Ashraf Aboulnaga

Department of Computer Science and Engineering
University of Texas at Arlington
Engineering Research Building (Room 648)
Arlington, Texas, USA

ashraf.aboulnaga@uta.edu
https://ashraf.aboulnaga.me
Phone: +1 (817) 272-3716

Education

- **Ph.D. in Computer Science**, University of Wisconsin - Madison, 2002.
  Thesis: *Cost Estimation Techniques for Database Systems*.
  Advisor: Professor Jeffrey Naughton.

- **M.S. in Computer Science**, University of Wisconsin - Madison, 1999.

- **M.S. in Computer Science**, Faculty of Engineering, Alexandria University, Egypt, 1996.
  Thesis: *An Algorithm and a Cost Model for Window Queries in Spatial Databases*.
  Advisor: Professor Walid Aref.

- **B.S. in Computer Science**, Faculty of Engineering, Alexandria University, Egypt, 1993.

Professional Experience

- **University of Texas at Arlington**, Arlington, Texas, USA.
  Department of Computer Science and Engineering.
  Professor, 2023–present.

- **Qatar Computing Research Institute**, Doha, Qatar.
  Chief Scientist, 2021–2023.
  Senior Research Director, 2019–2021.
  Research Director of the Distributed Systems Group, 2015–2018.
  Principal Scientist and Manager of the Distributed Systems Group, 2013–2015.

- **University of Waterloo**, Waterloo, Ontario, Canada.
  Cheriton School of Computer Science.
  Adjunct Associate Professor, 2015–2018.
  Associate Professor, 2009–2015.
  Assistant Professor, 2004–2009.

- **Google**, Waterloo, Ontario, Canada.
  Visiting Research Scientist, December 2009–June 2010.

- **IBM Almaden Research Center**, San Jose, California, USA.
  Research Staff Member, 2002–2004.

- **University of Wisconsin**, Madison, Wisconsin, USA.
  Research Assistant, 1997–2002.
  Teaching Assistant, 1996.

- **Microsoft Research**, Redmond, Washington, USA.
  Researcher Intern, May–August 1999 and May–August 1998.

- **Alexandria University**, Alexandria, Egypt.
  Lecturer, 1993–1996.

Awards

- IEEE Computer Society Distinguished Contributor, 2021.

- ACM Distinguished Scientist, 2015.
- Best Paper Award at the ACM Symposium on Cloud Computing (SoCC), 2015.
- Best Paper Award at the International Conference on Very Large Data Bases (VLDB), 2011.
- Ontario Early Researcher Award, 2009.
- Google Research Award, 2008.
- IBM Centre for Advanced Studies (CAS) Faculty Fellow, 2006–2015.
- IBM CAS Project of the Year Award (2008) and Innovation Impact of the Year Award (2009) for the "XML Design Advisor" project.
- Microsoft Research Graduate Fellowship, 1998–2000.
- B.S. with First Degree of Honor (Summa Cum Laude), 1993.

Professional Memberships

- ACM Distinguished Scientist.
- Senior Member of the IEEE and IEEE Computer Society.

Publications

Refereed Journal Papers

[J1] Aisha Mohamed, Ghadeer Abuoda, Abdurrahman Ghanem, Zoi Kaoudi, and Ashraf Aboulnaga. "RDF-Frames: Knowledge Graph Access for Machine Learning Tools," *The VLDB Journal (VLDBJ)*, volume 31, number 2, pp. 321–346, March 2022. *(special issue on big graph data management and processing)*

[J2] Yifei Yang, Matt Youill, Matthew E. Woicik, Yizhou Liu, Xiangyao Yu, Marco Serafini, Ashraf Aboulnaga, and Michael Stonebraker. "FlexPushdownDB: Hybrid Pushdown and Caching in a Cloud DBMS," *Proceedings of the VLDB Endowment (PVLDB)*, volume 14, number 11, pp. 2101–2113, July 2021. Presented at the *International Conference on Very Large Data Bases (VLDB '21)*, Copenhagen, Denmark, August 2021.

[J3] Xiaowei Zhu, Guanyu Feng, Marco Serafini, Xiaosong Ma, Jiping Yu, Lei Xie, Ashraf Aboulnaga, and Wenguang Chen. "LiveGraph: A Transactional Graph Storage System with Purely Sequential Adjacency List Scans," *Proceedings of the VLDB Endowment (PVLDB)*, volume 13, number 7, pp. 1020–1034, March 2020. Presented at the *International Conference on Very Large Data Bases (VLDB '20)*, Virtual Conference, August 2020.

[J4] Junjay Tan, Thanaa Ghanem, Matthew Perron, Xiangyao Yu, Michael Stonebraker, David J. DeWitt, Marco Serafini, Ashraf Aboulnaga, and Tim Kraska. "Choosing a Cloud DBMS: Architectures and Tradeoffs," *Proceedings of the VLDB Endowment (PVLDB)*, volume 12, number 12, pp. 2170–2182, August 2019. Presented at the *International Conference on Very Large Data Bases (VLDB '19)*, Los Angeles, California, USA, August 2019.

[J5] Ibrahim Abdelaziz, Essam Mansour, Mourad Ouzzani, Ashraf Aboulnaga, and Panos Kalnis. "Lusail: A System for Querying Linked Data at Scale," *Proceedings of the VLDB Endowment (PVLDB)*, volume 11, number 4, pp. 485–498, December 2017. Presented at the *International Conference on Very Large Data Bases (VLDB '18)*, Rio de Janeiro, Brazil, August 2018. *(extended version of [C27])*

[J6] Michael J. Mior, Kenneth Salem, Ashraf Aboulnaga, and Rui Liu. "NoSE: Schema Design for NoSQL Applications," *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, volume 29, number 10, pp. 2275–2289, October 2017. *(extended version of [C28], special issue for the best papers of ICDE 2016)*

[J7] Marco Serafini, Rebecca Taft, Aaron J. Elmore, Andrew Pavlo, Ashraf Aboulnaga, and Michael Stonebraker. "Clay: Fine-Grained Adaptive Partitioning for General Database Schemas," *Proceedings of the VLDB Endowment (PVLDB)*, volume 10, number 5, pp. 445–456, December 2016. Presented at the *International Conference on Very Large Data Bases (VLDB '17)*, Munich, Germany, August 2017.

[J8] Rebecca Taft, Essam Mansour, Marco Serafini, Jennie Duggan, Aaron J. Elmore, Ashraf Aboulnaga, Andrew Pavlo, and Michael Stonebraker. "E-Store: Fine-Grained Elastic Partitioning for Distributed Transaction Processing," *Proceedings of the VLDB Endowment (PVLDB)*, volume 8, number 3, pp. 245–256, November 2014. Presented at the *International Conference on Very Large Data Bases (VLDB '15)*, Kohala Coast, Hawaii, USA, August 2015.

[J9] Marco Serafini, Essam Mansour, Ashraf Aboulnaga, Kenneth Salem, Taha Rafiq, and Umar Farooq Minhas. "Accordion: Elastic Scalability for Database Systems Supporting Distributed Transactions," *Proceedings of the VLDB Endowment (PVLDB)*, volume 7, number 12, pp. 1035–1046, August 2014. Presented at the *International Conference on Very Large Data Bases (VLDB '14)*, Hangzhou, China, September 2014.

[J10] Essam Mansour, Ahmed El-Roby, Panos Kalnis, Aron Ahmadia, and Ashraf Aboulnaga. "RACE: A Scalable and Elastic Parallel System for Discovering Repeats in Very Long Sequences," *Proceedings of the VLDB Endowment (PVLDB)*, volume 6, number 10, pp. 865–876, August 2013. Presented at the *International Conference on Very Large Data Bases (VLDB '13)*, Riva del Garda, Italy, August 2013.

[J11] Iman Elghandour, Ashraf Aboulnaga, Daniel C. Zilio, and Calisto Zuzarte. "Recommending XML Physical Designs for XML Databases," *The VLDB Journal (VLDBJ)*, volume 22, number 4, pp. 447–470, August 2013. *(extended version of [C44] and [C51])*

[J12] Rui Liu, Ashraf Aboulnaga, and Kenneth Salem. "DAX: A Widely Distributed Multi-tenant Storage Service for DBMS Hosting," *Proceedings of the VLDB Endowment (PVLDB)*, volume 6, number 4, pp. 253–264, February 2013. Presented at the *International Conference on Very Large Data Bases (VLDB '13)*, Riva del Garda, Italy, August 2013.

[J13] Umar Farooq Minhas, Shriram Rajagopalan, Brendan Cully, Ashraf Aboulnaga, Kenneth Salem, and Andrew Warfield. "RemusDB: Transparent High Availability for Database Systems," *The VLDB Journal (VLDBJ)*, volume 22, number 1, pp. 29–45, February 2013. *(extended version of [J15], special issue for the best papers of VLDB 2011)*

[J14] Iman Elghandour and Ashraf Aboulnaga. "ReStore: Reusing Results of MapReduce Jobs," *Proceedings of the VLDB Endowment (PVLDB)*, volume 5, number 6, pp. 586–597, February 2012. Presented at the *International Conference on Very Large Data Bases (VLDB '12)*, Istanbul, Turkey, August 2012.

[J15] Umar Farooq Minhas, Shriram Rajagopalan, Brendan Cully, Ashraf Aboulnaga, Kenneth Salem, and Andrew Warfield. "RemusDB: Transparent High Availability for Database Systems," *Proceedings of the VLDB Endowment (PVLDB)*, volume 4, number 11, pp. 738–748, August 2011. Presented at the *International Conference on Very Large Data Bases (VLDB '11)*, Seattle, Washington, USA, August 2011.
(Best Paper Award)

[J16] Mumtaz Ahmad, Ashraf Aboulnaga, Shivnath Babu, and Kamesh Munagala. "Interaction-Aware Scheduling of Report Generation Workloads," *The VLDB Journal (VLDBJ)*, volume 20, number 4, pp. 589–615, August 2011. *(extended version of [C48] and [C52])*

[J17] Ahmed A. Soror, Umar Farooq Minhas, Ashraf Aboulnaga, Kenneth Salem, Peter Kokosielis, and Sunil Kamath. "Automatic Virtual Machine Configuration for Database Workloads," *ACM Transactions on Database Systems (TODS)*, volume 35, number 1, article 7 (47 pages), February 2010. *(extended version of [C50], special issue for the best papers of SIGMOD 2008)*

[J18] Ashraf Aboulnaga and Walid G. Aref. "Window Query Processing in Linear Quadtrees," *Distributed and Parallel Databases*, volume 10, number 2, pp. 111–126, September 2001.

Refereed Conference and Workshop Papers

[C19] Chaoyi Ruan, Yingqiang Zhang, Chao Bi, Xiaosong Ma, Hao Chen, Feifei Li, Xinjun Yang, Cheng Li, Ashraf Aboulnaga, and Yinlong Xu. "Persistent Memory Disaggregation for Cloud-Native Relational Databases," In *Proceedings of the Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS '23), Volume 3*, Vancouver, Canada, pp. 498–512, March 2023.

[C20] Ghadeer Abuoda, Saravanan Thirumuruganathan, and Ashraf Aboulnaga. "Accelerating Entity Lookups in Knowledge Graphs Through Embeddings," In *Proceedings of the IEEE International Conference on Data Engineering (ICDE '22)*, Virtual Conference, pp. 1111–1123, May 2022.

[C21] Aisha Mohamed, Shameem A. Parambath, Zoi Kaoudi, and Ashraf Aboulnaga. "Popularity Agnostic Evaluation of Knowledge Graph Embeddings," In *Proceedings of the Conference on Uncertainty in Artificial Intelligence (UAI '20)*, Virtual Conference, pp. 1059–1068, August 2020.

[C22] Xiangyao Yu, Matt Youill, Matthew E. Woicik, Abdurrahman Ghanem, Marco Serafini, Ashraf Aboulnaga, and Michael Stonebraker. "PushdownDB: Accelerating a DBMS Using S3 Computation," In *Proceedings of the IEEE International Conference on Data Engineering (ICDE '20)*, Virtual Conference, pp. 1802–1805, April 2020. *(poster)*

[C23] Xiongchao Tang, Haojie Wang, Xiaosong Ma, Nosayba El-Sayed, Jidong Zhai, Wenguang Chen, and Ashraf Aboulnaga. "Spread-n-share: Improving Application Performance and Cluster Throughput with Resource-aware Job Placement," In *Proceedings of the International Conference for High Performance Computing, Networking, Storage, and Analysis (SC '19)*, Denver, Colorado, USA, pp. 12:1–15, November 2019.

[C24] Ghadeer Abuoda, Gianmarco De Francisci Morales, and Ashraf Aboulnaga. "Link Prediction via Higher-Order Motif Features," In *Proceedings of the European Conference on Machine Learning and Knowledge Discovery in Databases (ECML/PKDD '19)*, Würzburg, Germany, pp. 412–429, September 2019.

[C25] Rebecca Taft, Nosayba El-Sayed, Marco Serafini, Yu Lu, Ashraf Aboulnaga, Michael Stonebraker, Ricardo Mayerhofer, and Francisco Andrade. "P-Store: An Elastic Database System with Predictive Provisioning," In *Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD '18)*, Houston, Texas, USA, pp. 205–219, June 2018.

[C26] Ahmed El-Roby and Ashraf Aboulnaga. "UFeed: Refining Web Data Integration Based on User Feedback," In *Proceedings of the ACM International Conference on Information and Knowledge Management (CIKM '17)*, Singapore, pp. 187–196, November 2017.

[C27] Ibrahim Abdelaziz, Essam Mansour, Mourad Ouzzani, Ashraf Aboulnaga, and Panos Kalnis. "Query Optimizations Over Decentralized RDF Graphs," In *Proceedings of the IEEE International Conference on Data Engineering (ICDE '17)*, San Diego, California, USA, pp. 139–142, April 2017. *(poster)*

[C28] Michael J. Mior, Kenneth Salem, Ashraf Aboulnaga, and Rui Liu. "NoSE: Schema Design for NoSQL Applications," In *Proceedings of the IEEE International Conference on Data Engineering (ICDE '16)*, Helsinki, Finland, pp. 181–192, May 2016.

[C29] Carlos H.C. Teixeira, Alexandre J. Fonseca, Marco Serafini, Georgios Siganos, Mohammed J. Zaki, and Ashraf Aboulnaga. "Arabesque: A System for Distributed Graph Pattern Mining," In *Proceedings of the ACM Symposium on Operating Systems Principles (SOSP '15)*, Monterey, California, USA, pp. 425–440, October 2015.

[C30] Jaemyung Kim, Kenneth Salem, Khuzaima Daudjee, Ashraf Aboulnaga, and Xin Pan. "Database High Availability Using SHADOW Systems," In *Proceedings of the ACM Symposium on Cloud Computing (SoCC '15)*, Kohala Coast, Hawaii, USA, pp. 209–221, August 2015. *(Best Paper Award)*

[C31] Tarek Elgamal, Maysam Yabandeh, Ashraf Aboulnaga, Waleed Mustafa, and Mohamed Hefeeda. "sPCA: Scalable Principal Component Analysis for Big Data on Distributed Platforms," In *Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD '15)*, Melbourne, Australia, pp. 79–91, June 2015.

[C32] Ahmed El-Roby and Ashraf Aboulnaga. "ALEX: Automatic Link Exploration in Linked Data," In *Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD '15)*, Melbourne, Australia, pp. 1839–1853, June 2015.

[C33] Jingen Xiang, Huangdong Meng, and Ashraf Aboulnaga. "Scalable Matrix Inversion Using MapReduce," In *Proceedings of the International Symposium on High-Performance Parallel and Distributed Computing (HPDC '14)*, Vancouver, Canada, pp. 177–190, June 2014.

[C34] Mostafa Ead, Herodotos Herodotou, Ashraf Aboulnaga, and Shivnath Babu. "PStorM: Profile Storage and Matching for Feedback-Based Tuning of MapReduce Jobs," In *Proceedings of the International Conference on Extending Database Technology (EDBT '14)*, Athens, Greece, pp. 1–12, March 2014.

[C35] Farhana H. Zulkernine, Patrick Martin, Ying Zou, Michael Bauer, Femida Gwadry-Sridhar, and Ashraf Aboulnaga. "Towards Cloud-Based Analytics-as-a-Service (CLAaaS) for Big Data Analytics in the Cloud,"

In *Proceedings of the IEEE International Congress on Big Data*, Santa Clara, California, USA, pp. 62–69, June 2013.

[C36] Jingen Xiang, Cong Guo, and Ashraf Aboulnaga. "Scalable Maximum Clique Computation Using MapReduce," In *Proceedings of the IEEE International Conference on Data Engineering (ICDE '13)*, Brisbane, Australia, pp. 74–85, April 2013.

[C37] Umar Farooq Minhas, Rui Liu, Ashraf Aboulnaga, Kenneth Salem, Jonathan Ng, and Sean Robertson. "Elastic Scale-out for Partition-Based Database Systems," In *Proceedings of the International Workshop on Self-Managing Database Systems (SMDB '12)*, Washington, DC, USA, April 2012.

[C38] Muhammad Bilal Sheikh, Umar Farooq Minhas, Omar Zia Khan, Ashraf Aboulnaga, Pascal Poupart, and David J. Taylor. "A Bayesian Approach to Online Performance Modeling for Database Appliances using Gaussian Models," In *Proceedings of the International Conference on Autonomic Computing (ICAC '11)*, Karlsruhe, Germany, pp. 121–130, June 2011.

[C39] Mumtaz Ahmad, Songyun Duan, Ashraf Aboulnaga, and Shivnath Babu. "Predicting Completion Times of Batch Query Workloads Using Interaction-aware Models and Simulation," In *Proceedings of the International Conference on Extending Database Technology (EDBT '11)*, Uppsala, Sweden, pp. 449–460, March 2011. *(extended version of [C41])*

[C40] Hatem A. Mahmoud and Ashraf Aboulnaga. "Schema Clustering and Retrieval for Multi-domain Pay-As-You-Go Data Integration Systems," In *Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD '10)*, Indianapolis, Indiana, USA, pp. 411–422, June 2010.

[C41] Mumtaz Ahmad, Songyun Duan, Ashraf Aboulnaga, and Shivnath Babu. "Interaction-Aware Prediction of Business Intelligence Workload Completion Times," In *Proceedings of the IEEE International Conference on Data Engineering (ICDE '10)*, Long Beach, California, USA, pp. 413–416, March 2010. *(short paper)*

[C42] Sean Tozer, Tim Brecht, and Ashraf Aboulnaga. "Q-Cop: Avoiding Bad Query Mixes to Minimize Client Timeouts Under Heavy Loads," In *Proceedings of the IEEE International Conference on Data Engineering (ICDE '10)*, Long Beach, California, USA, pp. 397–408, March 2010.

[C43] Ashraf Aboulnaga, Ziyu Wang, and Zi Ye Zhang. "Packing the Most Onto Your Cloud," In *Proceedings of the International Workshop on Cloud Data Management (CloudDB '09)*, Hong Kong, China, pp. 25–28, November 2009. *(short paper)*

[C44] Iman Elghandour, Ashraf Aboulnaga, Daniel C. Zilio, and Calisto Zuzarte. "Recommending XMLTable Views for XQuery Workloads," In *Proceedings of the International XML Database Symposium (XSym '09)*, Lyon, France, pp. 129–144, August 2009.

[C45] Chen Zhang, Hans De Sterck, Ashraf Aboulnaga, Haig Djambazian, and Robert Sladek. "Case Study of Scientific Data Processing on a Cloud Using Hadoop," In *Proceedings of the International Symposium on High Performance Computing Systems and Applications (HPCS '09)*, Kingston, Ontario, Canada, pp. 400–415, June 2009.

[C46] Mumtaz Ahmad, Ashraf Aboulnaga, and Shivnath Babu. "Query Interactions in Database Workloads," In *Proceedings of the International Workshop on Testing Database Systems (DBTest '09)*, Providence, Rhode Island, USA, article 11 (6 pages), June 2009.

[C47] Xin Liu, Ashraf Aboulnaga, Kenneth Salem, and Xuhui Li. "CLIC: CLient-Informed Caching for Storage Servers," In *Proceedings of the USENIX Conference on File and Storage Technologies (FAST '09)*, San Francisco, California, USA, pp. 297–310, February 2009.

[C48] Mumtaz Ahmad, Ashraf Aboulnaga, Shivnath Babu, and Kamesh Munagala. "Modeling and Exploiting Query Interactions in Database Systems," In *Proceedings of the ACM International Conference on Information and Knowledge Management (CIKM '08)*, Napa, California, USA, pp. 183–192, October 2008.

[C49] Ahmed Ataullah, Ashraf Aboulnaga, and Frank Wm. Tompa. "Records Retention in Relational Database Systems," In *Proceedings of the ACM International Conference on Information and Knowledge Management (CIKM '08)*, Napa, California, USA, pp. 873-882, October 2008.

[C50] Ahmed A. Soror, Umar Farooq Minhas, Ashraf Aboulnaga, Kenneth Salem, Peter Kokosielis, and Sunil Kamath. "Automatic Virtual Machine Configuration for Database Workloads," In *Proceedings of the ACM*

SIGMOD International Conference on Management of Data (SIGMOD '08), Vancouver, Canada, pp. 953–966, June 2008.

[C51] Iman Elghandour, Ashraf Aboulnaga, Daniel C. Zilio, Fei Chiang, Andrey Balmin, Kevin Beyer, and Calisto Zuzarte. "XML Index Recommendation with Tight Optimizer Coupling," In *Proceedings of the IEEE International Conference on Data Engineering (ICDE '08)*, Cancun, Mexico, pp. 833–842, April 2008.

[C52] Mumtaz Ahmad, Ashraf Aboulnaga, Shivnath Babu, and Kamesh Munagala. "QShuffler: Getting the Query Mix Right," In *Proceedings of the IEEE International Conference on Data Engineering (ICDE '08)*, Cancun, Mexico, pp. 1415–1417, April 2008. *(poster)*

[C53] Umar Farooq Minhas, Jitendra Yadav, Ashraf Aboulnaga, and Kenneth Salem. "Database Systems on Virtual Machines: How Much do You Lose?," In *Proceedings of the International Workshop on Self-Managing Database Systems (SMDB '08)*, Cancun, Mexico, pp. 35–41, April 2008.

[C54] Kareem El Gebaly and Ashraf Aboulnaga. "Robustness in Automatic Physical Database Design," In *Proceedings of the International Conference on Extending Database Technology (EDBT '08)*, Nantes, France, pp. 145–156, March 2008.

[C55] Ashraf Aboulnaga and Kareem El Gebaly. "μBE: User Guided Source Selection and Schema Mediation for Internet Scale Data Integration," In *Proceedings of the IEEE International Conference on Data Engineering (ICDE '07)*, Istanbul, Turkey, pp. 186–195, April 2007.

[C56] Ahmed A. Soror, Ashraf Aboulnaga, and Kenneth Salem. "Database Virtualization: A New Frontier for Database Tuning and Physical Design," In *Proceedings of the International Workshop on Self-Managing Database Systems (SMDB '07)*, Istanbul, Turkey, pp. 388–394, April 2007.

[C57] Ning Zhang, M. Tamer Özsu, Ihab F. Ilyas, and Ashraf Aboulnaga. "FIX: Feature-based Indexing Techniques for XML Documents," In *Proceedings of the International Conference on Very Large Data Bases (VLDB '06)*, Seoul, Korea, pp. 259–270, September 2006.

[C58] Ning Zhang, M. Tamer Özsu, Ashraf Aboulnaga, and Ihab F. Ilyas. "XSeed: Accurate and Fast Cardinality Estimation for XPath Queries," In *Proceedings of the IEEE International Conference on Data Engineering (ICDE '06)*, Atlanta, Georgia, USA, pp. 61–72, April 2006.

[C59] Xuhui Li, Ashraf Aboulnaga, Kenneth Salem, Aamer Sachedina, and Shaobo Gao. "Second-Tier Cache Management Using Write Hints," In *Proceedings of the USENIX Conference on File and Storage Technologies (FAST '05)*, San Francisco, California, USA, pp. 115–128, December 2005.

[C60] Ashraf Aboulnaga, Peter J. Haas, Mokhtar Kandil, Sam Lightstone, Guy M. Lohman, Volker Markl, Ivan Popivanov, and Vijayshankar Raman. "Automated Statistics Collection in DB2 UDB," In *Proceedings of the International Conference on Very Large Data Bases (VLDB '04)*, Toronto, Canada, pp. 1146–1157, August 2004.

[C61] Ihab F. Ilyas, Peter J. Haas, Volker Markl, Paul Brown, and Ashraf Aboulnaga. "CORDS: Automatic Discovery of Correlations and Soft Functional Dependencies," In *Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD '04)*, Paris, France, pp. 647–658, June 2004.

[C62] Ihab F. Ilyas, Peter J. Haas, Volker Markl, Paul Brown, and Ashraf Aboulnaga. "Automatic Relationship Discovery in Self-Managing Database Systems," In *Proceedings of the International Conference on Autonomic Computing (ICAC '04)*, New York, New York, USA, pp. 340–341, May 2004. *(poster)*

[C63] Ashraf Aboulnaga and Jeffrey F. Naughton. "Building XML Statistics for the Hidden Web," In *Proceedings of the ACM International Conference on Information and Knowledge Management (CIKM '03)*, New Orleans, Louisiana, USA, pp. 358–365, November 2003.

[C64] Ashraf Aboulnaga, Alaa R. Alameldeen, and Jeffrey F. Naughton. "Estimating the Selectivity of XML Path Expressions for Internet Scale Applications," In *Proceedings of the International Conference on Very Large Data Bases (VLDB '01)*, Rome, Italy, pp. 591–600, September 2001.

[C65] Ashraf Aboulnaga, Jeffrey F. Naughton, and Chun Zhang. "Generating Synthetic Complex-structured XML Data," In *Proceedings of the International Workshop on the Web and Databases (WebDB '01)*, Santa Barbara, California, USA, pp. 79–84, May 2001.

[C66] Ashraf Aboulnaga and Jeffrey F. Naughton. "Accurate Estimation of the Cost of Spatial Selections," In *Proceedings of the IEEE International Conference on Data Engineering (ICDE '00)*, San Diego, California, USA, pp. 123–134, February 2000.

[C67] Ashraf Aboulnaga and Surajit Chaudhuri. "Self-tuning Histograms: Building Histograms Without Looking at Data," In *Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD '99)*, Philadelphia, Pennsylvania, USA, pp. 181–192, June 1999.

System Demonstrations

[D68] Aisha Mohamed, Ghadeer Abuoda, Abdurrahman Ghanem, Zoi Kaoudi, and Ashraf Aboulnaga. "RDF-Frames: Knowledge Graph Access for Machine Learning Tools," Demonstration at the *International Conference on Very Large Data Bases (VLDB '20)*, Virtual Conference, *Proceedings of the VLDB Endowment (PVLDB)*, volume 13, number 12, pp. 2889–2892, August 2020. *(demonstration of the work in [J1])*

[D69] Essam Mansour, Ibrahim Abdelaziz, Mourad Ouzzani, Ashraf Aboulnaga, and Panos Kalnis. "A Demonstration of Lusail: Querying Linked Data at Scale," Demonstration at the *ACM SIGMOD International Conference on Management of Data (SIGMOD '17)*, Chicago, Illinois, USA, pp. 1603–1606, May 2017. *(demonstration of the work in [J5, C27])*

[D70] Eslam Hussein, Abdurrahman Ghanem, Vinícius Vitor dos Santos Dias, Carlos H. C. Teixeira, Ghadeer Abuoda, Marco Serafini, Georgos Siganos, Gianmarco De Francisci Morales, Ashraf Aboulnaga, and Mohammed J. Zaki. "Graph Data Mining with Arabesque," Demonstration at the *ACM SIGMOD International Conference on Management of Data (SIGMOD '17)*, Chicago, Illinois, USA, pp. 1647–1650, May 2017. *(demonstration of the work in [C29])*
(Best Demonstration Award, Honorable Mention)

[D71] Ahmed El-Roby, Khaled Ammar, Ashraf Aboulnaga, and Jimmy Lin. "Sapphire: Querying RDF Data Made Simple," Demonstration at the *International Conference on Very Large Data Bases (VLDB '16)*, New Delhi, India, *Proceedings of the VLDB Endowment (PVLDB)*, volume 9, number 13, pp. 1481–1484, September 2016.

[D72] Ahmed El-Roby and Ashraf Aboulnaga. "ALEX: Automatic Link Exploration in Linked Data," Demonstration at the *IEEE International Conference on Data Engineering (ICDE '16)*, Helsinki, Finland, pp. 1322–1325, May 2016. *(demonstration of the work in [C32])*

[D73] Essam Mansour, Andrei Vlad Sambra, Sandro Hawke, Maged Zereba, Sarven Capadisli, Abdurrahman Ghanem, Ashraf Aboulnaga, and Tim Berners-Lee. "A Demonstration of the Solid Platform for Social Web Applications," Demonstration at the *International Conference on World Wide Web (WWW '16)*, Montreal, Canada, pp. 223–226, April 2016.

[D74] Andrei Sambra, Sandro Hawke, Tim Berners-Lee, Lalana Kagal, and Ashraf Aboulnaga. "CIMBA - Client-Integrated MicroBlogging Architecture," Demonstration at the *International Semantic Web Conference (ISWC '14)*, Riva del Garda, Italy, pp. 57–60, October 2014.

[D75] Iman Elghandour and Ashraf Aboulnaga. "ReStore: Reusing Results of MapReduce Jobs in Pig," Demonstration at the *ACM SIGMOD International Conference on Management of Data (SIGMOD '12)*, Scottsdale, Arizona, USA, pp. 701–704, May 2012. *(demonstration of the work in [J14])*

[D76] Iman Elghandour, Ashraf Aboulnaga, Daniel C. Zilio, Fei Chiang, Andrey Balmin, Kevin Beyer, and Calisto Zuzarte. "An XML Index Advisor for DB2," Demonstration at the *ACM SIGMOD International Conference on Management of Data (SIGMOD '08)*, Vancouver, Canada, pp. 1267–1270, June 2008. *(demonstration of the work in [C51])*

[D77] Ashraf Aboulnaga, Kareem El Gebaly, and Daniel Wong. "μBE: User Guided Source Selection and Schema Mediation for Internet Scale Data Integration," Demonstration at the *IEEE International Conference on Data Engineering (ICDE '07)*, Istanbul, Turkey, pp. 1513–1514, April 2007. *(demonstration of the work in [C55])*

[D78] Ihab F. Ilyas, Volker Markl, Peter J. Haas, Paul G. Brown, and Ashraf Aboulnaga. "CORDS: Automatic Generation of Correlation Statistics in DB2," Demonstration at the *International Conference on Very Large Data Bases (VLDB '04)*, Toronto, Canada, pp. 1341–1344, August 2004. *(demonstration of the work in [C61])*

Tutorials

[T79] Ashraf Aboulnaga and Shivnath Babu. "Workload Management for Big Data Analytics," Tutorial at the *ACM SIGMOD International Conference on Management of Data (SIGMOD '13)*, New York, New York, USA, pp. 929–932, June 2013.

[T80] Ashraf Aboulnaga and Shivnath Babu. "Workload Management for Big Data Analytics," Advanced technology seminar at the *IEEE International Conference on Data Engineering (ICDE '13)*, Brisbane, Australia, p. 1249, April 2013.

[T81] Ashraf Aboulnaga, Cristiana Amza, and Kenneth Salem. "Virtualization and Databases: State of the Art and Research Challenges," Tutorial at the *International Conference on Extending Database Technology (EDBT '08)*, Nantes, France, pp. 746–747, March 2008.

[T82] Ashraf Aboulnaga, Cristiana Amza, and Kenneth Salem. "Virtualization and Databases: State of the Art and Research Challenges," Advanced technology seminar at the *IEEE International Conference on Data Engineering (ICDE '07)*, Istanbul, Turkey, April 2007.

Other Papers

[O83] Ashraf Aboulnaga, Azza Abouzied, Karima Echihabi, and Mourad Ouzzani. "Database Systems Research in the Arab World: A Tradition that Spans Decades," *Communications of the ACM*, volume 64, number 4, pp. 120–123, April 2021. *(invited, special section on computing in the Arab World)*

[O84] Ashraf Aboulnaga, Mohammed Al-Mannai, Hassan Al-Sayed, Sanjay Chawla, Ahmed Elmagarmid, "An AI-enabled Future for Qatar and the Region," *Communications of the ACM*, volume 64, number 4, pp. 59–61, April 2021. *(invited, special section on computing in the Arab World)*

[O85] Ashraf Aboulnaga. "Guest Editorial for Special Issue on the Best Papers of DaMoN 2019," *The VLDB Journal (VLDBJ)*, volume 29, number 6, p. 1221, November 2020.

[O86] Ashraf Aboulnaga, Beng Chin Ooi, and Patrick Valduriez. "Guest Editorial for Special Section on Data-Intensive Cloud Infrastructure," *The VLDB Journal (VLDBJ)*, volume 23, number 6, p. 843, December 2014.

[O87] Ashraf Aboulnaga and Carlo Curino. "Report on the Second International Workshop on Data Management in the Cloud (DMC 2013)," *IEEE Data Engineering Bulletin*, volume 36, number 2, pp. 2–4, June 2013. *(invited)*

[O88] Ashraf Aboulnaga and Chee Yong Chan. "Letter from the Guest Editors," *Proceedings of the VLDB Endowment (PVLDB)*, volume 6, number 4, p. x, February 2013.

[O89] Ashraf Aboulnaga and Kenneth Salem. "Report on the Fourth International Workshop on Self-Managing Database Systems (SMDB 2009)," *IEEE Data Engineering Bulletin*, volume 32, number 4, pp. 2–5, December 2009. *(invited)*

[O90] Ashraf Aboulnaga, Kenneth Salem, Ahmed A. Soror, Umar Farooq Minhas, Peter Kokosielis, and Sunil Kamath. "Deploying Database Appliances in the Cloud," *IEEE Data Engineering Bulletin*, volume 32, number 1, pp. 13–20, March 2009. *(invited)*

[O91] Jeffrey F. Naughton, David J. DeWitt, David Maier, Ashraf Aboulnaga, Jianjun Chen, Leonidas Galanis, Jaewoo Kang, Rajasekar Krishnamurthy, Qiong Luo, Naveen Prakash, Ravishankar Ramamurthy, Jayavel Shanmugasundaram, Feng Tian, Kristin Tufte, Stratis Viglas, Yuan Wang, Chun Zhang, Bruce Jackson, Anuragc Gupta, and Rushan Chen. "The Niagara Internet Query System," *IEEE Data Engineering Bulletin*, volume 24, number 2, pp. 27–33, June 2001. *(invited)*

Technical Reports

[R92] Xiangyao Yu, Matt Youill, Matthew E. Woicik, Abdurrahman Ghanem, Marco Serafini, Ashraf Aboulnaga, and Michael Stonebraker. "PushdownDB: Accelerating a DBMS using S3 Computation," Technical Report arXiv:2002.05837, Computing Research Repository (CoRR), February 2020. *(extended version of [C22])*

[R93] Aisha Mohamed, Ghadeer Abuoda, Abdurrahman Ghanem, Zoi Kaoudi, and Ashraf Aboulnaga. "RDF-Frames: Knowledge Graph Access for Machine Learning Tools," Technical Report arXiv:2002.03614, Computing Research Repository (CoRR), February 2020. *(extended version of [J1, D68])*

[R94] Xiaowei Zhu, Guanyu Feng, Marco Serafini, Xiaosong Ma, Jiping Yu, Lei Xie, Ashraf Aboulnaga, and Wenguang Chen. "LiveGraph: A Transactional Graph Storage System with Purely Sequential Adjacency List Scans," Technical Report arXiv:1910.05773, Computing Research Repository (CoRR), October 2019. *(extended version of [J3])*

[R95] Ghadeer Abuoda, Gianmarco De Francisci Morales, and Ashraf Aboulnaga. "Link Prediction via Higher-Order Motif Features," Technical Report arXiv:1902.06679, Computing Research Repository (CoRR), February 2019. *(extended version of [C24])*

[R96] Ahmed El-Roby, Khaled Ammar, Ashraf Aboulnaga, and Jimmy Lin. "Sapphire: Querying RDF Data Made Simple," Technical Report arXiv:1805.11728, Computing Research Repository (CoRR), May 2018. *(extended version of [D71])*

[R97] Michael J. Mior, Kenneth Salem, Ashraf Aboulnaga, and Rui Liu. "NoSE: Schema Design for NoSQL Applications," Technical Report CS-2015-17, Cheriton School of Computer Science, University of Waterloo, October 2015. *(extended version of [C28])*

[R98] Carlos H.C. Teixeira, Alexandre J. Fonseca, Marco Serafini, Georgios Siganos, Mohammed J. Zaki, and Ashraf Aboulnaga. "Arabesque: A System for Distributed Graph Pattern Mining," Technical Report arXiv:1510.04233, Computing Research Repository (CoRR), October 2015. *(extended version of [C29])*

[R99] Muhammad Bilal Sheikh, Umar Farooq Minhas, Omar Zia Khan, Ashraf Aboulnaga, Pascal Poupart, and David J. Taylor. "A Bayesian Approach to Online Performance Modeling for Database Appliances using Gaussian Models," Technical Report CS-2011-13, Cheriton School of Computer Science, University of Waterloo, December 2011. *(extended version of [C38])*

[R100] Mumtaz Ahmad, Ashraf Aboulnaga, Shivnath Babu, and Kamesh Munagala. "Interaction-Aware Scheduling of Report Generation Workloads," Technical Report CS-2010-22, Cheriton School of Computer Science, University of Waterloo, December 2010. *(preprint of [J16])*

[R101] Iman Elghandour, Ashraf Aboulnaga, Daniel C. Zilio, Fei Chiang, Andrey Balmin, Kevin Beyer, and Calisto Zuzarte. "XML Index Recommendation with Tight Optimizer Coupling," Technical Report CS-2007-22, Cheriton School of Computer Science, University of Waterloo, July 2007. *(extended version of [C51])*

[R102] Ashraf Aboulnaga and Jeffrey F. Naughton. "Towards Building XML Statistics for the Hidden Web," Technical Report TR-CS-03-1477, Computer Sciences Department, University of Wisconsin - Madison, April 2003. *(extended version of [C63])*

[R103] Ashraf Aboulnaga and Surajit Chaudhuri. "Accurate Query Optimization by Sub-plan Memoization," Technical Report MSR-TR-1999-102, Microsoft Research, December 1999. *(not otherwise published)*

Patents

[P104] Ashraf Aboulnaga, Andrey Balmin, Kevin Beyer, Iman Elghandour, Fei Chiang, Daniel C. Zilio, and Calisto Zuzarte. "Index Selection for XML Database Systems," US Patent No. 8,229,920, filed August 2007, issued July 2012.

[P105] Volker Markl, Peter J. Haas, Ashraf Aboulnaga, Vijayashankar Raman, and Felix Endres. "System and Method for Updating Database Statistics According to Query Feedback," US Patent No. 7,831,592, filed October 2004, issued November 2010.

[P106] Ashraf Aboulnaga, Peter J. Haas, Sam Lightstone, Volker Markl, Ivan Popivanov, and Vijayshankar Raman. "Method, System and Program for Prioritizing Maintenance of Database Tables," US Patent No. 7,363,324, filed December 2004, issued April 2008.

[P107] Surajit Chaudhuri and Ashraf Aboulnaga. "Query Optimization by Sub-plan Memoization," US Patent No. 6,850,925, filed May 2001, issued February 2005.

[P108] Ashraf Aboulnaga and Surajit Chaudhuri. "Self-tuning Histogram and Database Modeling," US Patent No. 6,460,045, filed March 1999, issued October 2002.

Research Funding at the University of Waterloo

All amounts are given as totals for the entire grant duration, in Canadian dollars. For grants with multiple investigators and/or where overhead was charged, my share of the grant, net of overhead, is given in parentheses.

- $50,000. Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant, "Systems and Algorithms for Big Data Analytics," 2013–2014.

- $318,316. Equipment Grant from Canada Foundation for Innovation (CFI) Leaders Opportunity Fund and Ontario Research Fund - Research Infrastructure (ORF-RI), "Research Infrastructure for Computational Analytics on Big Data in the Cloud," (Hans De Sterck co-PI), 2013–2014.

- $25,000. Natural Sciences and Engineering Research Council of Canada (NSERC) Strategic Network on Business Intelligence, "Feedback-driven Tuning of Hadoop Performance Parameters," PI: Frank Tompa (with 1 co-PI), 2012–2013.

- $25,000. Natural Sciences and Engineering Research Council of Canada (NSERC) Strategic Network on Business Intelligence, "Focusing on the Data in Web-Scale Data Integration," PI: M. Tamer Özsu (with 1 co-PI), 2012–2013.

- $3,347,385 (net: $113,395). Ontario Research Fund for Research Excellence, "Delivering Ultra-large-scale Services," PI: Patrick Martin, Queen's University (with 10 co-PIs), 2011–2016.

- $10,000. Amazon Web Services Research Grant (AWS credits in support of research), "Database Services on the Cloud," 2011.

- $190,000 (net: $150,000). Ontario Early Researcher Award (ERA) from the Ontario Ministry of Research and Innovation with matching from the University of Waterloo, "Improving Database Management in Cloud Computing Environments," 2009–2014.

- $93,250 (net: $71,200). Natural Sciences and Engineering Research Council of Canada (NSERC) Strategic Network on Business Intelligence, "Best-Effort Data Integration," PI: Frank Tompa (with 2 co-PIs), 2009–2012.

- $40,000 (net: $30,691). Google Research Award, "Structured Mashups on Demand," 2008.

- $135,000. Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant, "Robustness in Database Management," 2008–2013.

- $426,951 (net: $142,317). Natural Sciences and Engineering Research Council of Canada (NSERC) Collaborative Research and Development Grant, "Dynamic Resource Allocation for Virtual Database Appliances," PI: Kenneth Salem (with 2 co-PIs), 2008–2011.

- $1,724,802 (net: $178,666). Ontario Centres of Excellence (OCE), "Centre of Excellence for Research in Adaptive Systems (CERAS)," 11 co-PIs, 2007–2010.

- $85,000. IBM Centre for Advanced Studies (CAS), Toronto, "Database Virtualization," 2007–2009.

- $146,296 (net: $29,259). IBM Shared University Research (SUR) Equipment Grant, "Autonomic Computing Lab at the University of Waterloo," PI: Johnny Wong (with 4 co-PIs), 2006.

- $164,000. Natural Sciences and Engineering Research Council of Canada (NSERC) Collaborative Research and Development Grant, "Database Design and Tuning for XML Workloads," 2006–2010.

- $85,000. IBM Centre for Advanced Studies (CAS), Toronto, "XML Design Advisor," 2006–2008.

- $962,325 (net: $180,970). Equipment Grant from Canada Foundation for Innovation (CFI) New Opportunities Fund and Ontario Innovation Trust (OIT), "Resources For Computational Research Innovation," PI: Kevin Hare (with 13 co-PIs), 2006–2008.

- $66,000. Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant, "Advanced Querying of the Web," 2005–2008.

- $58,000. Software Telecommunications Group, University of Waterloo, "XML View Caching for the Efficient Execution of Web Queries," 2005–2007.

- $80,000. School of Computer Science, University of Waterloo, Startup Funds, 2004–2008.

Students Supervised

Current Ph.D. Students

- Humera Sabir.

Former Students and Postdoctoral Fellows

Qatar Computing Research Institute

Ph.D. Students

- Ghadeer Abuoda, May 2021. *(Ph.D. from Hamad Bin Khalifa University, co-supervised by Ting Yu)*
  Thesis: Advanced Data Processing for Better Machine Learning on Graphs.
  First Employment: Postdoctoral Fellow at Aalborg University, Denmark.

Research Assistants

- Aisha Mohamed, 2018–2020.
  First Employment: Ph.D. student at the University of Wisconsin - Madison.
- Eslam Hussein, 2016–2018.
  First Employment: Ph.D. student at Virginia Tech.

University of Waterloo

Postdoctoral Fellows

- Rui Liu, 2010–2013. *(co-supervised with Kenneth Salem)*
  First Employment: Senior Software Engineer at LinkedIn, Mountain View, California.
- Iman Elghandour, 2011–2012.
  First Employment: Assistant Professor at Alexandria University, Egypt.

Ph.D. Students

- Kareem El Gebaly, May 2018. *(co-supervised with Jimmy Lin and Lukasz Golab)*
  Thesis: Analytics for Everyone.
  First Employment: Software Engineer at the Huawei Canada Research Centre, Toronto.
- Ahmed El-Roby, April 2018.
  Thesis: Web Data Integration for Non-Expert Users.
  First Employment: Assistant Professor at Carleton University, Ottawa.
- Umar Farooq Minhas, December 2012.
  Thesis: Scalable and Highly Available Database Systems in the Cloud.
  First Employment: Research Staff Member at the IBM Almaden Research Center, San Jose, California.
- Mumtaz Ahmad, November 2012.
  Thesis: Query Interactions in Database Systems.
  First Employment: Software Developer at Auvik Networks, Waterloo.
- Iman Elghandour, October 2010.
  Thesis: Automatic Physical Design for XML Databases.
  First Employment: Postdoctoral Fellow at the University of Waterloo.

Master's (M.Math.) Students

- Huangdong Meng, January 2014.
  Thesis: High Availability for Database Systems in Geographically Distributed Cloud Computing Environments.
  First Employment: Software Development Engineer at Amazon, Toronto.

- Jingen Xiang, September 2013.
  Thesis: Scalable Scientific Computing Algorithms Using MapReduce.
  First Employment: Software Developer at SAP, Waterloo.

- Taha Rafiq, May 2013.
  Thesis: Elasca: Workload-Aware Elastic Scalability for Partition Based Database Systems.
  First Employment: Software Development Engineer at Amazon, Toronto.

- Mostafa Ead, December 2012.
  Thesis: PStorM: Profile Storage and Matching for Feedback-Based Tuning of MapReduce Jobs.
  First Employment: Software Development Engineer at Amazon, Vancouver.

- Hatem Mahmoud, June 2010.
  Thesis: Schema Clustering and Retrieval for Multi-domain Pay-As-You-Go Data Integration Systems.
  First Employment: Ph.D. student at the University of California - Santa Barbara.

- Sean Tozer, August 2009. *(co-supervised with Tim Brecht)*
  Thesis: Avoiding Bad Query Mixes to Minimize Unsuccessful Client Requests Under Heavy Loads.
  First Employment: Senior Developer at the Centre for Phenogenomics, Mount Sinai Hospital, Toronto.

- Ahmed Ataullah, May 2008. *(co-supervised with Frank Tompa)*
  Thesis: A Framework for Records Management in Relational Database Systems.
  First Employment: Ph.D. student at the University of Waterloo.

- Nickolay Tchervenski, February 2008.
  Essay: An Evaluation of Storage Load Balancing Alternatives for Database Systems.
  First Employment: Software Developer at the IBM Toronto Laboratory.

- Umar Farooq Minhas, December 2007.
  Thesis: A Performance Evaluation of Database Systems on Virtual Machines.
  First Employment: Ph.D. student at the University of Waterloo.

- Kareem El Gebaly, August 2007.
  Thesis: Robustness in Automatic Physical Database Design.
  First Employment: Software Developer at Ericsson Inc., Egypt.

Undergraduate Interns and Co-op Students

- Fábio Albuquerque (Mitacs Globalink intern), *Universidade Federal da Paraíba*, 2013.
- Aniruddha Gangopadhyay (Intern), *Birla Institute of Technology and Science (BITS)*, 2012.
- Sean Robertson (Co-op student, NSERC USRA), *University of Waterloo*, 2011.
- Jonathan Ng (Co-op student, NSERC USRA), *University of Waterloo*, 2011.
- Ziyu Wang (Co-op student, NSERC USRA), *University of Waterloo*, 2008.
- Zi Ye Zhang (Co-op student, NSERC USRA), *University of Waterloo*, 2008.
- Harikrishna Bikmal (Intern), *Indian Institute of Technology, Bombay*, 2008.
- Jitendra Yadav (Intern), *Indian Institute of Technology, Kanpur*, 2007.
- Hidayath Ansari (Intern), *Indian Institute of Technology, Bombay*, 2006.
- Mayank Singhal (Intern), *Indian Institute of Technology, Bombay*, 2006.
- Daniel Wong (Co-op student), *University of Waterloo*, 2006.
- Daniel Brotherston (Co-op student), *University of Waterloo*, 2006.

Membership on Thesis Committees

- Member of the Ph.D. thesis committee of:
  Mohammad Shaito, 2023–present.

- Member of the M.S. thesis committee of:
  Kiran Bolaj, 2023.
  Arshdeep Singh, 2023.

University of Waterloo

- Ph.D. external examiner for:
  Jiang Du, University of Toronto, 2019.
  Aysan Rasooli, McMaster University, 2013.
  Madalin Mihailescu, University of Toronto, 2012.
  Jun Hu, University of Ottawa, 2011.
  Samir Mohammad, Queen's University, 2011.

- Member of the Ph.D. thesis committee of:
  Khalid Alsubhi, 2015.
  Xin Liu, 2014.
  Patrick Kling, 2012.
  Xuhui Li, 2011.
  Chen Zhang, 2011.
  Rolando Blanco, 2010.
  Loubna Mekouar, 2010.
  David DeHaan, 2009.

- Master thesis reader for:
  Alexander Szlavik, 2013.
  Alexey Karyakin, 2011.
  Prashant Gaharwar, 2010.
  Tim Benke, 2009.
  Ye Hu, 2009.
  Mosharaf Chowdhury, 2009.
  Weihan Wang, 2007.
  Mohammed Abouzour, 2007.
  E. Cem Sözgen, 2006.
  Jie Zhang, 2005.

- Master essay reader for:
  Yasemin Ugur Ozekinci, 2007.

Selected Recent Talks

- Elastic Provisioning and Pushdown Computation for Cloud Databases.
  University of Sydney, September 2022.

- One-Stop Data Management and Service in the Cloud with Cloud Native Databases.
  Panel at the IEEE International Conference on Data Engineering (ICDE), May 2022.

- Databases on the Cloud - Architectural Implications.
  Mohammed VI Polytechnic University, Ben Guerir, Morocco, April 2021. *(on-line)*
  Keynote at the International Conference on Database Systems for Advanced Applications (DASFAA), Chiang Mai, Thailand, April 2019.

- Arabesque: A System for Distributed Graph Pattern Mining.
  Carnegie Mellon University, October 2016.
  University of Waterloo, October 2015.

- Elasticity and High Availability for Databases in the Cloud.
  Tsinghua University, September 2014.

- Self-Tuning Hadoop.
  Qatar University, May 2014.
  Carnegie Mellon University - Qatar, February 2014.
  University of Toronto, June 2012.

- Workload Management for Big Data Analytics.
  IBM Workshop on Big Data Analytics, June 2013.

- High Availability for Database Systems in Cloud Computing Environments.
  SAP Waterloo, December 2012.
  Workshop on Ultra-large Scale Services, CASCON Conference, November 2012.
  Keynote at the International Conference on Informatics and Systems (INFOS), Cairo, Egypt, May 2012.
  Qatar Computing Research Institute, May 2012.
  King Abdullah University of Science and Technology, April 2012.
  University of Washington, November 2011.
  Microsoft Research, November 2011.
  Amazon, November 2011.
  McMaster University, October 2011.
  Egyptian Japanese University of Science and Technology (E-JUST), July 2011.

- Multi-domain Pay-as-you-go Data Integration Systems.
  IBM Cognos, Ottawa, December 2010.

- Optimizing Complex Analytics Workloads on the Cloud.
  Workshop on Web Data Processing on the Cloud, CASCON Conference, November 2010.

Teaching Activities

Courses Taught

- CSE 4331/5331 - DBMS Models and Implementation Techniques.
  Fall 2023 (1 section)
- CSE 6331 - Advanced Topics in Database Systems (Data Infrastructure for Machine Learning).
  Spring 2024 (1 section)

Teaching at the University of Waterloo

Courses Taught

- CS 338 - Computer Applications in Business: Databases.
  Winter 2005 (1 section, 81 students).
- CS 348 - Introduction to Database Management.
  Winter 2013 (1 section, 94 students), Fall 2007 (1 section, 50 students), Winter 2007 (2 sections, 156 students).
- CS 350 - Operating Systems.
  Fall 2012 (2 sections, 113 students), Fall 2011 (2 sections, 119 students), Winter 2011 (2 sections, 129 students), Winter 2009 (2 sections, 121 students), Spring 2006 (2 sections, 128 students).
- CS 848 - Advanced Topics in Data Bases: Databases in Cloud Computing Environments.
  Winter 2013 (14 students), Winter 2012 (20 students), Fall 2008 (16 students).
- CS 848 - Advanced Topics in Data Bases: Self-managing Database Systems.
  Fall 2007 (11 students), Winter 2007 (11 students), Winter 2006 (6 students), Spring 2005 (13 students).

Course Development

- CS 350 - Operating Systems (with Tim Brecht and Kenneth Salem).
  Winter 2011: Migrated the infrastructure for course projects from Solaris to Linux.
  Winter 2007: Updated the infrastructure for course projects to use OS/161 instead of Nachos.

- CS 848 - Advanced Topics in Data Bases: Databases in Cloud Computing Environments.
  Fall 2008: New course introduced.
- CS 348 - Introduction to Database Management.
  Winter 2007: Updated course material and infrastructure for programming assignments.
- CS 848 - Advanced Topics in Data Bases: Self-managing Database Systems.
  Spring 2005: New course introduced.

Scientific Activities

Journal Editing

- Member of the Editorial Board of Distributed and Parallel Databases (DAPD), 2014–present.
- Associate Editor of the VLDB Journal, 2015–2021.
- Guest Editor of the VLDB Journal Special Section on the Best Papers of the Workshop on Data Management on New Hardware (DaMoN) 2019, November 2020.
- Associate Editor of the IEEE Transactions on Knowledge and Data Engineering (TKDE), 2012–2016.
- Guest Co-editor (with Beng Chin Ooi and Patrick Valduriez) of the VLDB Journal Special Section on Data-intensive Cloud Infrastructure, December 2014.

Conference and Workshop Organization

- Senior Program Committee Member of the Conference on Extending Database Technology (EDBT), Barcelona, Spain, March 2025.
- Associate Editor of the Proceedings of the VLDB Endowment (PVLDB), volume 17, 2023–2024.
  Papers published in this volume of PVLDB presented at the International Conference on Very Large Databases (VLDB), Guangzhou, China, August 2024.
- Associate Editor of the ACM SIGMOD International Conference on Management of Data (SIGMOD), Santiago, Chile, June 2024.
- Industry and Applications Track Program Committee Co-Chair (with Peter Boncz and Fatma Özcan) of the IEEE International Conference on Data Engineering (ICDE), Utrecht, Netherlands, April 2024.
- Senior Program Committee Member of the Conference on Extending Database Technology (EDBT), Paestum, Italy, March 2024.
- Industrial Track Program Committee Co-Chair (with Avrilia Floratou) of the ACM SIGMOD International Conference on Management of Data (SIGMOD), Philadelphia, Pennsylvania, USA, June 2022.
- Associate Editor of the Proceedings of the VLDB Endowment (PVLDB), volume 14, 2020–2021.
  Papers published in this volume of PVLDB presented at the International Conference on Very Large Databases (VLDB), Copenhagen, Denmark, August 2021.
- Program Committee Vice-chair for Distributed, Parallel, and P2P Data Management at the IEEE International Conference on Data Engineering (ICDE), Paris, France, April 2018.
- Associate Editor of the Proceedings of the VLDB Endowment (PVLDB), volume 10, 2016–2017.
  Papers published in this volume of PVLDB presented at the International Conference on Very Large Databases (VLDB), Munich, Germany, August 2017.
- Program Committee Track Chair for Cloud Computing and Database-as-a-Service at the IEEE International Conference on Data Engineering (ICDE), Helsinki, Finland, April 2016.
- Co-chair (with Barzan Mozafari and Ippokratis Pandis) of the Workshop on Cloud Data Management (CloudDM), 2016.
  Held in conjunction with the IEEE International Conference on Data Engineering (ICDE), Helsinki, Finland, April 2016.

- Associate Editor of the Proceedings of the VLDB Endowment (PVLDB), volume 6, 2012–2013.
  Papers published in this volume of PVLDB presented at the International Conference on Very Large Databases (VLDB), Riva del Garda, Italy, August 2013.

- Co-chair (with Carlo Curino) of the International Workshop on Data Management in the Cloud (DMC), 2013.
  Held in conjunction with the IEEE International Conference on Data Engineering (ICDE), Brisbane, Australia, April 2013.

- Senior Program Committee Member of the ACM Conference on Information and Knowledge Management (CIKM), Maui, Hawaii, USA, October 2012.

- Co-chair (with Kenneth Salem) of the International Workshop on Self-Managing Database Systems (SMDB), 2009.
  Held in conjunction with the IEEE International Conference on Data Engineering (ICDE), Shanghai, China, March 2009.

Membership on Professional Committees

- Member of the SIGMOD Jim Gray Doctoral Dissertation Award Committee, 2013–2017 (co-chair 2015–2017).

- Member of the Executive Committee of the IEEE Computer Society Data Engineering Workgroup on Self-Managing Database Systems, 2010–2020.

Advisory Board Membership

- ClevrU, 2011–2014.

Program Committee Membership

- ACM SIGMOD International Conference on Management of Data: 2006, 2007 (Demonstration Committee), 2009, 2010 (Best Demonstration Selection Committee), 2011–2014 (all years), 2016, 2017, 2017 (Demonstration Committee), 2018 (***Distinguished Program Committee Member Award***), 2018 (Demonstration Committee), 2019 (Core Program Committee, Best Paper Selection Committee), 2020, 2023, 2023 (Industrial Track, Best Paper Selection Committee).

- International Conference on Very Large Databases (VLDB) and Proceedings of the VLDB Endowment (PVLDB): 2007 (Demonstration Committee), 2009–2012 (all years), 2015, 2016, 2018, 2019 (Best Paper Selection Committee), 2020 (***Distinguished Review Board Member Award***), 2023.

- IEEE International Conference on Data Engineering (ICDE): 2005–2009 (all years), 2009 (Demonstration Committee), 2011, 2013, 2014, 2015, 2017, 2018 (Industrial Track), 2019, 2019 (Industrial Track), 2020, 2024.

- Conference on Extending Database Technology (EDBT): 2006, 2015, 2021, 2023.

- ACM Conference on Information and Knowledge Management (CIKM): 2006, 2008, 2011, 2013, 2014, 2015.

- ACM Symposium on Cloud Computing (SoCC): 2013, 2014, 2015, 2017, 2018.

- ACM SIGKDD Knowledge Discovery and Data Mining Conference (KDD): 2013.

- IEEE International Conference on Distributed Computing Systems (ICDCS): 2007, 2009.

- International Conference on Web Information System Engineering (WISE): 2012.

- ACM International Workshop on Web Information and Data Management (WIDM): 2004, 2005, 2007, 2008.

- International Conference on Scientific and Statistical Database Management (SSDBM): 2007, 2013.

- International XML Database Symposium (XSym): 2005, 2006, 2007, 2009.

- USENIX Conference on File and Storage Technologies (FAST): 2007.

- IEEE International Conference on Big Data (BigData): 2015.

- Workshop on Serverless Data Analytics (SDA): 2023.

- International Workshop on Self-Managing Database Systems (SMDB): 2011, 2012, 2014, 2019.

- International Workshop on Reliable Data Services and Systems (RDSS): 2014.
- International Workshop on End-to-end Management of Big Data (BigData): 2012.
- Conference of the Centre for Advanced Studies on Collaborative Research (CASCON): 2009, 2011 (Ten-year Most Influential Paper Selection Committee), 2015.
- SIGMOD/PODS Ph.D. Symposium: 2012.
- VLDB Ph.D. Workshop: 2007, 2008.

Grant Reviewing for Government Funding Agencies

- Natural Sciences and Engineering Research Council of Canada (NSERC):
  Reviewer for Discovery Grants, 2005–2014 (all years), 2017, 2018, 2020.
  Reviewer for Idea to Innovation Grants, 2015.
  Reviewer for Strategic Project Grants, 2013.
  Reviewer for Collaborative Research and Development Grants, 2005, 2013, 2016, 2017, 2018, 2019.
- National Science Foundation, USA:
  Panelist on Division of Information and Intelligent Systems (IIS) Grant Review Panel, 2003 (two panels), 2011.
  External reviewer for Division of Information and Intelligent Systems (IIS), 2010.
- Research Network on Mathematics of Information Technology and Complex Systems (MITACS), Canada:
  Reviewer for Accelerate Program, 2010, 2012, 2013, 2014, 2016, 2019.
  Reviewer for Elevate Program, 2013, 2016, 2018, 2019.
- Canada Research Coordinating Committee (CRCC):
  Member of the Review Panel for New Frontiers in Research Fund (Exploration), 2019, 2021.
- Canada Research Chairs:
  Reviewer for Tier 2 Canada Research Chairs, 2016.
- Qatar National Research Fund:
  Reviewer for National Priorities Research Program, 2009.
  Reviewer for Undergraduate Research Experience Program, 2007–2009 (all years).
- Research Grant Council (RGC) of Hong Kong:
  Reviewer for General Research Fund, 2013, 2014.
  Reviewer for Collaborative Research Fund, 2013.

Journal Reviewing

- ACM Transactions on Database Systems, 2001–2007 (all years), 2013, 2015, 2017.
- The VLDB Journal, 2004, 2007, 2008, 2009, 2010, 2013, 2014, 2017.
- IEEE Transactions on Knowledge and Data Engineering, 2003–2006 (all years), 2009, 2010, 2011.
- ACM Computing Surveys, 2008.
- IEEE Computer, 2011.
- Information Systems, 2011, 2012, 2014, 2015.
- Distributed and Parallel Databases, 2005, 2008, 2010, 2011, 2012, 2014.
- SIGMOD Record, 2012, 2015.
- Data and Knowledge Engineering, 2001, 2004, 2005.
- Egyptian Informatics Journal, 2013.
- Kuwait Journal of Science and Engineering, 2011.

Administrative Service

- Member of the following Department of Computer Science and Engineering committees:
  Industry Outreach Committee, 2023–2025.
  Ph.D. Admissions Committee, 2023–2025.
  Broadening Participation in Computing Committee, 2023–2025.
  Faculty Search Committee, 2023–2024.

Qatar Computing Research Institute

- Chief Scientist, 2021–2023.

- Senior Research Director, 2019–2021.

- Manager of the Distributed Systems Group, 2013–2018.

University of Waterloo

- Member of the Software Engineering Board, 2011–2013.

- Chair of the Cheriton School of Computer Science Web Committee, 2011–2013.

- Member of the Faculty of Mathematics Web Design Committee, 2011–2012.

- Member of the following Cheriton School of Computer Science committees:
  Graduate Academic Plans Committee, 2010–2013.
  Tenure and Promotion Committee, 2011–2012, 2012–2013.
  Graduate Recruiting Committee, 2010–2012.
  Undergraduate Recruiting Committee, 2005–2009.
  External Relations Committee, 2004–2005.

October 29, 2023