Curriculum Vitae

Boris Glavic
Department of Computer Science
851 S. Morgan St.
9th Floor SEO 940
Chicago, IL 60607, USA

Email: bglavic@uic.edu
Homepage: http://www.cs.uic.edu/~bglavic/
Group Webpage: http://www.cs.uic.edu/~bglavic/dbgroup
Phone: +1 (312) 355 2333 (work)

Research Interests

Database systems, in particular, data provenance, explanations, uncertain data management, query processing, and data cleaning and data integration.

Education

2010 **Doktor (Ph.D.)**, summa cum laude, in Computer Science
University of Zurich, Switzerland
Thesis Advisors: Michael H. Böhlen, Gustavo Alonso
Thesis: Perm: Efficient Provenance Support for Relational Databases

2005 **Diploma (M.Sc.)** in Computer Science
RWTH Aachen, Germany
Thesis Advisors: Ralph Krieger, Ira Assent
Thesis: Subspace Clustering for Sequences of Ordered Categorical Data

Employment and Appointments

2024– **Associate Professor**
University of Illinois at Chicago (UIC), Chicago, USA

2021– **Founder & Technical Advisory Board Member**
Breadcrumb Analytics, Buffalo, USA

2018–2023 **Associate Professor**
2012–2018 **Assistant Professor (Tenure-track)**
IIT, Chicago, USA

2017– **Technical Advisory Board Member**
Ocient, Chicago, USA

2016–2019 **Adjunct Assistant Professor (Status only)**
University of Toronto, Toronto, Canada

2010–2012 **Postdoctoral Fellow**
University of Toronto, Canada
Supervisor: Renée J. Miller

2005–2010 **Ph.D. Student and Research Assistant**
University of Zurich, Switzerland
Supervisors: Klaus R. Dittrich, Carl-Christian Kanne, Michael H. Böhlen

Awards

2023	PVLDBVol16 Distinguished Associate Editor
	EDBT 2023 Distinguished PC Member Award
	Mercator Fellow in FONDA – Foundations of Workflows for Large-Scale Scientific Data Analysis (Deutsche Forschungsgemeinschaft)
2022	SIGMOD Distinguished Program Committee Member Award
	Mercator Fellow in FONDA – Foundations of Workflows for Large-Scale Scientific Data Analysis (Deutsche Forschungsgemeinschaft)
2021	SIGMOD Distinguished Program Committee Member Award
	VLDB Distinguished Reviewer Award
2020	SIGMOD Distinguished Program Committee Member Award
	SIGMOD Reproducibility Award
2019	SIGMOD Distinguished Reviewer Award
	VLDB Distinguished Reviewer Award
	ICDE Outstanding Reviewer Award
2017	Best of ICDE invited IEEE TKDE Journal article

Publications

Proceedings Editor

[1] **Boris Glavic**, Vanessa Braganholo, David Koop. *Provenance and Annotation of Data and Processes - 8th and 9th International Provenance and Annotation Workshop,* **IPAW 2020** + **IPAW 2021**, *Virtual Event, July 19-22, 2021, Proceedings.* Lecture Notes in Computer Science, Springer 2021, ISBN 978-3-030-80959-1.

[2] Marta Mattoso, **Boris Glavic**. *Provenance and Annotation of Data and Processes - 6th International Provenance and Annotation Workshop,* **IPAW 2016**, *McLean, VA, USA, June 7-8, 2016, Proceedings*, Lecture Notes in Computer Science 9672, Springer 2016, ISBN 978-3-319-40592-6.

Guest Editorials

[1] Mohamed Eltabakh, **Boris Glavic**. *Guest editorial: large-scale data curation and metadata management*, Distributed and Parallel Databases (**DaPD**), Springer 2018, ISBN 0926-8782

Peer-Reviewed Journals

[1] Alexandru Iulian Orhean, Anna Giannakou, Lavanya Ramakrishnan, Kyle Chard, **Boris Glavic**, Ioan Raicu. *SCIPIS: Scalable and Concurrent Persistent Indexing and Search in High-End Computing Systems.* Journal of Parallel and Distributed Computing **J. Parallel Distributed Comput.**, vol. 189, 104878, 2024.

[2] Ricardo Salazar-Díaz, **Boris Glavic**, Tilmann Rabl. *InferDB: In-Database Machine Learning Inference Using Indexes.* Proceedings of the VLDB Endowment (**PVLDB**), vol. 16 (6), 1346 - 1358, 2024.

[3] Su Feng, **Boris Glavic**, Oliver Kennedy. *Efficient Approximation of Certain and Possible Answers for Ranking and Window Queries over Uncertain Data.* Proceedings of the VLDB Endowment (**PVLDB**), vol. 16 (6), 1346-1358, 2023.

[4] Oliver Kennedy, **Boris Glavic**, Juliana Freire, Mike Brachmann. *The Right Tool for the Job: Data-Centric Workflows in Vizier.* **IEEE Data Engineering Bulletin**, vol. 45 (3), 129–144, 2022.

[5] Chenjie Li, Juseung Lee, Zhengjie Miao, **Boris Glavic**, Sudeepa Roy. *CaJaDE: Explaining Query Results by Augmenting Provenance with Context.* Proceedings of the VLDB Endowment (Demonstration Track) (**PVLDB**), vol. 15 (12), 2022.

[6] Xing Niu, Ziyu Liu, Pengyuan Li, **Boris Glavic**, Dieter Gawlick, Vasudha Krishnaswamy, Zhen Hua Liu, Danica Porobic. *Provenance-Based Data Skipping.* Proceedings of the VLDB Endowment (**PVLDB**), vol. 15 (3), 451 - 464, 2021.

[7] Ralf Diestelkämper, Seokki Lee, Melanie Herschel, **Boris Glavic**. *Debugging Missing Answers for Spark Queries over Nested Data with Breadcrumb.* Proceedings of the VLDB Endowment (Demonstration Track) (**PVLDB**), vol. 14 (12), 2731 - 2734, 2021.

[8] **Boris Glavic**, Alexandra Meliou, Sudeepa Roy. *Trends in Explanations: Understanding and Debugging Data-driven Systems.* Foundations and Trends in Databases (**FnT**), vol. 11 (3), 226-318, 2021.

[9] **Boris Glavic**. *Data Provenance - Origins, Applications, Algorithms, and Models.* Foundations and Trends in Databases (**FnT**), vol. 9, 209 - 441, 2021.

[10] Seokki Lee, Bertram Ludäscher, **Boris Glavic**. *Approximate Summaries for Why and Why-not Provenance.* (**PVLDB**), vol. 13 (6), 912 - 924, 2020.

[11] Seokki Lee, Bertram Ludäscher, **Boris Glavic**. *PUG: a framework and practical implementation for why and why-not provenance.* (**VLDBJ**), vol. 28, 47—71, 2018.

[12] Zhengjie Miao, Qitian Zeng, Chenjie Li, **Boris Glavic**, Oliver Kennedy, and Sudeepa Roy. *Cape: explaining outliers by counterbalancing.* Proceedings of the VLDB Endowment (**PVLDB**) (Demonstration Track), vol. 12, no. 12, 1806-1809, 2019.

[13] Anton Dignös, **Boris Glavic**, Xing Niu, Michael H. Böhlen and Johann Gamper, *Snapshot Semantics for Temporal Multiset Relations* Proceedings of the VLDB Endowment (**PVDLB**), vol. 12, no. 6, 639–652, 2019.

[14] Xing Niu, Raghav Kapoor, **Boris Glavic**, Dieter Gawlick, Zhen Hua Liu, Vasudha Krishnaswamy, Venkatesh Radhakrishnan, *Heuristic and Cost-based Optimization for Diverse Provenance Tasks* IEEE Transactions on Knowledge and Data Engineering (**TKDE**), vol. 31, no. 7, 1267–1280, 2019.

[15] Seokki Lee, Bertram Ludäscher, **Boris Glavic**, *Provenance Summaries for Answers and Non-Answers.* Proceedings of the VLDB Endowment (**PVLDB**) (Demonstration Track), vol. 11, no. 12, 1954–1957, 2018.

[16] Bahareh Arab, Su Feng, **Boris Glavic**, Seokki Lee, Xing Niu, Qitian Zeng. *GProM - A Swiss Army Knife for Your Provenance Needs.* (**Data Eng. Bull.**), vol. 41 (1), 51-62, 2018.

[17] Bahareh S. Arab, Dieter Gawlick, Vasudha Krishnaswamy, Venkatesh Radhakrishnan, **Boris Glavic**. *Using Reenactment to Retroactively Capture Provenance for Transactions* IEEE Transactions on Knowledge and Data Engineering (**TKDE**), vol. 30 (3), 599-612, 2018.

[18] Xing Niu, Bahareh S. Arab, Seokki Lee, Su Feng, Xun Zou, Dieter Gawlick, Vasudha Krishnaswamy, Zhen Hua Liu, **Boris Glavic**. *Debugging Transactions and Tracking their Provenance with Reenactment.* Proceedings of the VLDB Endowment (**PVLDB**), vol. 10, no. 12, 1857-1860, 2017.

[19] Patricia C. Arocena, **Boris Glavic**, Giansalvatore Mecca, Renée J. Miller, Paolo Papotti, and Donatello Santoro, *Benchmarking Data Curation Systems*, In **IEEE Data Engineering Bulletin**, vol. 39, no.2, 47-62, 2016.

[20] Patricia C. Arocena, **Boris Glavic**, Radu Ciucanu, and Renée J. Miller. *The iBench Integration Metadata Generator.* In Proceedings of the VLDB Endowment (**PVLDB**), vol. 9, no. 3, 108-119, 2015.

[21] Patricia C. Arocena, **Boris Glavic**, Giansalvatore Mecca, and Renée J. Miller, Paolo Papotti, Donatello Santoro. *Messing Up with BART: Error Generation For Evaluating Data-Cleaning Algorithms.* In Proceedings of the VLDB Endowment (**PVLDB**), vol. 9, no. 2, 36-47, 2015.

[22] Patricia C. Arocena, Radu Ciucanu, **Boris Glavic**, and Renée J. Miller. *Gain Control over your Integration Evaluations.* In Proceedings of the VLDB Endowment (**PVLDB**) (Demonstration Track), vol. 8, no. 12, 1960 - 1971, 2015.

[23] Quan Pham, Richard Whaling, **Boris Glavic**, and Tanu Malik. *Sharing and reproducing database applications.* Proceedings of the VLDB Endowment (**PVLDB**) (Demonstration Track), vol. 8, no. 12, 1988 - 1999, 2015.

[24] **Boris Glavic**, Kyumars Sheykh Esmaili, Peter M. Fischer, and Nesime Tatbul. *Efficient Stream Provenance via Operator Instrumentation.* In Transactions on Internet Technology (**TOIT**),vol. 14, no. 1, 7:1-7:26, 2014.

[25] **Boris Glavic**, Jiang Du, Renée J. Miller, Gustavo Alonso, and Laura M. Haas. *Debugging Data Exchange with Vagabond.* In Proceedings of the Very Large Data Bases Endowment (**PVLDB**) (Demonstration Track), vol. 4, no. 12, 1383-1386, 2011.

[26] **Boris Glavic**, Gustavo Alonso, Renée J. Miller, and Laura M. Haas. *TRAMP: Understanding the Behavior of Schema Mappings through Provenance.* In Proceedings of the Very Large Data Bases Endowment (**PVLDB**), vol. 3, no. 1, 1314-1325, 2010.

[27] Ira Assent, Ralph Krieger, **Boris Glavic**, and Thomas Seidl. *Clustering Multidimensional Sequences in Spatial and Temporal Databases.* In International Journal on Knowledge and Information Systems (**KAIS**), vol. 16, no. 1, 29-51, 2008.

Peer-Reviewed Conferences

[1] Nils Strassenburg, **Boris Glavic**, Tilmann Rabl. *Alsatian: Optimizing Model Search for Deep Transfer Learning.* SIGMOD (**SIGMOD**), 2025 (to appear).

[2] Aaron Huber, Oliver Kennedy, Atri Rudra, Zhuoyue Zhao, Su Feng, **Boris Glavic**. *FastPDB: Towards Bag-Probabilistic Queries at Interactive Speeds.* Proceedings of the 51st International Conference on Management of Data (**SIGMOD**), 2025 (to appear).

[3] Jiongli Zhu, Su Feng, **Boris Glavic**, Babak Salimi. *Learning from Uncertain Data: From Possible Worlds to Possible Models.* NeurIPS (**NeurIPS**), 2024 (to appear).

[4] **Boris Glavic**, Giansalvatore Mecca, Renée J. Miller, Paolo Papotti, Donatello Santoro, Enzo Veltri. *Similarity Measures for Incomplete Database Instances.* Proceedings of the International Conference on Extending Database Technology (**EDBT**), pp. 461–473, 2024.

[5] **Boris Glavic**, Pengyuan Li, Ziyu Liu, Dieter Gawlick, Vasudha Krishnaswamy, Danica Porobic, Zhen Hua Liu. *Towards an Objective Metric for Data Value through Relevance.* Proceedings of the 14th Conference on Innovative Data Systems (**CIDR**), 2024.

[6] Jiongli Zhu, Romila Pradhan, **Boris Glavic**, Babak Salimi. *Generating Interpretable Data-Based Explanations for Fairness Debugging Using Gopher.* Proceedings of the 48th International Conference on Management of Data (SIGMOD) (Demonstration Track) (**SIGMOD**), 2433–2436, 2022.

[7] Babak Salimi, Romila Pradhan, Jiongli Zhu, **Boris Glavic**. *Interpretable Data-Based Explanations for Fairness Debugging.* Proceedings of the 48th International Conference on Management of Data (**SIGMOD**), 247–261, 2022.

[8] Felix Campbell, Bahareh Arab, **Boris Glavic**. *Efficient Answering of Historical What-If Queries.* Proceedings of the 48th International Conference on Management of Data (**SIGMOD**), 1556-1569, 2022.

[9] Ralf Diestelkämper, Seokki Lee, Melanie Herschel, **Boris Glavic**. *To not miss the forest for the trees - A holistic approach for explaining missing answers over nested data.* Proceedings of the 46th International Conference on Management of Data (**SIGMOD**), 405–417, 2021.

[10] Chenjie Li, Zhengjie Miao, Qitian Zeng, **Boris Glavic**, Sudeepa Roy. *Putting Things into Context: Rich Explanations for Query Answers using Join Graphs.* Proceedings of the 46th International Conference on Management of Data (**SIGMOD**), 1051–1063, 2021.

[11] Su Feng, Aaron Huber, **Boris Glavic**, Oliver Kennedy. *Efficient Uncertainty Tracking for Complex Queries with Attribute-level Bounds.* Proceedings of the 46th International Conference on Management of Data (**SIGMOD**), 528 – 540, 2021.

[12] Poonam Kumari, Michael Brachmann, Oliver Kennedy, Su Feng, and Boris Glavic. *DataSense: Display Agnostic Data Documentation.* In Proceedings of the 11th Conference on Innovative Data Systems (**CIDR**), 2021.

[13] Michael Brachmann, William Spoth, Oliver Kennedy, Boris Glavic, Heiko Müller, Sonia Castel, Carlos Bautista, and Juliana Freire. *Your notebook is not crumby enough, replace it.* In Proceedings of the 10th Conference on Innovative Data Systems (**CIDR**), 2020.

[14] Zhengjie Miao, Qitian Zeng, **Boris Glavic**, Sudeepa Roy. *Going Beyond Provenance: Explaining Query Answers with Pattern-based Counterbalances*, Proceedings of the 44th International Conference on Management of Data (**SIGMOD**), 485-502, 2019

[15] Jason Arnold, **Boris Glavic**, Ioan Raicu. *A High-Performance Distributed Relational Database System for Scalable OLAP Processing*, Proceedings of the 33rd IEEE International Parallel and Distributed Processing Symposium (**IPDPS**), 2019

[16] Su Feng, Aaron Huber, **Boris Glavic**, Oliver Kennedy, *Uncertainty Annotated Databases - A Lightweight Approach for Approximating Certain Answers*, Proceedings of the 44th International Conference on Management of Data (**SIGMOD**), 1313-1330, 2019

[17] Mike Brachmann, Carlos Bautista, Sonia Castelo, Su Feng, Juliana Freire, **Boris Glavic**, Oliver Kennedy, Heiko Müller, Rémi Rampin, William Spoth, Ying Yang, *Data Debugging and Exploration with Vizier*, Proceedings of the 44th International Conference on Management of Data (**SIGMOD**) (Demonstration Track), 1877-1880, 2019

[18] Donatello Santoro, Patricia C. Arocena, **Boris Glavic**, Giansalvatore Mecca, Renée J. Miller, Paolo Papotti, *Let's Make It Dirty with BART!*, Proceedings of the 26th Italian Symposium on Advanced Database Systems, 2018

[19] Seokki Lee, Sven Köhler, Bertram Ludäscher, **Boris Glavic**, *A SQL-Middleware Unifying Why and Why-Not Provenance for First-Order Queries*, In Proceedings of the 33rd IEEE International Conference on Data Engineering (**ICDE**), 485–496, 2017

[20] Xing Niu, Raghav Kapoor, **Boris Glavic**, Dieter Gawlick, Zhen Hua Liu, Vasudha Krishnaswamy, Venkatesh Radhakrishnan, *Provenance-aware Query Optimization*, In Proceedings of the 33rd IEEE International Conference on Data Engineering (**ICDE**), 473–484, 2017

[21] Jiang Du, **Boris Glavic**, Wei Tan, Renée J. Miller, *DeepSea: Adaptive Workload-Aware Partitioning of Materialized Views in Scalable Data Analytics*, In Proceedings of the 20th International Conference on Extending Database Technology (**EDBT**), 198-209, 2017

[22] William Spoth, Bahareh Arab, Eric S. Chan, Dieter Gawlick, Adel Ghoneimy, **Boris Glavic**, Beda Hammerschmidt, Oliver Kennedy, Seokki Lee, Zhen Hua Liu, Xing Niu, Ying Yang, *Adaptive Schema Databases*, In Proceedings of the 8th Biennial Conference on Innovative Data Systems (**CIDR**), 2017

[23] James Wagner, Alexander Rasin, **Boris Glavic**, Karen Heart, Jacob Furst, Lucas Bressan and Jonathan Grier. *Carving Database Storage to Detect and Trace Security Breaches*, Digital Forensics Research Conference (**DFRWS**), 2017

[24] Bahareh Arab, Dieter Gawlick, Vasudha Krishnaswamy, Venkatesh Radhakrishnan, and **Boris Glavic**, *Reenactment for Read-Committed Snapshot Isolation*, In Proceedings of the 25th International Conference on Information and Knowledge Management (**CIKM**), 841-850, 2016

[25] Donatello Santoro, Patricia C. Arocena, **Boris Glavic**, Giansalvatore Mecca, Renée J. Miller, and Paolo Papotti, *BART in Action: Error Generation and Empirical Evaluations of Data-Cleaning Systems*, In Proceedings of the 42nd International Conference on Management of Data (**SIGMOD**) (Demonstration Track), 2161–2164, 2016

[26] Quan Pham, Tanu Malik, **Boris Glavic**, and Ian Foster. *LDV: Light-weight Database Virtualization*, In Proceedings of the 31st IEEE International Conference on Data Engineering (**ICDE**), 1179-1190, 2015

[27] Patricia C. Arocena, **Boris Glavic**, and Renée J. Miller. *Value Invention for Data Exchange.* In Proceedings of the 39th International Conference on Management of Data (**SIGMOD**), 157-168, 2013

[28] **Boris Glavic**, Kyumars Sheykh Esmaili, Peter M. Fischer, and Nesime Tatbul. *Ariadne: Managing Fine-Grained Provenance on Data Streams.* In Proceedings of the 7th ACM International Conference on Distributed Event-Based Systems (**DEBS**), 39-50, 2013

[29] Christian Tilgner, **Boris Glavic**, Michael H. Böhlen, and Carl-Christian Kanne. *Declarative Serializable Snapshot Isolation.* In Proceedings of the 15th International Conference on Advances in Database and Information Systems (**ADBIS**), 170-184, 2011

[30] Christian Tilgner, **Boris Glavic**, Michael H. Böhlen, and Carl-Christian Kanne. *Smile: Enabling Easy and Fast Development of Domain-Specific Scheduling Protocols.* In Proceedings of the 28th British National Conference on Databases (**BNCOD**), 128-131, 2011

[31] **Boris Glavic** and Gustavo Alonso. *The Perm Provenance Management System in Action.* In Proceedings of the 35th ACM SIGMOD International Conference on Management of Data (**SIGMOD**) (Demonstration Track), 1055-1058, 2009

[32] **Boris Glavic** and Gustavo Alonso. *Perm: Processing Provenance and Data on the same Data Model through Query Rewriting.* In Proceedings of the 25th IEEE International Conference on Data Engineering (**ICDE**), 174-185, 2009

[33] **Boris Glavic** and Gustavo Alonso. *Provenance for Nested Subqueries.* In Proceedings of the 12th International Conference on Extending Database Technology (**EDBT**), 982-993, 2009

[34] **Boris Glavic** and Klaus R. Dittrich. *Data Provenance: A Categorization of Existing Approaches.* In Proceedings of the 12th GI Conference on Database Systems for Business, Technology, and Web (**BTW**), 227-241, 2007

[35] **Boris Glavic**. *Subspace Sequence Clustering - Datamining zur Entscheidungsunterstützung in der Hydrologie.* In Proceedings of the 11th GI Conference on Database Systems for Business, Technology, and Web (**BTW**) (Student Track), 15-17, 2005.

Book Chapters

[1] Oliver Kennedy and **Boris Glavic**. *Analyzing uncertain tabular data.* In Éloi Bossé and Galina Rogova, editors, Information quality in information fusion and decision making, page 291-320. Springer, 2019.

[2] **Boris Glavic**, Gustavo Alonso, and Renée J. Miller. *Using SQL for Efficient Generation and Querying of Provenance Information.* In search of elegance in the theory and practice of computation: a Festschrift in honour of Peter Buneman, Lecture Notes in Computes Science, 291-320, 2013.

Workshops

[1] Shatha Algarni, **Boris Glavic**, Seokki Lee, Adriane Chapman. *Solving Why Not Questions for Aggregate Constraints through Query Repair.* TaPP (**TaPP**), pp. 592-596, 2024.

[2] Seokki Lee, **Boris Glavic**, Adriane Chapman, Bertram Ludäscher. *Hybrid Query and Instance Explanations and Repairs.* Companion Proceedings of the ACM Web Conference, Proceedings of the 15th International Workshop on the Theory and Practice of Provenance (**TaPP**), 2023.

[3] Nachiket Deo, **Boris Glavic**, Oliver Kennedy. *Runtime Provenance Refinement for Notebooks.* Proceedings of the 14th International Workshop on the Theory and Practice of Provenance (**TaPP**), 2022.

[4] Qitian Zeng, Kyle Hale, **Boris Glavic**. *Playing Fetch with CAT - Composing Cache Partitioning and Prefetching for Task-based Query Processing.* International Workshop on Data Management on New Hardware (**DaMoN**), 2021.

[5] Ralf Diestelkämper, **Boris Glavic**, Melanie Herschel, and Seokki Lee. *Query-based why-not explanations for nested data.* In Proceedings of the 11th USENIX Workshop on the Theory and Practice of Provenance, 2019.

[6] Bahareh S. Arab, **Boris Glavic**. *Answering Historical What-if Queries with Provenance, Reenactment, and Symbolic Execution* In Proceedings of the 8th USENIX Workshop on the Theory and Practice of Provenance (**TaPP**), 2017

[7] Seokki Lee, Xing Niu, Betram Ludäscher, **Boris Glavic**. *Integrating Approximate Summarization with Provenance Capture.* In Proceedings of the 8th USENIX Workshop on the Theory and Practice of Provenance (**TaPP**), 2017

[8] Juliana Freire, **Boris Glavic**, Oliver Kennedy, and Heiko Müller. *The Exception that Improves the Rule* In SIGMOD Workshop on Human-In-the-Loop Data Analytics (**HILDA**), 2016

[9] Xing Niu, Bahareh Arab, Dieter Gawlick, Zhen Hua Liu, Vasudha Krishnaswamy, Oliver Kennedy, and **Boris Glavic**, *Provenance-aware Versioned Dataworkspaces* In Proceedings of the 8th USENIX Workshop on the Theory and Practice of Provenance (**TaPP**), 2016

[10] Xing Niu, Raghav Kapoor, and **Boris Glavic**. *Heuristic and Cost-based Optimization for Provenance Computation.* In Proceedings of the 7th USENIX Workshop on the Theory and Practice of Provenance (**TaPP**) (Poster), 2015.

[11] Xing Niu, Raghav Kapoor, Dieter Gawlick, Zhen Hua Liu, Vasudha Krishnaswamy, Venkatesh Radhakrishnan, and **Boris Glavic**. *Interoperability for provenance-aware databases using prov and json.* In Proceedings of the 7th USENIX Workshop on the Theory and Practice of Provenance (**TaPP**), 2015.

[12] **Boris Glavic**, Tanu Malik, and Quan Pham. *Making database applications shareable.* In Proceedings of the 7th USENIX Workshop on the Theory and Practice of Provenance (**TaPP**) (Poster), 2015.

[13] **Boris Glavic**, Sven Köhler, Sean Riddle, and Bertram Ludäscher. *Towards constraint-based explanations for answers and non-answers.* In Proceedings of the 7th USENIX Workshop on the Theory and Practice of Provenance (**TaPP**), 2015.

[14] Bahareh Arab, Dieter Gawlick, Venkatesh Radhakrishnan, Hao Guo, **Boris Glavic**. *A Generic Provenance Middleware for Queries, Updates, and Transactions.* In Proceedings of the 6th USENIX Workshop on the Theory and Practice of Provenance (**TaPP**), 2014.

[15] **Boris Glavic**. Invited talk. *Big Data Provenance: Challenges and Potential Implications for Benchmarking.* In Proceedings of the 2nd Workshop on Big Data Benchmarking (**WBDB**), 2013.

[16] **Boris Glavic**. *Provenance for Data Mining.* In Proceedings of the 5th USENIX Workshop on the Theory and Practice of Provenance (**TaPP**), 2013.

[17] **Boris Glavic** and Renée J. Miller. *Reexamining Some Holy Grails of Data Provenance.* In Proceedings of the 3rd USENIX Workshop on the Theory and Practice of Provenance (**TaPP**), 2011.

[18] **Boris Glavic**, Kyumars Sheykh Esmaili, Peter M. Fischer, and Nesime Tatbul. *The Case for Fine-Grained Stream Provenance.* In Proceedings of the 1st Workshop on Data Streams and Event Processing (**DSEP**) collocated with BTW, 58-61, 2011

[19] Ira Assent, Ralph Krieger, **Boris Glavic**, and Thomas Seidl. *Spatial Multidimensional Sequence Clustering.* In Proceedings of the 1st International Workshop on Spatial and Spatio-temporal Data Mining (**SSTDM**) collocated with ICDM, 343-348, 2006.

[20] **Boris Glavic** and Klaus Dittrich. *Sesam: Ensuring Privacy for an Interdisciplinary Longitudinal Study.* In Proceedings of the 1st Workshop Elektronische Datentreuhänderschaft - Anwendungen, Verfahren, Grundlagen collocated with GI Jahrestagung, 736-743, 2006.

Keynote Talks

2023   *Quid pro quo - Provenance for Visualization and Visualization for Provenance*
1st IEEE Workshop on Visualization and Provenance Across Domains @ IEEE VIZ

2023   *Provenance, Relevance-based Data Management, and the Value of Data*
ProvenanceWeek 2023 @ WWW

2023   *How-to-PhD - A Dummies Guide towards Research*
EDBT Ph.D. workshop @ EDBT

Invited Talks

2024   *Lightweight Models for Uncertain Query Processing and Machine Learning*
University of California, Berkeley, Skylab seminar

Relevance-based Data Management and The Value of Data
SUNY Buffalo

Lessons Learned from Building Systems for Provenance and Explanations
Dagstuhl Seminar 24032 Representation, Provenance, and Explanations in Database Theory and Logic

2023   *Approximating Certain and Possible Answers Beyond Sets*
Simons Institute Workshop on Logic and Algebra for Query Evaluation

Uncertain Data Management for Responsible Data Management
DATA Lab Seminar, **Northeastern University**

2022   *Interpretable Data-Based Explanations for Fairness Debugging*
DATA Lab Seminar, **Northeastern University**

Capturing Factorized Provenance
Factorized Databases Workshop, Zurich

Provenance-based Data Skipping
FONDA project lecture series **Humbolt University at Berlin, Germany**

Efficient Answering of Historical What-if Queries
Research School "Data Science and Engineering", **Hasso Plattner Institute (HPI), Germany**

Uncertainty-Annotated Databases - A Lightweight and Efficient Model for Uncertain Data
Seminar, **University of Massachusetts, Amherst**

2021   *Efficient Uncertainty Tracking for Complex Queries with Attribute-level Bounds*
Seminar, **Hasso Plattner Institute (HPI)**

2019   *Relevance-Based Data Management*
Seminar, **Northeastern University**

2017   *Reenacting Transactions for Provenance, Debugging, and World Peace*
DB Seminar, **Brown University**

2016   *HRDBMS: How to scale like an elephant and perform like a fruit*
Midwest Big Data Opportunities and Challenges (MBDOC) Workshop, sponsored by NSF Midwest Big Data Hub
Towards Rigorous Evaluation of Data Integration Systems - It's All About the Tools
QDB workshop (co-located with VLDB 2016)

Reenacting Transactions to Compute their Provenance

UBDB Seminar, **SUNY Buffalo**

Answering Why- and Why-not Provenance Questions Efficiently and Comprehensibly.
CMD Research Colloquium, **DePaul University**

Reenacting Transactions to Compute their Provenance.
The Yahoo!-DAIS Seminar, **University of Illinois at Urbana Champaign**

2015   *Reenacting Transactional Histories to Compute Their Provenance.*
DIS Seminar Series, **Free University of Bozen**, Italy

Big Data Benchmarking needs Big Metadata Generation.
Workshop on Big Data Benchmarking (**WBDB**)

Reenacting Transactions to Compute their Provenance.
Database Seminar @ NCSA, **University of Illinois at Urbana Champaign**

Vagabond: Automatic Generation and Ranking of Explanations for Data Exchange Errors.
CMD Research Colloquium, **DePaul University**

2014   *Vagabond: Automatic Generation and Ranking of Explanations for Data Exchange Errors.*
Workshop on Best Practices in Empirical Evaluation of Big Data Integration and Cleaning

Computing Provenance for Database Updates and Transactions.
CMD Research Colloquium, **DePaul University**

Computing Provenance for Database Updates and Transactions.
IIT Data Science Seminar

2012   *Big Data Provenance: Challenges and Implications for Benchmarking.*
Workshop on Big Data Benchmarking (**WBDB**)

2010   *Integrating Data Provenance Support in Database Systems.*
DIS Seminar Series, **Free University of Bozen**, Italy

Technical Reports

[1] Su Feng, Aaron Huber, **Boris Glavic**, Oliver Kennedy. *Uncertainty annotated databases - a lightweight approach for dealing with uncertainty.* Technical report IIT/CS-DB-2018-01, Illinois Institute of Technology, 2018.

[2] Seokki Lee, Sven Köhler, Bertram Ludäscher, **Boris Glavic**. *Efficiently computing provenance graphs for queries with negation.* Technical report IIT/CS-DB-2016-03, Illinois Institute of Technology, 2016.

[3] Xing Niu **Boris Glavic**. *Optimizing provenance capture and queries - algebraic transformations and cost-based optimization.* Technical report IIT/CS-DB-2016-02, Illinois Institute of Technology, 2016.

[4] Bahareh Arab, Dieter Gawlick, Vasudha Krishnaswamy, Venkatesh Radhakrishnan, **Boris Glavic**. *Reenactment for read-committed snapshot isolation (long version).* Technical report, Illinois Institute of Technology, 2016.

[5] Arindam Nandi, Ying Yang, Oliver Kennedy, **Boris Glavic**, Ronny Fehling, Zhen Hua Liu, and Dieter Gawlick. *Mimir: Bringing CTables into Practice*, Technical report, CoRR, 2016.

[6] Bahareh Arab, Dieter Gawlick, Vasudha Krishnaswamy, Venkatesh Radhakrishnan, and **Boris Glavic**. *Formal Foundations of Reenactment and Transaction Provenance*, Technical report, IIT, 2016.

[7] Seokki Lee, Yuchen Tang, Sven Köhler, Bertram Ludäscher, and **Boris Glavic**. *An Efficient Implementation Of Game Provenance In DBMS*, Technical Report, IIT, 2015.

[8] Patricia C. Arocena, **Boris Glavic**, Gianni Mecca, Renée Miller, Paolo Papotti, and Donatello Santoro. *Error Generation for Evaluating Data Cleaning Algorithms.* Technical report TR-01-2015, Universitá della Basilicata, 2015.

[9] Patricia C. Arocena, **Boris Glavic**, Radu Ciucanu, and Renée J. Miller. *The iBench Integration Metadata Generator.* Technical report, University of Toronto, 2015.

[10] Seokki Lee, Zhen Wang, **Boris Glavic**, and Renée J. Miller. *Automatic Generation and Ranking of Explanations for Mapping Errors.* Technical report, IIT, 2015.

[11] Quan Pham, Tanu Malik, **Boris Glavic**, Ian Foster. *LDV: Light-weight Database Virtualization.* Technical Report, IIT, 2014.

[12] Bahareh Arab, Dieter Gawlick, Vasudha Krishnaswamy, Venkatesh Radhakrishnan, **Boris Glavic**. *Reenacting Transactions to Compute their Provenance.* Technical Report, IIT, 2014.

[13] **Boris Glavic**. *A Primer on Database Provenance.* Technical Report, IIT, 2014.

[14] Patricia C. Arocena, Mariana D'Angelo, **Boris Glavic**, Renée J. Miller. *iBench First Cut.* Technical Report, University of Toronto, 2013.

[15] Javed Siddique, **Boris Glavic**, Renée J. Miller. *Provenance Management for Frequent Itemsets.* Technical report, University of Toronto, 2013.

[16] **Boris Glavic**, Kyumars Sheykh Esmaili, Peter M. Fischer, Nesime Tatbul. *Ariadne: Managing Fine-Grained Provenance on Data Streams.* Technical report, ETH Zurich, 2012.

[17] **Boris Glavic**. *Formal Foundation of Contribution Semantics and Provenance Computation through Query Rewrite in TRAMP.* Technical report, University of Zurich, 2010.

[18] Christian Tilgner, **Boris Glavic**, Michael H. Böhlen, and Carl-Christian Kanne. *Correctness Proof of the Declarative SS2PL Protocol Implementation.* Technical report, University of Zurich, 2010.

Extended Abstracts and Posters

[1] Xing Niu, and **Boris Glavic**, *Debugging Transactions and Tracking their Provenance with Reenactment*, 6th Greater Chicago Area System Research Workshop (**GCASR**) - Student Poster, 2017

[2] Seokki Lee, Sven Köhler, Bertram Ludäscher, and **Boris Glavic**, *A SQL-Middleware Unifying Why and Why-Not Provenance for First-Order Queries*, 6th Greater Chicago Area System Research Workshop (**GCASR**) - Student Poster, 2017

[3] Xing Niu, Raghav Kapoor and **Boris Glavic**, *Heuristic and Cost-based Optimization for Provenance Computation*, 5th Greater Chicago Area System Research Workshop (**GCASR**) - Student Poster, 2016

[4] Seokki Lee, Sven Köhler, Bertram Ludäscher, and **Boris Glavic**, *Implementing Unified Why- and Why-Not Provenance Through Games*, In Proceedings of the 8th USENIX Workshop on the Theory and Practice of Provenance (**TaPP**) (Poster), 2016

[5] Jason Arnold, **Boris Glavic**, Ioan Raicu. *HRDBMS: A NewSQL Database for Analytics*, In Proceedings of the IEEE International Conference on Cluster Computing (**Cluster**) (Poster), 2015

[6] Richard Whaling. *LDV – Lightweight Database Virtualization for Scientific Reproducibility.* 4th Greater Chicago Area System Research Workshop (**GCASR**) - Student Poster, 2015

[7] Seokki Lee. *Automatic Generation and Ranking of Explanations for Mapping Errors.* 4th Greater Chicago Area System Research Workshop (**GCASR**) - Student Poster, 2015

[8] Xing Niu, Raghav Kapoor. *Heuristic and Cost-based Optimization for Provenance Computation.* 4th Greater Chicago Area System Research Workshop (**GCASR**) - Student Poster, 2015

[9] Bahareh Arab, Hao Guo. *Provenance for Database Transactions.* 3nd Greater Chicago Area System Research Workshop (**GCASR**) - Student Poster, 2014

[10] **Boris Glavic**. *The (Potential) Perks of Integrating Provenance Support into Database Engines*, 2nd Greater Chicago Area System Research Workshop (**GCASR**), 2013

			Funding

2025-2026	$16,000 ($16,000)	PI	**NSF** REU Supplement for *III: Medium: Collaborative Research: U4U - Taming Uncertainty with Uncertainty-Annotated Databases*
2022-2023	$34,000	PI	Unrestricted Gift from the **Oracle** Cooperation (cloud credits) *Exploiting Provenance to Enhance Data Management - Cloud credits (2nd extension)*
2022-2023	$98,077	PI	Unrestricted Gift from the **Oracle** Cooperation *Self-tuning Database Operations by Assessing the Importance of Data (2nd extension)*
2021-2022	$16,000 ($16,000)	PI (collaborative)	**NSF** REU Supplement for *III: Medium: Collaborative Research: U4U - Taming Uncertainty with Uncertainty-Annotated Databases*
2021-2025	$1,199,874 ($357,000)	PI (collaborative)	**NSF** *III: Medium: Collaborative Research: From Open Data to Open Data Curation*
2021-2022	$25,000	PI	Unrestricted Gift from the **Oracle** Cooperation (cloud credits) *Exploiting Provenance to Enhance Data Management - Cloud credits (extension)*
2021-2022	$98,077	PI	Unrestricted Gift from the **Oracle** Cooperation *Self-tuning Database Operations by Assessing the Importance of Data (extension)*
2020-2021	$25,000	PI	Unrestricted Gift from the **Oracle** Cooperation (cloud credits) *Exploiting Provenance to Enhance Data Management - Cloud credits*
2020-2025	$999,492 ($466,569)	Lead PI (collaborative)	**NSF** *III: Medium: Collaborative Research: U4U - Taming Uncertainty with Uncertainty-Annotated Databases*
2019-2020	$97,507	PI	Unrestricted Gift from the **Oracle** Cooperation *Self-tuning Database Operations by Assessing the Importance of Data*
2018-2019	$95,248	PI	Unrestricted Gift from the **Oracle** Cooperation *Exploiting Provenance to Enhance Data Management*
2017-2018	$24,000 ($8,000)	PI at IIT (collaborative)	**NSF** REU Supplement for *CIF21 DIBBs: EI: Vizier, Streamlined Data Curation*
2017-2020	$2,725,699 ($404,979)	PI at IIT	**NSF** *CIF21 DIBBs: EI: Vizier, Streamlined Data Curation*
2016-2017	$90,871	PI	Unrestricted Gift from the **Oracle** Cooperation *Provenance for Big Data*
2014-2015	$95,829	PI	Unrestricted Gift from the **Oracle** Cooperation *Provenance using temporal databases (extension)*
2013-2014	$85,000	PI	Unrestricted Gift from the **Oracle** Cooperation *Provenance using temporal databases*
2013-2013	$25,000	PI	**IIT** Educational and Research Initiative Fund (ERIF) Grant *An interdisciplinary approach for assessing treatment motivation among patients undergoing antiretroviral therapy: Integrating multidimensional scaling with data provenance techniques*

Program Chair

2020 ProvenanceWeek 2020 / 2021 (Senior PC chair) (**ProvenanceWeek**)
2016 International Provenance and Annotation Workshop (**IPAW**)

Associate Editor

2023- The VLDB Journal (**VLDBJ**)
2023-2024 Proceedings of the VLDB Endowment (**PVLDB**)
2022-2023 Proceedings of the VLDB Endowment (**PVLDB**)
2021-2022 Proceedings of the VLDB Endowment (**PVLDB**)
2018- IEEE Transactions on Knowledge and Data Engineering (**TKDE**)
2020- Distributed and Parallel Databases (**DAPD**)

Guest Editor

2016 Distributed and Parallel Databases (**DAPD**) - Special issue on Large-Scale Data Curation and Metadata Management

Panel Organization

2021 VLDB Round Table - Data Governance and Provenance (co-organized with Sudeepa Roy)
2020 VLDB Round Table - Making Uncertain Data Management Practical (co-organized with Oliver Kennedy)

Committees

2023 VLDB 2023 Best Paper Award Committee

Program Committee Member

2026 **Area Editor** ACM SIGMOD International Conference on Management of Data (**SIGMOD**)
2025 **Area Editor** ACM SIGMOD International Conference on Management of Data (**SIGMOD**)
Proceedings of the VLDB Endowment (**PVLDB**)
ProvenanceWeek 2025 (**ProvenanceWeek**)
2024 ACM SIGMOD International Conference on Management of Data (**SIGMOD**)
International Conference on Extending Database Technology (**EDBT**)
Proceedings of the VLDB Endowment reproducibility (**PVLDB reproducibility**)
ACM International Conference on Information and Knowledge Management (**CIKM**) - [**senior PC member**]
2023 ACM International Conference on Information and Knowledge Management (**CIKM**) - [**senior PC member**]
ACM SIGMOD International Conference on Management of Data, demonstrations (**SIGMOD**)
ACM SIGMOD International Conference on Management of Data (**SIGMOD**)
International Conference on Extending Database Technology (**EDBT**)
IEEE International Conference on Data Engineering (**ICDE**)
PVLDB reproducibility (**PVLDB**)

ACM International Conference on Information and Knowledge Management, short papers (**CIKM**) - [**senior PC member**]

2022 ACM SIGMOD International Conference on Management of Data (**SIGMOD**)
IEEE International Conference on Data Engineering (**ICDE**)
IEEE International Conference on Data Engineering, demonstrations (**ICDE**)
International Conference on Extending Database Technology (**EDBT**)
ACM SIGIR Conference on Research and Development in Information Retrieval (**SIGIR**)
USENIX Workshop on the Theory and Practice of Provenance (**TaPP**)
International Conference on Distributed Event-Based Systems (**DEBS**)

2021 International Conference on Very Large Databases (**VLDB**)
ACM SIGMOD International Conference on Management of Data (**SIGMOD**)
IEEE International Conference on Data Engineering, demonstrations (**ICDE**)
International Conference on Very Large Databases, PhD Workshop (**VLDB**)
International Conference on Extending Database Technology (**EDBT**)
ACM SIGIR Conference on Research and Development in Information Retrieval (**SIGIR**)

2020 ACM SIGMOD International Conference on Management of Data (**SIGMOD**)
IEEE International Conference on Data Engineering (**ICDE**)
International Conference on Very Large Databases, demonstrations (**VLDB**)
ACM International Conference on Information and Knowledge Management (**CIKM**)
International Conference on Distributed Event-Based Systems (**DEBS**)
International Conference on Database Systems for Advanced Applications (**DASFAA**)
International Conference on Scientific and Statistical Database Management (**SSDBM**)

2019 ACM SIGMOD International Conference on Management of Data, demonstrations (**SIGMOD**)
ACM International Conference on Information and Knowledge Management (**CIKM**)
International Conference on Extending Database Technology (**EDBT**)
ACM SIGMOD International Conference on Management of Data (**SIGMOD**)
International Conference on Very Large Databases (**VLDB**)
International Conference on Very Large Databases, demonstrations (**VLDB**)
IEEE International Conference on Data Engineering (**ICDE**)
International Conference on Scientific & Statistical Database Management (**SSDBM**)
USENIX Workshop on the Theory and Practice of Provenance (**TaPP**)
International Conference on Database Systems for Advanced Applications (**DASFAA**)

2018 ACM International Conference on Information and Knowledge Management (**CIKM**)
International Workshop on Bidirectional Transformations (**Bx**)
USENIX Workshop on the Theory and Practice of Provenance (**TaPP**)
IEEE International Conference on Data Engineering (**ICDE**)
International Conference on Extending Database Technology (**EDBT**)
International Conference on Very Large Databases, demonstrations (**VLDB**)
International Workshop on Big Data Stream Processing (**BigStream**)

2017 International Workshop on Data Management on Virtualized Active Systems (**Active**)
ACM International Conference on Information and Knowledge Management (**CIKM**)
International Conference on Data Management Technologies and Applications (**DATA**)
IEEE International Conference on Data Engineering (**ICDE**), Ph.D. symposium
ACM SIGMOD International Conference on Management of Data (**SIGMOD**)
International Conference on Very Large Databases (**VLDB**)

2016 IEEE International Conference on Data Science and Advanced Analytics (**DSSA**)
IEEE International Conference on Data Engineering (**ICDE**), demonstrations
ACM SIGMOD International Conference on Management of Data (**SIGMOD**)

2015 Workshop on Big Data Benchmarking Canada (**WBDB Canada**)
International Workshop on the Web and Databases (**WebDB**)
USENIX Workshop on the Theory and Practice of Provenance (**TaPP**)
International Conference on Data Management Technologies and Applications (**DATA**)
ACM SIGMOD International Conference on Management of Data (**SIGMOD**)

IEEE International Conference on Data Engineering (**ICDE**)
International Conference on Extending Database Technology (**EDBT**)
ACM SIGMOD International Conference on Management of Data, demonstrations (**SIGMOD**)
2014 ACM SIGMOD International Conference on Management of Data, demonstrations (**SIGMOD**)
USENIX Workshop on the Theory and Practice of Provenance (**TaPP**)
International Conference on Data Management Technologies and Applications (**DATA**)
2013 Workshop on Big Data Benchmarking China (**WBDB China**)
Workshop on Big Data Benchmarking US (**WBDB.us**)
USENIX Workshop on the Theory and Practice of Provenance (**TaPP**)
International Conference on Scientific and Statistical Database Management, demonstrations (**SSDBM**)
2012 Alberto Mendelzon Workshop on Foundations of Data Management (**AMW**)
2011 ACM SIGMOD International Conference on Management of Data (**SIGMOD**)

Journal Reviews

2025 ACM Journal of Data and Information Quality (**JDIQ**)
The VLDB Journal
2024 The VLDB Journal
ACM Transactions on Software Engineering and Methodology (**TOSEM**)
2023 ACM Transactions on Database Systems (**TODS**)
ACM Journal of Data and Information Quality (**JDIQ**)
The VLDB Journal
ACM Transactions on Software Engineering and Methodology (**TOSEM**)
2022 The VLDB Journal
ACM Computing Surveys (**CSUR**)
2021 Elsevier Information Systems
ACM Transactions on Database Systems (**TODS**)
The VLDB Journal
ACM Computing Surveys (**CSUR**)
2020 ACM Computing Surveys (**CSUR**)
ACM Transactions on Database Systems (**TODS**)
The VLDB Journal
ACM Journal of Data and Information Quality (**JDIQ**)
2019 ACM Transactions on Database Systems (**TODS**)
IEEE Computer Graphics and Applications (**CGA**)
Elsevier Information Systems
2018 Elsevier Information Systems
ACM Transactions on Database Systems (**TODS**)
IEEE Transactions on Knowledge and Data Engineering (**TKDE**)
2017 International Journal on Semantic Web and Information Systems (**IJSWIS**)
ScienceDirect Science of Computer Programming Journal (**SCICO**)
Springer Information Systems Frontiers (**ISF**)
The VLDB Journal
2016 IEEE Transactions on Software Engineering (**TSE**)
ACM Transactions on Internet Technology (**TOIT**)
The VLDB Journal
IEEE Transactions on Knowledge and Data Engineering (**TKDE**)
2015 Elsevier Journal of Biomedical Informatics (**JBI**)
Springer Artificial Intelligence Review (**ARTIF INTELL REV**)
Elsevier Information Systems
The VLDB Journal

2014 The VLDB Journal
      ACM Journal of Data and Information Quality (**JDIQ**)
      IEEE Computer (**Computer**)
2013 IEEE Transactions on Knowledge and Data Engineering (**TKDE**)
      ACM Transactions on Internet Technology (**TOIT**)
2012 The VLDB Journal
      ACM Transactions on Database Systems (**TODS**)
2011 ACM Transactions on Database Systems (**TODS**)
      IEEE Transactions on Knowledge and Data Engineering (**TKDE**)

NSF Panels

2023 Information and Intelligent Systems (CISE-IIS)
2023 Information and Intelligent Systems (CISE-IIS)
2022 Information and Intelligent Systems (CISE-IIS)
2021 Information and Intelligent Systems (CISE-IIS)
2020 Information and Intelligent Systems (CISE-IIS)
2019 Information and Intelligent Systems (CISE-IIS)
2013 Information and Intelligent Systems (CISE-IIS)

Other Proposal Reviews

2025 Israel Science Foundation (ISF) - Proposal Review
2024 Israel Science Foundation (ISF) - Proposal Review
2023 U.S-Israel Binational Science Foundation (BSF) - Proposal Review
2023 Israel Science Foundation (ISF) - Proposal Review
2022 Canadian Natural Sciences and Engineering Research Council (NSERC) - Proposal Review
2020 Canadian Natural Sciences and Engineering Research Council (NSERC) - Proposal Review
2019 European Research Council (ERC) - Proposal Review
2018 European Research Council (ERC) - Proposal Review
2017 Canadian Natural Sciences and Engineering Research Council (NSERC) - Proposal Review
2016 Israel Science Foundation (ISF) - Proposal Review
2015 Fields Institute, Toronto - Event Proposal Review
2014 Marsden Fund (The Royal Society of New Zealand) - Proposal Review

External Reviewer

2019 International Conference on Database Theory (**ICDT**)
      ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages (**POPL**)
2015 ACM/IEEE Symposium on Logic In Computer Science (**LICS**)
2010 IEEE International Conference on Data Engineering (**ICDE**)
2009 GI Conference on Database Systems for Business, Technology, and Web (**BTW**)
2008 IEEE International Conference on Data Engineering (**ICDE**)
2007 International Conference on Very Large Databases (**VLDB**)
      ACM SIGMOD International Conference on Management of Data (**SIGMOD**)
      International Conference on Objects, Models, Components, Patterns (**TOOLS**)
      International Conference on Business Information Systems (**BIS**)
2006 International Conference on Very Large Databases (**VLDB**)

Teaching and Advisor Experience

Graduated Ph.D. Students

2023	S. Feng	Efficient Management of Uncertain Data
2021	X. Niu	Integrating Provenance Management and Query Optimization
2020	S. Lee	Why- and Why-not Provenance for Queries with Negation
2019	B. Arab	Provenance for Updates and Transactions

Current Ph.D. Students

C. Li	Explanations beyond provenance
Z. Liu	Cost estimation for relevance-based data management
P. Li	Maintaining provenance sketches under updates
D. Lee	Translating procedural code into batch-oriented SQL

Courses Developed and Taught

2024	CS594 - Provenance & Explanations
	UIC
	CS480 - Database Systems
	UIC
2023	CS525 - Advanced Database Organization
	IIT
2022	CS520 - Data Integration, Warehousing, and Provenance
	IIT
	CS425 - Database Organization
	IIT
2021	CS595 - Topics in Big Data Analytics
	IIT
	CS331 - Datastructures and Algorithms
	IIT
2020	CS520 - Data Integration, Warehousing, and Provenance
	IIT
	CS525 - Advanced Database Organization
	IIT
2019	CS116 - Object-oriented Programming II
	IIT
2018	CS520 - Data Integration, Warehousing, and Provenance
	IIT
	CS425 - Database Organization
	IIT
2017	CS425 - Database Organization
	IIT
	CS525 - Advanced Database Organization
	IIT
2016	CS425 - Database Organization
	IIT

CS520 - Data Integration, Warehousing, and Provenance
IIT

2015 CS525 - Advanced Database Organization
IIT
CS520 - Data Integration, Warehousing, and Provenance
IIT

2014 CS425 - Database Organization
IIT
CS525 - Advanced Database Organization
IIT
CS695 - Doctoral Seminar
IIT

2013 CS425 - Database Organization
IIT
CS525 - Advanced Database Organization
IIT

2012 CS595 - Hot topics in database systems: Data Provenance
IIT

2009 327 - Informatik für Ökonomen II (Teil 1) - Einführung in Datenbanken (Computer science for economists II (part 1) - introduction to databases)
University of Zurich

Ph.D. Student Thesis Committee Member

2023	UBC	O. Al Omeir
	IIT CS	H. Wang
2021	IIT CS	U. Das
	University of Stuttgart, Germany	R. Diestelkämper
	IIT CS	B. Bekiroglu
2020	Tel Aviv University, Israel	A. Gilad
	University of Tübingen, Germany	T. Müller
	DePaul University	J. Wagner
2019	University of Toronto CS	J. Du
2017	IIT CS	X. Zhang
	IIT CS	T. Jung
	University of Basilicata CS, Italy	E. Veltri
	IIT CS	J. Kroculic
2016	IIT CS	I. Sadooghi
	IIT CS	M. Sharma
	IIT Psychology	K. Niel
	IIT ECE	Q. Tong
2015	IIT CS	X. Zhang
	IIT CS	T. Li
	IIT CS	U. Das
2013	IIT CS	R. Summerscales

Supervised Master and Diploma Theses

2023	M. Sarkar	Effect of Pre-processing Data on Fairness and Fairness Debugging Using Gopher
2022	J. Lee	Enhancing Explanation Generation in the CaJaDE system using Interactive User Feedback

2018	S. Samson	Optimizing Communication in Distributed Query Execution
2015	A. Cornudella	Computing Candidate Keys of Relational Operators for Optimizing Rewrite-Based Provenance Computations
2014	Z. Wang	Efficient Ranking of Explanations for Data Exchange Errors in Vagabond
2010	B. Steiger	Data lineage/provenance in X-query (co-supervised with P.M. Fischer and K.S. Esmaili)
	A. Vagliardo	MidPerm - Realisierung einer Provenance SQL-Erweiterung als Middleware Lösung
2009	O. Wirz	Entwurf und Implementierung eines SQL-DDL-Präprozessors zur Unterstützung von Datenbankentwurfsmustern
2008	S. Gerster	Entwurf und Umsetzung eines Prototyps zur effizienten Speicherung von Hoch-Volumen Prozessdaten
	S. Blatti	Entwurf und Implementierung eines Provenance Browsers für die Visualisierung von Data Provenance
	S. Nedkoff	DBDoc Entwurf und Implementierung einer Anwendung zur partiellen Automation des Dokumentations-Prozesses für Datenbanken
2007	P. Hochstrasser	Entwurf und Implementierung einer Anwendung zur computergestützten Durchführung von klinischen Interviews
	C. Humard	Entwurf und Implementierung einer modular erweiterbaren Anwendung für Eingabe und Import von heterogenen Daten
	T. Schlaginhaufen	Design and Implementation of a Database Client Application for Inserting, Modifying, Presentation and Export of Bitemporal Personal Data

Supervised Bachelor Theses

2008	M. Keller	Visualisierung von datenbankunterstützten Prozessen
	S. Schurgast	Export von Datenbankinhalten in Datenformate von Statistikprogrammen

Supervised Student Projects

2024	E. Federmeyer	Uncertain Query Processing with Z-tables
2022	A. Veer	CS597 - Hierarchical Provenance Sketches for JSON data
	J. Calveres Garcia	CS597 - Enhancing Explanations in CaJaDe with Causal Reasoning
2021	F. Campbell	NSF REU - Historical What-if queries
	J. Lee	Lenses in Vizier, A demonstration of Cajade
2020	F. Campbell	Historical What-if queries
2019	C. Li	CS597 - Cape - a GUI frontend for explaining outliers
	F. Campbell	CS497 - Supporting time travel and audit logs in Postgres
2018	P. Rotti	CS497 - GProM - Creating uncertainty labelings from incomplete databases
	P. Rotti	Summer Research - Attribute level uncertainty labelings
	J. Sun	CS497 - GProM - Creating uncertainty labelings from probabilistic databases
	V. Polavarapu	CS597 - HRDBMS - Non-blocking queues
2017	T. Nguyen	HRDBMS - External Table types
	A. Alsabbagh	NSF REU - Vizier
	F. Morgan	CS597 - HRDBMS - Fast external load
	M. Kuzmyn	CS597 - HRDBMS - External table support
	F. Morgan	CS597 - HRDBMS - External table support

	N. Brahmankar	CS597 - HRDBMS - expression compilation
	K. Juneja	CS597 - HRDBMS - expression compilation
	J. Chen	CS597 - HRDBMS - expression compilation
	M. Mistry	CS597 - GProM - SparkSQL backend
2016	S. Kunali	CS597 - HRDBMS - External table support
	M. Kuzmyn	CS597 - An effective encoding of temporal data and its implementation in an ORM tool
	P. Li	CS597 - Developing a debugger GUI for transactional workloads
	C. Liu	CS597 - Extending reenactment with temporal aspects
	D. Patel	CS597 - User-defined domain specification for Datalog provenance
	L. Patil	CS597 - Adding support for JSON querying to HRDBMS
	A. Zhen	Summer Internship - Datageneration in iBench
	R. Alves Da Silva	BSMP Undergraduate Summer Intern - iBench
	D. Campos Faial	BSMP Undergraduate Summer Intern - GProM
	R. Leitao Oliveira	BSMP Undergraduate Summer Intern - HRDBMS
	P. Miranda Vandeveld	BSMP Undergraduate Summer Intern - HRDBMS
	J. Monte Carvalho	BSMP Undergraduate Summer Intern - HRDBMS
	K. Queiroz Giulianetti C Moura	BSMP Undergraduate Summer Intern - HRDBMS
	F. Soares Di Biase Moura	BSMP Undergraduate Summer Intern - iBench
2015	R. Sreenidhi	CS597 - Vagabond - Interactive confirmation of ranked explanations
	S. Dasadawala	CS597 - GProM - LDV JDBC instrumentation
	J. Zhu	CS597 - GProM - Enhancing Algebra Graph Support
	R. Kapoor	CS597 - GProM - JSON support and PROV import/export
	C. Abreu	BSMP Undergraduate Summer Intern - GProM
	R. Barsanulfo Fernandes	BSMP Undergraduate Summer Intern - iBench
	C. Borges Da Silva	BSMP Undergraduate Summer Intern - iBench
	W. Castro	BSMP Undergraduate Summer Intern - GProM
	C. Dos Santos Sousa	BSMP Undergraduate Summer Intern - Vagabond
	R. Franco	BSMP Undergraduate Summer Intern - iBench
	A. Hirai	BSMP Undergraduate Summer Intern - iBench
	A. Paiva	BSMP Undergraduate Summer Intern - Vagabond
	M. Rocha	BSMP Undergraduate Summer Intern - iBench
	J. Vieira	BSMP Undergraduate Summer Intern - Vagabond
2014	R. Kapoor	CS597 - GProM - Heuristic and Cost-based Optimization
	R. Li	CS597 - Vagabond GUI - Fixes and Support for Query-based Error Marking
	J. Zhu	CS597 - GProM - Add Support for Algebra Graphs to Provenance Rewriter
	M. Lucera	BSMP Undergraduate Summer Intern - A GUI for Transaction Provenance
	A. Iype	CS597 - Functional Dependency Compliant Data Generator
	H. Guo	CS597 - Implementing a Query-Rewrite Based Provenance Database Middleware - Provenance Support for Update and Transactions

2013	B. Dan	CS597 - Implementing a Query-Rewrite Based Provenance Database Middleware - Temporal Features
	H. Guo	CS597 - Implementing a Query-Rewrite Based Provenance Database Middleware - Provenance Support for Update and Transactions
	Z. Lin	CS597 - Implementing a Query-Rewrite Based Provenance Database Middleware - PI-CS Provenance
	Y. Ni	CS597 - Implementing a Query-Rewrite Based Provenance Database Middleware - Nested Subqueries
	P. Purandare	CS597 - Implementing a Query-Rewrite Based Provenance Database Middleware - PI-CS Provenance
	S. Xie	CS597 - Implementing a Query-Rewrite Based Provenance Database Middleware - Collections and internal model
	A. Iype	CS597 - Database-side Batch Handling of Explanations and Error Sets
	S. Xie	CS597 - MDS Provenance
	G. Shankar	CS597 - Implementation and Integration of a Provenance-aware Aggregation Operator
2011	V. Suen	Undergraduate Student Research Awards Summer Project (NSERC USRA): Automatic Generation of Explanations for Data Exchange Errors
	G. Pourshahid	Project for Research Topics in Data Management (CSC 2525): A User Interface for TRAMP
2010	K. Bühler	Performance Analysis of Provenance Computation using Query Rewrite Techniques on a Commercial Database System