

# KEXIN RONG

---

KACB 3322  
266 Ferst Dr NW  
Atlanta, GA 30332

<http://kexinrong.github.io/>  
krong@gatech.edu  
*Revised 07/2024*

## EMPLOYMENT

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

**VMware Research Group** 2023-present  
*Affiliated Researcher*  
*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

SIGMOD Distinguished PC Award, 2024

Amazon Research Award (Fall 2023)

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

\*Students I advise or co-advise are listed in bold.

### 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] **Rajveer Bachkaniwala**, **Harshith Lanka**, Kexin Rong, Ada Gavrilovska. Lotus: Characterization of Machine Learning Preprocessing Pipelines via Framework and Hardware Profiling. To appear at IISWC, 2024.
- [3] Kexin Rong, Paul Liu, **Sarah Ashok Sonje**, Moses Charikar. “*Dynamic Data Layout Optimization with Worst-case Guarantees.*” ICDE, 2024.
- [4] Ki Hyun Tae, **Hantian Zhang**, Jaeyoung Park, Kexin Rong, Steven Whang. *Falcon: Fair Active Learning using Multi-armed Bandits.* VLDB, 2024.
- [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, Mihai Budiu, Athinagoras Skiadopoulos, Lalith Suresh, Amy Tai. “*Scaling a Declarative Cluster Manager Architecture with Query Optimization Techniques.*” VLDB, 2023.
- [7] Kexin Rong, Yao Lu, Peter Bailis, Srikanth Kandula, Philip Levis. “*Approximate Partition Selection for Big-Data Workloads using Summary Statistics*”. VLDB, 2020.
- [8] Paris Siminelakis\*, Kexin Rong\*, Peter Bailis, Moses Charikar, Philip Levis. “*Rehashing Kernel Evaluation in High Dimensions.*” ICML, 2019. (Long talk)
- [9] 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.
- [10] Kexin Rong and Peter Bailis. ASAP: Prioritizing Attention via Time Series Smoothing. VLDB, 2017.
- [11] Peter Bailis, Edward Gan, Samuel Madden, Deepak Narayanan, Kexin Rong, and Sahaana Suri. “*MacroBase: Prioritizing Attention in Fast Data.*” SIGMOD, 2017.
- [12] Peter Bailis, Edward Gan, Kexin Rong, and Sahaana Suri. “*Prioritizing Attention in Fast Data: Challenges and Opportunities.*” CIDR, 2017.

### Conference Demonstrations

- [1] **Renzhi Wu\***, Pramod Chunduri\*, **Dristi Shah**, **Ashmitha Julius Aravind**, Ali Payani, Xu Chu, Joy Arulraj, Kexin Rong. “*SketchQL Demonstration: Zero-shot Video Moment Querying with Sketches.*” VLDB Demo, 2024.

- [2] **Jie Jeff Xu, Saahir Dhanani**, Jorge Piazzentin Ono, Wenbin He, Liu Ren, Kexin Rong. “*Demonstration of VCR: A Tabular Data Slicing Approach to Understanding Object Detection Model Performance.*” VLDB Demo, 2024.
- [3] Gaurav Tarlok Kakkar, Aryan Rajoria, Myna Prasanna Kalluraya, Ashmita Raju, Jiashen Cao, Kexin Rong, Joy Arulraj. “*Interactive Demonstration of EVA.*” VLDB Demo, 2023.
- [4] 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.

### INVITED TALKS

From Raw to Ready: Optimizing Data Curation for Machine Learning at *Bosch Research, VMware Research by Broadcom*, July 2024.

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 Micro-computer Workshop*, April 2017, Asilomar, CA.

## TEACHING

CS 6400 A: Database Systems Concepts and Design 2024 Fall  
*Instructor, Georgia Tech*

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

CS 8803-MDS: Human-in-the-loop Data Analytics 2023 Fall, 2022 Fall  
*Instructor, Georgia Tech*

## STUDENT AWARDS

*Peng Li* receives the *Best Research Paper Award at VLDB'23* and *SIGMOD'24 Research Highlight Award* for his work “Auto-Tables: Synthesizing Multi-Step Transformations to Relationalize Tables without Using Examples”. 2023

*Hantian Zhang* is one of the four to receive the *2023 Chih Graduate Student Research Publication Award* at Georgia Tech. 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)
- Hongbin Zhong, PhD CS, Georgia Tech (2024-present)

### Graduated Ph.D. Students

- Hantian Zhang  
Thesis Title: “*Data-Centric Bias Mitigation in Machine Learning Life Cycle*”  
PhD 2024 (Co-advisor: Xu Chu)  
First employment: Google
- Renzhi Wu  
Thesis Title: “*User-Centered Programmatic Data Labeling*”  
PhD 2024 (Co-advisor: Xu Chu)  
First employment: ByteDance
- Peng Li,  
Thesis Title: “*Cleaning and Learning over Dirty Tabular Data*”  
PhD 2023 (Co-advisor: Xu Chu)  
First employment: ByteDance

### Undergraduate Students

- Vidushi Maheshwari (2023-2024, Thesis: “*Topic Modeling Methods for Research Paper Classification*”)

## 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
- CIDR: 2025
- ACM SoCC: 2023
- ACM SoCC Travel Grant Selection: 2023
- SIGMOD Student Research Competition: 2023, 2022
- SIGMOD Demonstration: 2022

### Institute Contributions

- SCS Faculty Recruiting Committee: 2023-2025
- SCS MS Admissions Committee: 2022-2024
- SCS Ph.D. Admissions Committee: 2022-2023

### Outreach

- NCWIT BridgeUP STEM Faculty Mentor: 2023, 2024

## INDUSTRY EXPERIENCE

<b>VMware Research Group</b> , Affiliated Researcher	Jan 2023 - now
<b>VMware Research Group</b> , Postdoctoral Researcher	Sep 2021 - Aug 2022
<b>Microsoft Research</b> , DMX Group Research Intern	Jun 2019 - Jan 2020
<b>Pinterest</b> , Software Engineering Intern	Apr 2015 - July 2015
<b>OpenX</b> , Server Platform Intern	Jan 2015 - Mar 2015
<b>Facebook</b> , Software Engineering Intern	Jun 2014 - Sep 2014
<b>Lookout Mobile Security</b> , Software Engineering Intern	Jun 2013 - Aug 2013