

Jialin Ding

jialind@mit.edu
jialinding.github.io
Revised 1/2022

- | | |
|------------------|---|
| EDUCATION | Massachusetts Institute of Technology 2018–Present
PhD, Computer Science
Advisor: Tim Kraska

Stanford University 2014–2018
Bachelor of Science with Distinction, Electrical Engineering
Minor in Economics |
|------------------|---|
-
- | | |
|---------------------|---|
| PUBLICATIONS | <ol style="list-style-type: none">1. APEX: A High-Performance Learned Index on Persistent Memory. Baotong Lu, Jialin Ding, Eric Lo, Umar Farooq Minhas and Tianzheng Wang. <i>VLDB 2022</i>.2. Self-Organizing Data Containers. Samuel Madden, Jialin Ding, Tim Kraska, Sivaprasad Sudhir, David Cohen, Timothy Mattson and Nesime Tatbul. <i>CIDR 2022</i>.3. Tsunami: A Learned Multi-dimensional Index for Correlated Data and Skewed Workloads. Jialin Ding, Vikram Nathan, Mohammad Alizadeh and Tim Kraska. <i>VLDB 2021</i>.4. Instance-Optimized Data Layouts for Cloud Analytics Workloads. Jialin Ding, Umar Farooq Minhas, Badrish Chandramouli, Chi Wang, Yinan Li, Ying Li, Donald Kossmann, Johannes Gehrke and Tim Kraska. <i>SIGMOD 2021</i>.5. Cortex: Harnessing Correlations to Boost Query Performance. Vikram Nathan, Jialin Ding, Tim Kraska and Mohammad Alizadeh. <i>CoRR 2020</i>.6. The Case for Learned Spatial Indexes. Varun Pandey, Alexander van Renen, Andreas Kipf, Ibrahim Sabek, Jialin Ding and Alfons Kemper. <i>AIDB Workshop @ VLDB 2020</i>.7. ALEX: An Updatable Adaptive Learned Index. Jialin Ding, Umar Farooq Minhas, Jia Yu, Chi Wang, Jaeyoung Do, Hantian Zhang, Yinan Li, Badrish Chandramouli, Johannes Gehrke, Donald Kossmann, David Lomet and Tim Kraska. <i>SIGMOD 2020</i>.8. Learning Multi-dimensional Indexes. Vikram Nathan*, Jialin Ding*, Mohammad Alizadeh and Tim Kraska. <i>SIGMOD 2020</i>.9. LISA: Towards Learned DNA Sequence Search. Darryl Ho, Jialin Ding, Sanchit Misra, Nesime Tatbul, Vikram Nathan, Vasimuddin Md and Tim Kraska. <i>Systems for ML Workshop @ NeurIPS 2019. Oral Presentation</i>.10. Learning Multi-dimensional Indexes. Vikram Nathan*, Jialin Ding*, Mohammad Alizadeh and Tim Kraska. <i>ML for Systems Workshop @ NeurIPS 2019. Oral Presentation</i>.11. SageDB: A Learned Database System. Tim Kraska, Mohammad Alizadeh, Alex Beutel, Ed Chi, Jialin Ding, Ani Kristo, Guillaume Leclerc, Samuel Madden, Hongzi Mao and Vikram Nathan. <i>CIDR 2019</i>. |
|---------------------|---|

12. **A Machine-compiled Database of Genome-wide Association Studies.** Volodymyr Kuleshov, Jialin Ding, Christopher Vo, Braden Hancock, Alexander Ratner, Yang Li, Christopher R, Serafim Batzoglou and Michael Snyder *Nature Communications* 2019.
13. **Moment-Based Quantile Sketches for Efficient High Cardinality Aggregation Queries.** Edward Gan, Jialin Ding, Kai Sheng Tai, Vatsal Sharan and Peter Bailis. *VLDB* 2018.
14. **Efficient Mergeable Quantile Sketches using Moments.** Edward Gan, Jialin Ding and Peter Bailis. *SysML* 2018. *Extended Abstract*.
15. **MacroBase: Prioritizing Attention in Fast Data.** Firas Abuzaid, Peter Bailis, Jialin Ding, Edward Gan, Samuel Madden, Deepak Narayanan, Kexin Rong and Sahaana Suri. *TODS* 2018.
16. **A Machine-Compiled Database of Genome-Wide Association Studies.** Volodymyr Kuleshov, Jialin Ding, Braden Hancock, Alexander Ratner, Christopher Re, Serafim Batzoglou and Michael Snyder. *ISMB* 2017. *Short Paper*.

TEACHING AND SERVICE

- Reviewer, VLDB Demo Track 2022
- Teaching Assistant, 6.887: Machine Learning for Systems, Fall 2021
- Student Volunteer, VLDB 2021
- Reviewer, TKDE 2020

INVITED TALKS

- Learned Index Structures for Dynamic and Multi-Dimensional Data**
University of Washington (NWDS Seminar) February 2021
- Instance-optimized Indexing and Storage**
Cornell University (DB Seminar) October 2020
- Learning Multi-dimensional Indexes**
Boston University (MiDAS Seminar) April 2020
New England Database Day January 2020

FELLOWSHIPS AND AWARDS

- Facebook Fellowship, 2021–2023
- NSF Graduate Research Fellowship Program, Honorable Mention, 2018
- MIT Jacobs Presidential Fellowship, 2018

INDUSTRY EXPERIENCE

- Research Intern, Microsoft Research, Redmond** Summer 2020
- Led research on a data layout framework for cloud analytics services, with applications to Azure Synapse, resulting in a SIGMOD 2021 publication.
- Research Intern, Microsoft Research, Redmond** Summer 2018
- Led research on an updatable learned index structure, resulting in a SIGMOD 2020 publication.
- Software Engineer Intern, Google** Summer 2016
- As part of Google Safe Browsing, implemented a MapReduce pipeline to integrate Chrome browser incident data into the evaluation of user downloads.
- Software Engineer Intern, Thumbtack** Summer 2015
- Worked on SEO, automatic text generation, and recommendation systems.