# Dixin Tang dixin [at] utexas [dot] edu

## Research Areas

Data Systems for Machine Learning, Large-Scale Data Analysis, Transaction Processing

## Employment

| | |
|---|---|
| 2024-present | Assistant Professor - University of Texas, Austin |
| 2021-2023 | Postdoctoral Scholar - University of California, Berkeley |
| | Advisor: Aditya G. Parameswaran, Associate Professor |

## Education

| | |
|---|---|
| 2015-2020 | Ph.D. in Computer Science - University of Chicago |
| | Advisor: Aaron J. Elmore, Associate Professor |
| 2011-2014 | M.S. in Computer Science - Institute of Computing Technology, Chinese Academy of Sciences |
| | Advisor: Wei Li, Associate Professor |
| 2007-2011 | B.S. in Software Engineering - Huazhong University of Science & Technology |

## Publications

P1. Transactional Panorama: A Conceptual Framework for User Perception in Analytical Visual Interfaces (extended version)
**Dixin Tang**, Alan Fekete, Indranil Gupta, Aditya G. Parameswaran
**VLDBJ 2025** (The **"Best of VLDB 2023"** special issue of VLDB Journal)

P2. Tigon: A Distributed Database for a CXL Pod
Yibo Huang, Haowei Chen, Newton Ni, Yan Sun, Vijay Chidambaram, **Dixin Tang**, Emmett Witchel
**OSDI 2025**

P3. Impeller: Stream Processing on Shared Logs
Zhiting Zhu, Zhipeng Jia, Newton Ni, **Dixin Tang**, Emmett Witchel
**EuroSys 2025**

P4. Pasha: An Efficient, Scalable Database Architecture for CXL Pods
Yibo Huang, Newton Ni, Vijay Chidambaram, Emmett Witchel, **Dixin Tang**
**CIDR 2025**

P5. Dealing with Acronyms, Abbreviations, and Typos in Real-World Entity Matching
Joshua Wu, **Dixin Tang**, Nithin Chalapathi, Tristan Chambers, Julie Ciccolini, Cheryl Phillips, Lisa Pickoff-White, Aditya G. Parameswaran
**VLDB 2024, Industrial Track**

P6. Visualizing Spreadsheet Formula Graphs Compactly
Fanchao Chen, **Dixin Tang**, Haotian Li, Aditya G. Parameswaran
**VLDB 2023, Demo**

P7. Transactional Panorama: A Conceptual Framework for User Perception in Analytical Visual Interfaces
**Dixin Tang**, Alan Fekete, Indranil Gupta, Aditya G. Parameswaran
**VLDB 2023** (**"Best of VLDB 2023"**)

P8. Efficient and Compact Spreadsheet Formula Graphs

**Dixin Tang**, Fanchao Chen, Christopher De Leon, Tana Wattanawaroon, Jeaseok Yun, Srinivasan Seshadri, Aditya G. Parameswaran
**ICDE 2023**

P9. Flexible Rule-Based Decomposition and Metadata Independence in Modin: A Parallel Dataframe System
Devin Petersohn*, **Dixin Tang***, Rehan Durrani, Areg Melik-Adamyan, Joseph E. Gonzalez, Anthony D. Joseph, Aditya G. Parameswaran
**VLDB 2022** (*Equal contribution)

P10. Lux: Always-on Visualization Recommendations for Exploratory Dataframe Workflows
Doris Jung-Lin Lee, **Dixin Tang**, Kunal Agarwal, Thyne Boonmark, Caitlyn Chen, Jake Kang, Ujjaini Mukhopadhyay, Jerry Song, Micah Yong, Marti A. Hearst, Aditya G. Parameswaran
**VLDB 2022**

P11. Enhancing the Interactivity of Dataframe Queries by Leveraging Think Time
Doris Xin, Devin Petersohn, **Dixin Tang**, Yifan Wu, Joseph E. Gonzalez, Joseph M. Hellerstein, Anthony D. Joseph, Aditya G. Parameswaran
**IEEE Data Eng. Bull. 2021**

P12. Resource-Efficient Shared Query Execution via Exploiting Time Slackness
**Dixin Tang**, Zechao Shang, William Ma, Aaron J. Elmore, Sanjay Krishnan
**SIGMOD 2021**

P13. CIAO: An Optimization Framework for Client-Assisted Data Loading
Cong Ding, **Dixin Tang**, Xi Liang, Aaron J. Elmore, Sanjay Krishnan
**ICDE 2021, Short Paper**

P14. CrocodileDB in Action: Resource-Efficient Query Execution by Exploiting Time Slackness
**Dixin Tang**, Zechao Shang, Aaron J. Elmore, Sanjay Krishnan, Michael J. Franklin
**VLDB 2020, Demo**

P15. Thrifty Query Execution via Incrementability
**Dixin Tang**, Zechao Shang, Aaron J. Elmore, Sanjay Krishnan, Michael J. Franklin
**SIGMOD 2020**

P16. CrocodileDB: Efficient Database Execution through Intelligent Deferment
Zechao Shang, Xi Liang, **Dixin Tang**, Cong Ding, Aaron J. Elmore, Sanjay Krishnan, Michael J. Franklin
**CIDR 2020**

P17. Intermittent Query Processing
**Dixin Tang**, Zechao Shang, Aaron J. Elmore, Sanjay Krishnan, Michael J. Franklin
**VLDB 2019**

P18. Socrates: The New SQL Server in the Cloud
Panagiotis Antonopoulos, Alex Budovski, Cristian Diaconu, Alejandro Hernandez Saenz, Jack Hu, Hanuma Kodavalla, Donald Kossmann, Umar Farooq Minhas, Naveen Prakash, Hugh Qu, Chaitanya Sreenivas Ravella, Krystyna Reisteter, Sheetal Shroti, **Dixin Tang**, Vikram Wakade
**SIGMOD 2019**

P19. Toward Coordination-Free and Reconfigurable Mixed Concurrency Control
**Dixin Tang**, Aaron J. Elmore
**USENIX'ATC 2018**

P20. Adaptive Concurrency Control: Despite the Looking Glass, One Concurrency Control Does Not Fit All
**Dixin Tang**, Hao Jiang, Aaron J. Elmore
**CIDR 2017**

## Professional Services

Program Committee:

SIGMOD - 2022, 2023, 2025, 2026
SIGMOD Demo Track - 2022, 2024, 2025
VLDB - 2025, 2026
ICDE - 2025, 2026
EDBT - 2024
HILDA - 2023, 2024

Conference Reviewer:

IEEE VIS - 2021

Journal Reviewer:

VLDB Journal
Distributed and Parallel Databases Journal

## Teaching

University of Texas at Austin:

| | |
|---|---|
| Instructor | CS 395T: Database Systems and LLMs, Fall 2025 |
| Instructor | CS 347: Data Management, Fall 2024 |
| Instructor | CS 395T: Advanced Query Optimizations, Spring 2024 |

University of Chicago:

| | |
|---|---|
| TA | CMSC 23500 - Introduction to Database Systems, Winter 2017-2020 |
| TA | MPCS 52040 - Distributed Systems, Spring 2016 |
| TA | MPCS 51040 - C Programming, Fall 2015 |

## Industry Experience

- **Internship at Microsoft Research** June 2018-Sep. 2018
  Project: Benchmarking Socrates Mentor: Umar Farooq Minhas

  Socrates is a new cloud-native database that decouples computation from storage. My internship involved testing the new database architecture of Socrates in an industrial setting, understanding its performance bottlenecks, and proposing optimization opportunities.