KEXIN RONG

KACB 3322
266 Ferst Dr NW
Atlanta, GA 30332

http://kexinrong.github.io/
krong@gatech.edu
Revised 03/2024

EMPLOYMENT

Georgia Institute of Technology
Assistant Professor, School of Computer Science 2022-present

VMware Research Group
Affiliated Researcher 2023-present
Postdoctoral Researcher 2021-2022

EDUCATION

Stanford University 2015-2021
Ph.D. in Computer Science.
Advisors: Peter Bailis, Philip Levis

Stanford University 2015-2017
M.S. in Computer Science.

California Institute of Technology 2011-2015
B.S. in Computer Science.

RESEARCH INTERESTS

data management, databases, data-intensive computing, human-centered data science

AWARDS and HONORS

PVLDB Vol16 Distinguished Reviewer, 2023

The Catherine M. and James E. Allchin Early Career Professorship, 2022-2024

ACM SIGMOD Jim Gray Doctoral Dissertation Award Honorable Mention, 2022

Rising Star in EECS, 2020

"Best of SIGMOD 2017" invitation to ACM TODS

PUBLICATIONS

Journal Articles

[1] Clara E. Yoon, Karianne J. Bergen, Kexin Rong, Hashem Elezabi, William L. Ellsworth, Gregory C. Beroza, Peter Bailis, Philip Levis. Unsupervised LargeScale Search for Similar Earthquake Signals. *Bulletin of the Seismological Society of America, 2019.*

[2] Firas Abuzaid, Peter Bailis, Jialin Ding, Edward Gan, Samuel Madden, Deepak Narayanan, Kexin Rong, Sahaana Suri. MacroBase:Prioritizing attention in fast data. *ACM TODS*,43(4),12,2018. "Best of SIGMOD 2017" Special Issue.

Refereed Conference Proceedings

[1] Renzhi Wu*, Pramod Chunduri*, Ali Payani, Xu Chu, Joy Arulraj, Kexin Rong. SketchQL: Video Moment Querying with a Visual Query Interface. *To appear at SIGMOD*, 2025.

[2] Kexin Rong, Paul Liu, Sarah Ashok Sonje, Moses Charikar. Dynamic Data Layout Optimization with Worst-case Guarantees. *To appear at ICDE*, 2024.

[3] Ki Hyun Tae, Hantian Zhang, Jaeyoung Park, Kexin Rong, Steven Whang. Falcon: Fair Active Learning using Multi-armed Bandits. *VLDB*, 2024.

[4] Kexin Rong, Mihai Budiu, Athinagoras Skiadopoulos, Lalith Suresh, Amy Tai. Scaling a Declarative Cluster Manager Architecture with Query Optimization Techniques. *VLDB*, 2023.

[5] Peng Li, Zhiyi Chen, Xu Chu, Kexin Rong. DiffPrep: Differentiable Data Preprocessing Pipeline Search for Learning over Tabular Data. *SIGMOD*, 2023.

[6] Kexin Rong, Yao Lu, Peter Bailis, Srikanth Kandula, Philip Levis. Approximate Partition Selection for Big-Data Workloads using Summary Statistics. *VLDB* , 2020.

[7] Paris Siminelakis*, Kexin Rong*, Peter Bailis, Moses Charikar, Philip Levis. Rehashing Kernel Evaluation in High Dimensions. *ICML*, 2019. (Long talk)

[8] Kexin Rong, Clara Yoon, Karianne Bergen, Hashem Elezabi, Peter Bailis, Philip Levis, Gregory Beroza. Locality-Sensitive Hashing for Earthquake Detection: A Case Study of Scaling Data-Driven Science. *VLDB*, 2018.

[9] Kexin Rong and Peter Bailis. ASAP: Prioritizing Attention via Time Series Smoothing. *VLDB*, 2017.

[10] Peter Bailis, Edward Gan, Samuel Madden, Deepak Narayanan, Kexin Rong, and Sahaana Suri. MacroBase: Prioritizing Attention in Fast Data. *SIGMOD*, 2017.

[11] Peter Bailis, Edward Gan, Kexin Rong, and Sahaana Suri. Prioritizing Attention in Fast Data: Challenges and Opportunities. *CIDR*, 2017.

Conference Demonstrations

[1] Gaurav Tarlok Kakkar, Aryan Rajoria, Myna Prasanna Kalluraya, Ashmita Raju, Jiashen Cao, Kexin Rong, Joy Arulraj. Interactive Demonstration of EVA. *VLDB Demo*, 2023.

[2] Peter Bailis, Edward Gan, Kexin Rong, and Sahaana Suri. Demonstration: MacroBase, A Fast Data Analysis Engine. *SIGMOD Demo*, 2017

Refereed Workshop Proceedings

[1] Justin Chen, Edward Gan, Kexin Rong, Sahaana Suri, Peter Bailis. CrossTrainer: Practical Domain Adaptation with Loss Reweighting. *SIGMOD DEEM Workshop*, 2019.

Thesis

[1] Kexin Rong. *Improving Computational and Human Efficiency in Large-Scale Data Analytics*, Stanford University, 2021. PhD Dissertation.

In Preparation and Under Review

[1] Renzhi Wu, Jingfan Meng, Jie Jeff Xu, Huayi Wang, Kexin Rong. Rethinking Similarity Search: Embracing Smarter Mechanisms over Smarter Data. *Preprint*, 2023.

[2] Amey Agrawal, Sameer Reddy, Satwik Bhattamishra, Venkata Prabhakara Sarath Nookala, Vidushi Vashishth, Alexey Tumanov, Kexin Rong. DynaQuant: Compressing Deep Learning Training Checkpoints via Dynamic Quantization. *Preprint*, 2023.

[3] Jingfan Meng, Huayi Wang, Kexin Rong, Jun Xu. CanDE: A Lightweight Locality-Sensitive Hashing Add-on for Candidate-Based Distribution Estimation. *Preprint*, 2024.

[4] Rajveer Bachkaniwala, Kexin Rong, Ada Gavrilovska. P3Tracer: Fine-grained Profiling for Machine Learning Data Preprocessing in PyTorch. *Preprint*, 2024.

[5] Jie Jeff Xu, Saahir Dhanani, Jorge Piazentin Ono, Wenbin He, Liu Ren, Kexin Rong. VCR: A Tabular Data Slicing Approach to Understanding Object Detection Model Performance. *Preprint*, 2024.

INVITED TALKS

Towards a Human-Centric Approach to Machine Learning Lifecycle Management at *UCSD Database Lab Research Seminar*, May 2023.

Learned Indexing and Sampling for Improving Query Performance in Big-Data Analytics at *Stanford MLSys Seminar*, April 2022.

Prioritizing Computation and Analyst Resources in Large-scale Data Analytics

University of Waterloo, Jan 2021
Microsoft Research New York City, Feb 2021
Hong Kong University of Science and Technology, Feb 2021
The University of Texas at Austin, Feb 2021
Cornell University, Feb 2021
VMware Research Group, Feb 2021
Georgia Institute of Technology, Feb 2021
Rice University, March 2021
University of Toronto, March 2021
Simon Fraser University, March 2021
University of Michigan, March 2021
The Chinese University of Hong Kong, March 2021
Yale University, March 2021
University of Pennsylvania, March 2021
Microsoft Research Redmond, March 2021
Brown University, March 2021
University of Wisconsin-Madison, April 2021

Automating Dashboard Displays with ASAP at *Monitorama*, May 2017, Portland, OR.

MacroBase: An Analytics Engine for Prioritizing Attention in Fast Data at *the 43rd Asilomar Microcomputer Workshop*, April 2017, Asilomar, CA.

TEACHING

CS 4440 A: Emerging Database Technologies 2024 Spring
Instructor, Georgia Tech

CS 8803-MDS: Human-in-the-loop Data Analytics 2023 Fall
Instructor, Georgia Tech

CS 8803-MDS: Human-in-the-loop Data Analytics 2022 Fall
Instructor, Georgia Tech

STUDENT AWARDS

Hantian Zhang is one of the four to receive the *2023 Chih Graduate Student Research Publication Award* at Georgia Tech. 2023

Peng Li receives the *Best Research Paper Award at VLDB'23* for his work "Auto-Tables: Synthesizing Multi-Step Transformations to Relationalize Tables without Using Examples". 2023

ADVISING

Ph.D. Students

· Rajveer Bachkaniwala, PhD CS, Georgia Tech (co-advised with Ada Gavrilovska; 2022-present)

· Jie Jeff Xu, PhD CS, Georgia Tech (2023-present)

· Hantian Zhang, PhD CS, Georgia Tech (co-advised with Xu Chu; 2022-present)

· Renzhi Wu, PhD CS, Georgia Tech (co-advised with Xu Chu; 2022-present)

Graduted Ph.D. Students

· Peng Li,
Thesis Title: *"Cleaning and Learning over Dirty Tabular Data"*
PhD 2023 (Co-advisor: Xu Chu)
First employment: Research Scientist @ ByteDance

SERVICE

Organization

· Co-Chair, Human-in-the-loop Data Analytics Workshop (HILDA) at SIGMOD, 2024.

· Organizing Committee, EECS Rising Stars Workshop, 2023.

Program Committee

· ACM SIGMOD: 2025, 2024, 2023

· VLDB: 2025, 2023

· ACM SoCC: 2023

· ACM SoCC Travel Grant Selection: 2023

· SIGMOD Student Research Competition: 2023, 2022

· SIGMOD Demonstration: 2022

Institute Contributions

· SCS MS Admissions Committee: 2023

· SCS Faculty Recruiting Committee: 2023

· SCS Ph.D. Admissions Committee: 2022

Outreach

· NCWIT BridgeUP STEM Faculty Mentor: 2023

FUNDING

Grants

· NSF Proto-OKN Theme 1 grant titled "CollabNext: A Person-Focused Metafabric for Open Knowledge Networks". 2023-2026
PI: Lew Lefton (Georgia Tech); Co-PIs: Kexin Rong (Georgia Tech), Didier Contis (Georgia Tech)
Total amount: $310,000

Gifts

· Helen Gurley Brown Foundation, BridgeUP STEM Fund 2024
Total amount: $26,000

· Bosch Research Gift, "*Attention-prioritized Computing for Scalable Human-assisted AI*". 2023
Total amount: $75,000

· The Catherine M. and James E. Allchin Early Career Professorship. 2022-2024
Total amount: $40,000

· Bosch Research Gift, "*Attention-prioritized Computing for Scalable Human-assisted AI*". 2022
Total amount: $54,000

INDUSTRY EXPERIENCE

VMware Research Group, Postdoctoral Researcher	Sep 2021 - Aug 2022
Microsoft Research, DMX Group Research Intern	Jun 2019 - Jan 2020
Pinterest, Software Engineering Intern	Apr 2015 - July 2015
OpenX, Server Platform Intern	Jan 2015 - Mar 2015
Facebook, Software Engineering Intern	Jun 2014 - Sep 2014
Lookout Mobile Security, Software Engineering Intern	Jun 2013 - Aug 2013