Yuanyuan Tian

yuanyuantian@microsoft.com
https://humming80.github.io   Phone: (734) 709 8497

Current Position	Principal Scientist Manager, Graph Architect	Microsoft
	ACM Distinguished Member	

Expertise	Graph Analytics, ML-for-Systems, Big Data, Cloud Databases

Education

Sep 2003 – Aug 2008	**University of Michigan**	Ann Arbor, MI

Ph.D. in Computer Science & Engineering   Advisor: Jignesh M. Patel
Thesis: Querying Graph Databases
Designed and developed a graph querying system to support various sophisticated analytics on graphs including graph matching and graph summarizations.

Sep 2003 – Dec 2005	**University of Michigan**	Ann Arbor, MI

M.S. in Computer Science   GPA: 8.77 / 9.00

Sep 1999 – Jul 2003	**Peking University**	Beijing, China

B.S. in Computer Science & Technology   GPA: 91.52 / 100

Professional Experience

Aug 2021 – Present	**Microsoft**	Mountain View, CA

Principal Scientist Manager, Graph Architect
- Leading graph analytics, ML-for-Systems, and query optimization research areas in Gray Systems Lab (GSL)
- Architecting the graph analytics engine for the Fabric Graph team

Oct 2008 – Jun 2021	**IBM Research - Almaden**	San Jose, CA

Principal Research Staff Member
- Founded and led a project on supporting graph analytics inside relational databases, which is released as IBM Db2 Graph.
- Co-founded and technically co-led the Wildfire system for hybrid transactional and analytical processing (HTAP), which is released as the IBM Db2 Event Store.
- Led multiple SQL-for-Big-Data projects with techniques incorporated into the IBM BigInsights and IBM Db2 Big SQL products.
- Co-founded and contributed as a lead developer of a distributed machine learning system, which was open sourced as Apache SystemML (now SystemDS)

May 2004 – Aug 2008	**University of Michigan**	Ann Arbor, MI

Graduate Student Research Assistant, Advisor: Jignesh M. Patel
- Invented novel and efficient methods for querying and mining graphs and sequences.
- Designed and developed an efficient graph querying system that was applied to solve the core driving biological problems in the National Center for Integrative Biomedical Informatics (NCIBI).

Jun 2007 - Aug 2007	**Nokia Research Center**	Palo Alto, CA

Research Intern

1

- Applied graph mining methods to analyze call logs.
- Researched on methods for summarizing large social networks.

Jan 2005 - Apr 2005	**University of Michigan**	Ann Arbor, MI

Graduate Student Instructor
- Led the weekly 1-hour discussion section of EECS 485: Web Databases.
- Held office hours to help students with lecture material and projects.
- Designed and graded course projects.

May 2002 - Jun 2003	**Peking University**	Beijing, China

Undergraduate Student Research Assistant
- Worked on the web database design for a departmental online information system.

Jul 2002 - Mar 2003	**Bright Ocean Corporation**	Beijing, China

Part-time Intern
- Designed and developed data visualization tools used in China Mobile Corporation's Decision Support System for telecommunication network operation, administration and maintenance.

Selected Awards

External Awards

2023	**DaMoN 2023 Best Short Paper Award**	DaMoN 2023

R. Sen, Y. Tian: Microarchitectural Analysis of Graph BI Queries on RDBMS

2020	**ACM Distinguished Member**	ACM

For Outstanding Scientific Contributions to Computing

2019	**ACM SIGMOD Research Highlight Award**	SIGMOD 2019

B. Hentschel, P. J. Haas, Y. Tian: Online Model Management via Temporally Biased Sampling

2019	**VLDB 2019 Distinguished Reviewer Award**	VLDB 2019

2018	**EDBT 2018 Best Paper Award**	EDBT 2018

B. Hentschel, P. J. Haas, Y. Tian: Temporally-Biased Sampling for Online Model Management

IBM Awards

2020	**Outstanding Technical Achievement Award**	IBM Research - Almaden

For the contribution to the IBM Db2 Event Store product

2019	**Outstanding Technical Achievement Award**	IBM Research - Almaden

For the work in large-scale graph analytics and infrastructure

2019	**Research Division Award**	IBM Research - Almaden

For technical contributions to declarative machine/deep learning

2019	**Invention Achievement Award**	IBM Research - Almaden

2019	**IBM A-Level Accomplishment**	IBM Research - Almaden

For contribution to the IBM Db2 Event Store product

2018	**IBM A-Level Accomplishment**	IBM Research - Almaden

2018   **IBM A-Level Accomplishment**   IBM Research - Almaden
For impactful research on next generation large graph analytics and infrastructure

2018   **IBM A-Level Accomplishment**   IBM Research - Almaden
For contributions to large scale declarative machine/deep learning

2016   **Outstanding Technical Achievement Award**   IBM Research - Almaden
For technical contributions to efficient join algorithms for Big Data.

2015   **IBM A-Level Accomplishment**   IBM Research - Almaden
For impactful research on efficient join algorithms for Big Data.

2015   **IBM A-Level Accomplishment**   IBM Research - Almaden
For contributions to the Apache SystemML open source project

2013   **High Value Patent Application Award**   IBM Research - Almaden
For US and International patents on systems and methods for processing machine learning algorithms in a MapReduce environment

Graduate School Awards

2008   **Distinguished Achievement Award**   University of Michigan
Awarded to selected doctoral students with research excellence.
(The only recipient in CSE division.)

2007   **2nd Place, CSE Honor Competition**   University of Michigan
For the best research and presentations among CSE graduate students.

2007   **Rackham Predoctoral Fellowship**   University of Michigan
Awarded to selected doctoral students with outstanding dissertation research
(The only recipient in CSE division.)

2003   **Rackham Graduate Fellowship**   University of Michigan
Awarded to selected incoming graduate students with outstanding academic records

Publications

- ## Books and Book Chapter

Systems for Big Graph Analytics (Book)
D. Yan, **Y. Tian**, J. Cheng. *SpringerBriefs in Computer Science, Springer, 2017.*

Big Graph Analytics Platforms (Book)
D. Yan, Y. Bu, **Y. Tian**, A. Deshpande. *Foundations and Trends in Databases, Vol. 7: No. 1-2, pp 1-195, 2017.*

Interactive Graph Summarization (Book Chapter)
Y. Tian, J. M. Patel. *Link Mining: Models, Algorithms and Applications, Springer, 2010.*

- ## Encyclopedia Articles

Hybrid Systems Based on Traditional Database Extensions
Y. Tian. *Encyclopedia of Big Data Technologies, 2018.*

Wildfire: HTAP for Big Data
R. Barber, V. Raman, R. Sidle, **Y. Tian**, P. Tozun. *Encyclopedia of Big Data Technologies, 2018.*

- **Journal Articles**

A Roadmap to Graph Analytics
A. Bonifati, M. T. Özsu, **Y. Tian**, H. Voigt, W. Yu, and W. Zhang. *ACM SIGMOD Record 53 (4), 2025.*

MLOS in Action: Bridging the Gap Between Experimentation and Auto-Tuning in the Cloud
B. Kroth, S. Matusevych, R. Alotaibi, Y. Zhu, A. Gruenheid, **Y. Tian**. *The Proceedings of the VLDB Endowment (PVLDB), 2024.*

Optimizing Data Pipelines for Machine Learning in Feature Stores
R. Liu, K. Park, F. Psallidas, X. Zhu, J. Mo, R. Sen, M. Interlandi, K. Karanasos, **Y. Tian**, J. Camacho-Rodríguez. *The Proceedings of the VLDB Endowment (PVLDB), 2024.*

Exact PPS Sampling with Bounded Sample Size
B. Hentschel, P. J. Haas, **Y. Tian**. *Information Processing Letters, 2023.*

The World of Graph Databases from An Industry Perspective
Y. Tian. *ACM SIGMOD Record 51 (4), 2022.*

Db2 Event Store: A Purpose-Built IoT Database Engine
C. Garcia-Arellano, A. Storm, D. Kalmuk, H. Roumani, R. Barber, **Y. Tian**, R. Sidle, F. Ozcan, M. Spilchen, J. Tiefenbach, D. Zilio, L. Pham, K. Rakopoulos, A. Cheung, D. Pepper, I. Sayyid, G. Gershinsky, G. Lushi, H. Pirahesh. *The Proceedings of the VLDB Endowment (PVLDB), 2020.*

General Temporally-Biased Sampling Schemes for Online Model Management
B. Hentschel, P. Haas, **Y. Tian**. *ACM Transactions on Database Systems (TODS), 2019.* **(Invited as Best of EDBT 2018)**

Online Model Management via Temporally Biased Sampling
B. Hentschel, P. Haas, **Y. Tian**. *SIGMOD Record, 2019.* **(ACM SIGMOD 2019 Research Highlight Award)**

HERMIT in Action: Succinct Secondary Indexing Mechanism via Correlation Exploration
Y. Wu, J. Yu, **Y. Tian**, R. Sidle, R. Barber. *The Proceedings of the VLDB Endowment (PVLDB), 2019.*

Synergistic Graph and SQL Analytics Inside IBM Db2
Y. Tian, S. J. Tong, H. Pirahesh, W. Sun, E. L. Xu, W. Zhao. *The Proceedings of the VLDB Endowment (PVLDB), 2019.*

Building a Hybrid Warehouse: Efficient Joins Between Data Stored in HDFS and Enterprise Warehouse
Y. Tian, F. Ozcan, T. Zou, R. Goncalves, H. Pirahesh. *ACM Transactions on Database Systems (TODS), 2016.* **(Invited as Best of EDBT 2015)**

SystemML's Optimizer: Plan Generation for Large-Scale Machine Learning Programs
M. Boehm, D. Burdick, A. Evfimievski, B. Reinwald, F. R. Reiss, P. Sen, S. Tatikonda, **Y. Tian**. *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering, 37(3), 2014.*

Hybrid Parallelization Strategies for Large Scale Machine Learning in SystemML
M. Boehm, S. Tatikonda, B. Reinwald, P. Sen, **Y. Tian**, D. R. Burdick, S. Vaithyanathan. *The Proceedings of the VLDB Endowment (PVLDB), 7(7), 2014.*

From "Think Like a Vertex" to "Think Like a Graph"
Y. Tian, A. Balmin, S. A. Corsten, S. Tatikonda, J. McPherson. *The Proceedings of the VLDB Endowment (PVLDB), 7(3), 2013. **(Best of VLDB 2014)***

A Platform for eXtreme Analytics
A. Balmin, K. Beyer, V. Ercegovac, J. McPherson, F. Ozcan, H. Pirahesh, E. Shekita, Y. Sismanis, S. Tata, **Y. Tian***. IBM Journal of Research and Development, 57(3/4), 2013.*

CoHadoop: Flexible Data Placement and Its Exploitation in Hadoop
M. Eltabakh, **Y. Tian**, F. Ozcan, R. Gemulla, A. Krettek, J. McPherson. *The Proceedings of the VLDB Endowment (PVLDB), 2011.*

Michigan Molecular Interactions R2: From Interacting Proteins to Pathways
G. Tarcea, T. Weymouth, A. Ade, A. Bookvich, J. Gao, V. Mahavisno, Z. Wright, A. Chapman, M. Jayapandian, A. Ozgur, **Y. Tian**, J. Cavalcoli, B. Mirel, J. Patel, D. Radev, B. Athey, D. States and H. V. Jagadish. *Nucleic Acids Research 37 (Database issue):D642-6, 2009.*

Periscope/GQ: A Graph Querying Toolkit
Y. Tian, J. M. Patel, V. Nair, S. Martini, M. Kretzler. *The Proceedings of the VLDB Endowment (PVLDB), 1(2), 2008.*

SAGA: A Subgraph Matching Tool for Biological Graphs
Y. Tian, R. C. McEachin, C. Santos, D. J. States, J. M. Patel. *Bioinformatics Journal, 23(2): 232-239, 2007.*

Practical Methods for Constructing Suffix Trees
Y. Tian, S. Tata, R. A. Hankins, J. M. Patel. *Very Large Data Base Journal (VLDBJ), 14(3): 281-299, 2005.*

- **Peer-Reviewed Conference Papers**

Towards Query Optimizer as a Service (QOaaS) in a Unified LakeHouse Ecosystem: Can One QO Rule Them All?
R. Alotaibi, **Y. Tian**, S. Grafberger, J. Camacho-Rodr'iguez, N. Bruno, B. Kroth, S. Matusevych, A. Agrawal, M. Behera, A Gosalia, C. Galindo-Legaria, M. Joshi, M. Potocnik, B. Sezgin, X. Li, C. Curino. *The 2025 Conference on Innovative Data Systems Research (CIDR), 2025.*

The Future of Graph Analytics
A. Bonifati, T. Özsu, **Y. Tian**, H. Voigt, W. Yu. *The 44th ACM SIGMOD International Conference on Management of Data (SIGMOD), 2024.*

Sibyl: Forecasting Time-Evolving Query Workloads
H. Huang, T Siddiqui, R. Alotaibi, C. Curino, J. Leeka, A. Jindal, J. Zhao, J. Camacho-

Rodríguez, **Y. Tian**. *The 44th ACM SIGMOD International Conference on Management of Data (SIGMOD), 2024.*

GEqO: ML-Accelerated Semantic Equivalence Detection
B. Haynes, R. Alotaibi, A. Pavlenko, J. Leeka, A. Jindal, **Y. Tian**. *The 44th ACM SIGMOD International Conference on Management of Data (SIGMOD), 2024.*

Microarchitectural Analysis of Graph BI Queries on RDBMS
R. Sen, **Y. Tian**. *The 19th International Workshop on Data Management on New Hardware (DaMoN), 2023*. *(DaMoN 2023 Best Short Paper Award)*

Towards Building Autonomous Data Services on Azure
Y. Zhu, **Y. Tian**, J. Cahoon, S. Krishnan, A. Agarwal, R. Alotaibi, J. Camacho-Rodríguez, B. Chundatt, A. Chung, N. Dutta, A. Fogarty, A. Gruenheid, B. Haynes, M. Interlandi, M. Iyer, N. Jurgens, S. Khushalani, B. Kroth, M. Kumar, J. Leeka, S. Matusevych, M. Mittal, A. Mueller, K. Muthyala, H. Nagulapalli, Y. Park, H. Patel, A. Pavlenko, O. Poppe, S. Ravindran, K. Saur, R. Sen, S. Suh, A. Tarafdar, K. Waghray, D., C. Curino, R. Ramakrishnan. *The 43rd ACM SIGMOD International Conference on Management of Data (SIGMOD), Industry Paper, 2023.*

IBM Db2 Graph: Supporting Synergistic and Retrofittable Graph Queries Inside IBM Db2
Y. Tian, E. L. Xu, W. Zhao, H. Pirahesh, S. J. Tong, W. Sun, T. Kolanko, S. H. Apu, H. Peng. *The 40th ACM SIGMOD International Conference on Management of Data (SIGMOD), Industry Paper, 2020.*

Enabling Rich Queries over Heterogeneous Data from Diverse Sources in HealthCare
A. Quamar, J. Straube, **Y. Tian**. *The 2020 Conference on Innovative Data Systems Research (CIDR), 2020.*

WiSer: A Highly Available HTAP DBMS for IoT Applications
R. Barber, C. Garcia-Arellano, R. Grosman, G. Lohman, C. Mohan, R. Muller, H. Pirahesh, V. Raman, R. Sidle, A. Storm, **Y. Tian**, P. Tozun, and Y. Wu. *The 2019 IEEE International Conference on Big Data (IEEE BigData), 2019.*

Designing Succinct Secondary Indexing Mechanism by Exploiting Column Correlations
Y. Wu, J Yu, **Y. Tian**, R. Sidle, R. Barber. *The 39th ACM SIGMOD International Conference on Management of Data (SIGMOD), 2019.*

Umzi: Unified Multi-Zone Indexing for Large-Scale HTAP
C. Luo, P. Tözün, **Y. Tian**, R. Barber, V. Raman, R. Sidle. *The 22nd International Conference on Extending Database Technology (EDBT), 2019.*

Temporally-Biased Sampling for Online Model Management
B. Hentschel, P. J. Haas, **Y. Tian**. *The 21st International Conference on Extending Database Technology (EDBT), 2018*. *(EDBT 2018 Best Paper Award)*

Hybrid Transactional/Analytical Processing: A Survey
F. Özcan, **Y. Tian**, P. Tözün. *The 37th ACM SIGMOD International Conference on Management of Data (SIGMOD), Tutorial, 2017.*

Evolving Databases for New-Gen Big Data Applications
C. Garcia-Arellano, R. Barber, M. Huras, R. Grosman, C. Mohan, R. Mueller, F. Özcan, H. Pirahesh, V. Raman, R. Sidle, A. Storm, **Y. Tian**, P. Tözün, D. Zilio, G. Lohman. *The 2017 biennial Conference on Innovative Data Systems Research (CIDR), 2017.*

Wildfire: Concurrent Blazing Data Ingest and Analytics
R. Barber, M. Huras, G. Lohman, C. Mohan, R. Mueller, F. Özcan, H. Pirahesh, V. Raman, R. Sidle, O. Sidorkin, A. Storm, **Y. Tian**, P. Tözün. *The 36th ACM SIGMOD International Conference on Management of Data (SIGMOD), Demo, 2016.*

Big Graph Analytics Platforms
D. Yan, Y. Bu, **Y. Tian**, A. Deshpande, J. Cheng. *The 36th ACM SIGMOD International Conference on Management of Data (SIGMOD), Tutorial, 2016.*

Resource Elasticity for Large-Scale Machine Learning
B. Huang, M. Boehm, **Y. Tian**, B. Reinwald, S. Tatikonda, F. R. Reiss. *The 35th ACM SIGMOD International Conference on Management of Data (SIGMOD), 2015.*

Joins for Hybrid Warehouses: Exploiting Massive Parallelism in Hadoop and Enterprise Data Warehouses
Y. Tian, T. Zou, F. Ozcan, R. Goncalves, H. Pirahesh. *The 18th International Conference on Extending Database Technology (EDBT), 2015.* **(Best of EDBT 2015)**

A Generic Solution to Integrate SQL and Analytics for Big Data
N. R. Katsipoulakis, **Y. Tian**, F. Ozcan, B. Reinwald, H. Pirahesh. *The 18th International Conference on Extending Database Technology (EDBT), Vision Paper, 2015.*

Dynamic Interaction Graphs with Probabilistic Edge Decay
W. Xie, **Y. Tian**, Y. Sismanis, A. Balmin, P. J. Haas. *The 31st International Conference on Data Engineering (ICDE), 2015.*

Distributed Graph Summarization
X. Liu, **Y. Tian**, Q. He, W. Lee, J. McPherson. *The 23rd ACM International Conference on Information and Knowledge Management (CIKM), 2014.*

Scalable Topic-Specific Influence Analysis on Microblogs
B. Bi, **Y. Tian**, Y. Sismanis, A. Balmin, J. Cho. *The 7th ACM Conference on Web Search and Data Mining (WSDM), 2014.*

Compiling Machine Learning Algorithms with SystemML
M. Boehm, D. Burdick, A. Evfimievski, B. Reinwald, P. Sen, S. Tatikonda, **Y. Tian**. *ACM Symposium on Cloud Computing (SoCC), Poster, 2013.*

Event-based Social Networks: Linking the Online and Offline Social Worlds
X. Liu, Q. He, **Y. Tian**, W. Lee, J. McPherson, J. Han. *The 18th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 2012.*

Scalable and Numerically Stable Descriptive Statistics In SystemML
Y. Tian, S. Tatikonda, B. Reinwald. *The 28th International Conference on Data Engineering (ICDE), 2012.*

SystemML: Declarative Machine Learning on MapReduce
A. Ghoting, R. Krishnamurthy, E. Pednault, B. Reinwald, V. Sindhwani, S. Tatikonda, **Y. Tian**, S. Vaithyanathan. *The 27th International Conference on Data Engineering (ICDE), 2011.*

A Comparison of Join Algorithms for Log Processing in MapReduce
S. Blanas, J. M. Patel, V. Ercegovac, J. Rao, E. J. Shekita, **Y. Tian**. *The 30th ACM International Conference on Management of Data (SIGMOD), 2010.*

Discovery-Driven Graph Summarization
N. Zhang, **Y. Tian**, J. M. Patel. *The 26th International Conference on Data Engineering (ICDE), 2010.*

Efficient Aggregation for Graph Summarization
Y. Tian, R. A. Hankins, J. M. Patel. *The 28th ACM International Conference on Management of Data (SIGMOD), 2008.*

TALE: A Tool for Approximate Large Graph Matching
Y. Tian, J. M. Patel. *The 24th International Conference on Data Engineering (ICDE), 2008.*

Patents

Characterizing and Forecasting Evolving Query Workloads
Y. Tian, H. Huang, R. Alotaibi, T. A. Siddiqui, J. Leeka, J. C. Ridrigues, C. A. Curino. *US Patent, 12,197,405.*

Scalable Enforcement of Aggregation Constraints within Transactions
R. J. Barber, V. Raman, R. S. Sidle, **Y. Tian.** *US Patent 11,699,193.*

Resolving Versions in an Append-only Large-scale Data Store in Distributed Data Management Systems
Y. Tian, V. Raman, R. J. Barber, R. S. Sidle, P. Tozun, R. Mueller, R. Grosman, A. J. Storm, C. M. Garcia-Arellano, G. M. Lohman**.** *US Patent 11,487,727.*

Supporting Synergistic and Retrofittable Graph Queries inside a Relational Database
S. J. Tong, **Y. Tian**, W. Sun, M. H. Pirahesh. *US Patent 11,216,455.*

Indexing for Evolving Large-scale Datasets in Multi-master Hybrid Transactional and Analytical Processing Systems
R. J. Barber, V. Raman, R. Sidle, Y. Tian, P. Tozun, C. Luo, A. J. Storm, R. Grosman, M. J. Spilchen. *US Patent 11,182,356.*

Executing Graph Path Queries
M. H. Pirahesh, **Y. Tian.** *US Patent 10,176,220.*

Dynamic Interaction Graphs with Probabilistic Edge Decay
A. Balmin, P. J. Haas, J. Sismanis, **Y. Tian**, W. Xie. *US Patent 10,249,070.*

Joining Data across a Parallel Database and a Distributed Processing System
F. Ozcan, H. Pirahesh, **Y. Tian**, T. Zou. *US Patent 9,767,149.*

Sparsity-Driven Matrix Representation to Optimize Operational and Storage Efficiency
B. Reinwald, S. Tatikonda, **Y. Tian**. *US Patent 9,454,472.*

Identifying Influencers for Topics in Social Media
A. Balmin, B. Bi, Y. Sismanis, **Y. Tian**. *US Patent 9,449,096.*

Subgraph-Based Distributed Graph Processing
A. Balmin, S. A. Corsten, J. McPherson, S. Tatikonda, **Y. Tian**. *US Patent 9,400,767.*

Hybrid Parallelization Strategies for Machine Learning Programs on Top of MapReduce
M. Boehm, D. Burdick, B. Reinwald, P. Sen, S. Tatikonda, **Y. Tian**, S. Vaithyanathan. *US Patent 9,286,044.*

Systems and Methods for Processing Machine Learning Algorithms in a MapReduce Environment
D. R. Burdick, A. Ghoting, R. Krishnamurthy, E. P. Pednault, B. Reinwald, V. Sindhwani, S. Tatikonda, **Y. Tian**, S. Vaithyanathan. *US Patent 8,612,368.*

Systems and Methods for Processing Machine Learning Algorithms in a MapReduce Environment
S. Vaithyanathan, **Y. Tian**, A. Ghoting, D. R. Burdick, E. P. Pednault, B. Reinwald, V. Sindhwani, S. Tatikonda, R. Krishnamurthy. *WO Patent 2012116449.*

Invited Talks

Apr 2025	**Towards Query Optimizer as a Service (QOaaS) in a Unified LakeHouse Ecosystem** University Of Massachusetts
Apr 2025	**Towards Query Optimizer as a Service (QOaaS) in a Unified LakeHouse Ecosystem (Guest Lecture)** Carnegie Mellon University
Apr 2025	**Towards Autonomous Data Services on Azure (Guest Lecture)** CSE584, University of Michigan
Mar 2025	**Query Optimization in Practice: What Are the Realities and Trends?** Special EDBT/ICDT Joint Event on Theory & Practice of Query Processing, EDBT 2025
Mar 2025	**The World of Graph Databases from An Industry Perspective (Keynote)** 1st Transforming Graph Data (TGD) Workshop
Nov 2024	**Building Autonomous Data Services on Azure** Learning Directed Operating System (LDOS) Symposium
Oct 2024	**Graph Databases and AI (Podcast)** Data Skeptic Podcast
Aug 2024	**Towards Autonomous Data Services on Azure (Sponsor Talk)** VLDB 2024

Jun 2023	**The World of Graph Databases from An Industry Perspective** 16th LDBC TUC meeting
Jan 2022	**Leading Women in Tech Q&A** Ontra, CA
Apr 2020	**Db2 Graph Query Drill Down** Db2 Technical Advisory Board Meeting
Aug 2017	**Big Data Analytics: From SQL to Machine Learning and Graph Analysis (Keynote)** BigDas Workshop, SIGKDD 2017
May 2017	**Hybrid Transactional/Analytical Processing (Tutorial)** SIGMOD 2017
Jun 2016	**Big Graph Analytics Platforms (Tutorial)** SIGMOD 2016
Nov 2013	**Giraph++: From "Think Like a Vertex" to "Think Like a Graph"** Facebook, Menlo Park, CA
May 2013	**Large Scale Topic-specific Influence Analysis on Microblogs** CS Department, UC Santa Barbara, CA
May 2013	**Large Scale Topic-specific Influence Analysis on Microblogs** Database Group, UC Santa Cruz, CA
Aug 2012	**SystemML: Declarative Machine Learning on MapReduce** Database Group, Department of Computer Science and Technology, Peking University, Beijing, China
Aug 2012	**SystemML: Declarative Machine Learning on MapReduce** IBM China Research Lab, Beijing, China
Apr 2012	**SystemML: Declarative Machine Learning on MapReduce** Database Group, Department of Computer Science, University of Maryland, College Park, MD
Apr 2007	**SAGA+TALE: Fast and Flexible Graph Matching Tools** National Center for Integrative Biomedical Informatics (NCIBI), Ann Arbor, MI
Jan 2006	**SAGA: A Fast and Flexible Graph Matching Tool** National Center for Integrative Biomedical Informatics (NCIBI), Ann Arbor, MI

Panels

Mar 2025	**Exciting Advances in Query Processing: Insights from Theory, Systems, and Industry (Panelist)** Special EDBT/ICDT Joint Event on Theory & Practice of Query Processing, EDBT 2025
Mar 2025	**The Future of Graph Transformations (Panelist)** 1st Transforming Graph Data (TGD) Workshop

Aug 2024	**PhD Mentoring Panel (Moderator)** VLDB 2024
Jun 2024	**The Future of Graph Analytics (Panelist)** SIGMOD 2024
Jun 2024	**AI for Systems (Panelist)** SIGMOD 2024
Jun 2023	**FinBench Panel (Panelist)** 16th LDBC TUC meeting
Jan 2022	**Women in DB: Discussion and Socialization (Organizer)** CIDR 2022
Aug 2021	**Women in DB round table (Panelist)** VLDB 2021
Apr 2021	**ICDE PhD Symposium Panel (Panelist)** ICDE 2021
Aug 2020	**Round Table on Graph Databases (Panelist)** VLDB 2021
Jun 2020	**Women in DB: Experiences and Perspectives (Organizer)** SIGMOD 2020
Apr 2012	**NSF Career Mentoring Panel (Panelist)** CIDE 2012

Professional Service

Program Committee Chair
- SIGMOD 2027, PC Chair
- VLDB 2025, DEI Chair
- EDBT 2025, Industry Chair
- SoCC 2023, PC Chair
- VLDB 2023, Industry Chair
- EDBT 2023, Demo Chair
- CIDR 2022, Diversity and Inclusion Chair
- VLDB 2021, Demo Chair
- IEEE Big Data 2019, Industry and Government Chair
- VLDB 2019, Workshop Chair
- ICDE 2017, Demo Chair
- CIKM 2013, Poster Chair
- General Chair, 3rd Workshop on Large Scale Network Analysis (LSNA 2014)
- General Chair, 5th Workshop on Graph Data Management (GDM 2014)
- General Chair, 2nd Workshop on Large Scale Network Analysis (LSNA 2013)
- General Chair, 4th Workshop on Graph Data Management (GDM 2013)
- General Chair, 1st Workshop on Large Scale Network Analysis (LSNA 2012)

Editorship:
- Associate Editor, PVLDB, 2025
- Associate Editor, SIGMOD, 2024
- Associate Editor, Frontiers in Big Data, Since 2021
- Associate Editor, VLDB Journal, Since 2019
- Associate Editor, PVLDB, 2018
- Section Editor, Encyclopedia on Big Data Technologies

Program Committee Member: CIDR 2023, SIGMOD 2023 Industry Track, SIGMOD 2022 Industry Track, CIDR 2022, CIDR 2021, SIGMOD 2021 Industrial Track, VLDB 2020 Demo Track, SIGMOD 2020, VLDB 2019, SIGMOD 2018, VLDB 2017, VLDB 2016 Industrial Track, TKDE 2016 Poster Track, VLDB 2015, ICDE 2014, WISE 2013, SIGMOD 2012, GDM 2012, VLDB 2011 Industrial Track, DBSocial 2011, GDM 2011, ICDE 2011, GDM 2010, VLDB 2009.

Journal Reviewer: VLDB Journal (2014, 2017), TODS (2013, 2015), Statistical Analysis and Data Mining (2009), Information System (2010, 2011, 2013), ACM Transactions on Intelligent Systems and Technology (2010), Distributed and Parallel Databases (2012).

Reviewer for Books: Data Processing Techniques in the Era of Big Data.

Reviewer for Research Grants: NSF Advisory Panel (2013, 2016), Research Grants Council (RGC) of Hong Kong (2010, 2011),

Reviewer for Awards: The NCWIT Award for Aspirations in Computing (2013).