop NAPAW* (Aug. 28, 1992).
112. "Supporting Real-Time Concurrency," I. Lee, S.B. Davidson and V. Wolfe. *7th Workshop on Real-Time Operating Systems* (May 11-12, 1990), pp. 49-54.
113. "Federated Approximations for Heterogeneous Databases", P. Buneman, S.B. Davidson and A. Watters. *Workshop on Heterogeneous Database Systems*, sponsored by NSF (Evanston, IL, Dec. 1990).
114. "Partial Computation in Real-Time Database Systems," S.B. Davidson and A. Watters, *5th Workshop on Real-Time Operating Systems* (May 12-13,1988), pp. 117-121.
115. "Motivating Time as a First Class Entity," I. Lee, S.B. Davidson and V. Wolfe. *4th Workshop on Real-Time Operating Systems* (July 3-4, 1987), pp. 165-169.

#### **Refereed Posters/Demos:**

116. "PROX: Approximated Summarization of Data Provenance," by Eleanor Ainy, Pierre Bourhis, Susan B. Davidson, Daniel Deutch, and Tova Milo. *EDBT 2016*: 620-623.
117. "Ontology Assisted Crowd Mining," by Yael Amsterdamer, Susan B. Davidson, Tova Milo, Slava Novgorodov, and Amit Somech. *PVLDB 7(13)*: 1597-1600 (2014) (Demo paper).
118. "PDiffView: Viewing the Difference in Provenance of Workflow Results," by Zhuowei Bao, Sarah Cohen-Boulakia, Susan B. Davidson and Pierrick Girard. *VLDB 2009* (Demo paper).
119. "WOLVES: Achieving Correct Provenance Analysis by Detecting and Resolving Unsound Workflow Views," by Peng Sun, Ziyang Liu, Susan Davidson and Yi Chen. *VLDB 2009* (Demo paper).
120. "Zoom\*UserViews: Querying Relevant Provenance in Workflow Systems" by Olivier Biton, Sarah Cohen-Boulakia and Susan B. Davidson. *VLDB 2007*, pp. 1366-1369 ( Demo paper).
121. "Querying Biologically Relevant Provenance information in Scientific Workflow Systems with Zoom\*UserViews," by Sarah Cohen-Boulakia, Olivier Biton, Susan Davidson. *4th International Workshop on Data Integration in the Life Sciences* (June 2007).
122. "Crimson: A Data Management System to Support Evaluating Phylogenetic Tree Reconstruction Algorithms," by Yifeng Zheng, Susan Davidson, Stephen Fischer and Junhyong Kim. *VLDB2006* (Sept. 2006), p. 1231-1234.
123. "SHARQ Guide: Finding relevant biological data and queries in a peer data management system," by Sarah Cohen-Boulakia, Olivier Biton, Shirley Cohen, Zachary Ives, Val Tannen and Susan Davidson. *3rd International Workshop on Data Integration in the Life Sciences* (July 2006), Poster #14.

## Book Chapters

124. “The Information Integration System K2” by S. Davidson, S. Harker and V. Tannen. In *Bioinformatics: Managing Scientific Data* co-edited by Terence Critchlow and Zoe Lacroix, Morgan Kaufmann Publishers (2003).
125. “The Kleisli Approach to Data Integration and Transformation” by L. Wong and S. Davidson. In *The Functional Approach to Data Management: Modeling, Analyzing, and Integrating Heterogeneous Data*, edited by Peter Gray, Larry Kerschberg, Peter King, Alex Poulouvassilis, chapter 6, pages 135–165. Springer-Verlag, September 2003.
126. “Creating and Maintaining Curated View Databases”, by S. B. DAvidson, H. Liefke and L. Wong. In *Knowledge Discovery and Data Mining in Biological Databases*, edited by V. Brusica, J. Zeleznikow and L. Wong, World Scientific Publishing/Singapore University Press, 1999.
127. “BioKleisli: Integrating Biomedical Data and Analysis Packages”, by Susan Davidson, Peter Buneman, Jonathan Crabtree, Val Tannen, Chris Overton, Limsoon Wong. In *Bioinformatics*, edited by Stan Letovsky, Kluwer Academic Publishers, Boston, 1999.
128. Foundations of Real-Time Computing: Formal Specifications and Methods, edited by Andre M. van Tilborg and Gary M. Koob.
129. “Replicated Data and Partition Failures”, *An Advanced Course on Distributed Systems*, edited by Sape Mullender, Addison-Wesley. This book was used in conjunction with the Fingerlakes ’89 course taught at Cornell in July 1989.

## Invited Papers

130. "Engineering the Future," by Susan B. Davidson, Yasmin Kafai and Michele Grab. International Innovation North America, May 2013, pp. 26-28.
131. "To Show or Not to Show in Workflow Provenance," by Susan B. Davidson, Sanjeev Khanna, and Tova Milo. In Search of Elegance in the Theory and Practice of Computation 2013: 217-226.
132. "Hiding Data and Structure in Workflow Provenance," by Susan B. Davidson, Zhuowei Bao, and Sudeepa Roy. DNIS 2011: 41-48
133. "On provenance and privacy," by Susan B. Davidson, Sanjeev Khanna, Sudeepa Roy, Julia Stoyanovich, Val Tannen, and Yi Chen. ICDT 2011: 3-10.
134. "On User Views in Scientific Workflow Systems", by Susan B. Davidson, Yi Chen, Peng Sun, and Sarah Cohen Boulakia. SWPM 2009.
135. "Active XML and Data Activation (Extended Abstract)," S. Abiteboul, S.B. Davidson and T. Milo. 12th International Workshop on Abstract State (ASM) 2005.
136. "Penn's Programs: A Comprehensive Solution," S.B. Davidson. Science's Next Wave, Sept. 2000.
137. "Transforming and Integrating Biomedical Data using Kleisli: A Perspective," S.B. Davidson, P. Buneman, S. Harker, C. Overton and V. Tannen. SIGBio. July 1999.
138. "Specifying Database Transformations in WOL," S.B. Davidson and A. Kosky. IEEE Bulletin, Special Issue on Database Transformation Technology. Summer 1999.
139. "Challenges in Integrating Biological Data Sources," S.B. Davidson, C. Overton and P. Buneman. (Panel Position Paper) Proceedings MIMBD (July 1995), Cambridge, England.
140. "Maintaining Consistency Over a Network in Real-Time Control Applications," (Regular paper with Insup Lee and Victor Wolfe) *American Control Conference*, (June 1989).
141. "The Optimistic Protocol for Partitioned Distributed Database Systems," (Panel Position Paper) Proceedings COMPSAC82 (November 1982), p. 453.
142. "Using Semantics to Reduce Blocking Delays," (Panel Position Paper with Hector Garcia-Molina) *Proceedings IEEE/ACM Second Symposium on Reliability in Distributed Software and Database Systems* (July 1982), p. 62.
143. "Partitions in a Distributed Database," with Hector Garcia-Molina. *Jornados en Computacion: Perspectivas* (June 1982).