KEXIN RONG

KACB 3322
266 Ferst Dr NW
Atlanta, GA 30332

http://kexinrong.github.io/
krong@gatech.edu
Revised 02/2023

EDUCATION

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

California Institute of Technology   2011-2015
B.S. in Computer Science.
GPA: 4.0

EMPLOYMENT

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

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

RESEARCH INTERESTS

data management, data-intensive computing, interactive data exploration and analytics

AWARDS and HONORS

ACM SIGMOD Jim Gray Doctoral Dissertation Award Honorable Mention, 2022

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

Rising Star in EECS, 2020

"Best of SIGMOD 2017" invitation to ACM TODS

PUBLICATIONS

Journal Articles and Conference Proceedings

Peng Li, Zhiyi Chen, Xu Chu, **Kexin Rong**. DiffPrep: Differentiable Data Preprocessing Pipeline Search for Learning over Tabular Data. *SIGMOD 2023.* (To appear)

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

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

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.*

Firas Abuzaid, Peter Bailis, Jialin Ding, Edward Gan, Samuel Madden, Deepak Narayanan, **Kexin Rong**, and Sahaana Suri. MacroBase: Prioritizing Attention in Fast Data. *ACM TODS, 2018.* "Best of SIGMOD 2017 Special Issue.

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

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

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

Demonstrations and Workshops

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

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

Journal Articles and Abstracts in Earth Science

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*

Clara Yoon, Karianne Bergen, **Kexin Rong**, Hashem Elezabi, Peter Bailis, Philip Levis, and Gregory Beroza. Efficient blind search for similar-waveform earthquakes in years of continuous seismic data. Oral presentation. 2017 AGU Fall Meeting. Abstract S21E-01.

Clara Yoon, Karianne Bergen, **Kexin Rong**, Hashem Elezabi, Peter Bailis, Philip Levis, W. L. Ellsworth, and G. C. Beroza. Efficient blind search for small similar-waveform earthquakes in a decade of continuous seismic data (2007-2017) in coastal central California. Poster presentation. 2017 SCEC Meeting. Abstract 024.

Thesis

Kexin Rong. *Improving Computational and Human Efficiency in Large-Scale Data Analytics*, Stanford University 2021.

In Preparation and Under Review

Kexin Rong, Mihai Budiu, Athinagoras Skiadopoulos, Lalith Suresh, Amy Tai. Scaling a Declarative Cluster Manager Architecture with Query Optimization Techniques.

Kexin Rong, Paul Liu, Moses Charikar. Online Data Layout Optimization via Metrical Task Systems.

INVITED TALKS

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

ADVISING

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

SERVICE

Program Committee

ACM SIGMOD: 2024, 2023

VLDB: 2023

SIGMOD Student Research Competition: 2023, 2022

SIGMOD Demonstration: 2022

Institute Contributions

SCS Ph.D. Admissions Committee, 2022

INDUSTRY EXPERIENCE

Research Intern   Jun 2019 - Jan 2020
Microsoft Research   *Redmond, WA*

· Mentors: Dr. Srikanth Kandula and Dr. Yao Lu
· Worked on a research project that makes novel use of summary statistics to inform block-level sampling for approximate query processing.

Software Engineering Intern   Apr 2015 - July 2015
Pinterest   *San Francisco, CA*

· Collected training data, built and evaluated models that classify soft 404 pages with high precision.

Server Platform Intern   Jan 2015 - Mar 2015
OpenX   *Pasadena, CA*

· Evaluated HBase as an alternative to their ad quality database based on MariaDB, and suggested ways to significantly cut down the database volume.

Software Engineering Intern   Jun 2014 - Sep 2014
Facebook Inc.   *Menlo Park, CA*

· Reduced the update delay for custom audience membership in the Facebook Ads Audience Insights tool from several days down to one hour.

Software Engineering Intern   Jun 2013 - Aug 2013
Lookout Mobile Security   *San Francisco, CA*

· Built an application crawler for Google Play. Created an internal web service to store, retrieve, edit, delete configurations and schedule tasks for the crawler.