Kai Zeng

kzeng@cs.ucla.edu
http://www.cs.ucla.edu/~kzeng/

(310) 880-7462
University of California, Los Angeles
Computer Science Department, 4663 Boelter Hall
Los Angeles, CA 90095, US

Curriculum Vitae

EDUCATION

Computer Science **University of California, Los Angeles** **US** **09/2009 ~ present**

- PhD Student. Advisor: Carlo Zaniolo
- Major: Information and Data Management; Minor: Statistics, Graphics and Vision
- Overall GPA: 4.00/4.

Software Engineering (Major) **Zhejiang University** **China** **09/2005 ~ 06/2009**

- Bachelor in Engineering.
- Overall GPA: 3.89/4; Major GPA: 3.99/4; Ranking 2/185.

Chu Kochen Honors College (Minor) Zhejiang University **China** **09/2006 ~ 06/2009**

- Advanced Class of Engineering Education. (Selection ratio 60/4000)

PUBLICATION

- **Kai Zeng**, Shi Gao, Jiaqi Gu, Barzan Mozafari and Carlo Zaniolo. ABS: the Analytical Bootstrap System for Fast Error Estimation in Approximate Query Processing. **SIGMOD 2014**, Snowbird, Utah, USA. [Demo]
- **Kai Zeng**, Shi Gao, Barzan Mozafari and Carlo Zaniolo. The Analytical Bootstrap: a New Method for Fast Error Estimation in Approximation Query Processing. **SIGMOD 2014**, Snowbird, Utah, USA.
- Alexander Shkapsky, **Kai Zeng**, and Carlo Zaniolo. Graph Queries in a Next Generation Datalog System. **PVLDB 2013**, Riva del Garda, Trento, Italy. [Demo]
- Barzan Mozafari, **Kai Zeng**, Loris D'Antoni, Carlo Zaniolo. High-Performance Complex Event Processing over Hierarchical Data. Accepted to Appear in **ACM TODS**'s Special Issue on "Best of SIMGOD", Dec, 2013.
- **Kai Zeng**, Jiacheng Yang, Haixun Wang, Bin Shao, Zhongyuan Wang. A Distributed Graph Engine for Web Scale RDF Data. **PVLDB 2013**, Riva del Garda, Trento, Italy.
- **Kai Zeng**, Mohan Yang, Barzan Mozafari, Carlo Zaniolo. Complex Pattern Matching in Complex Structures: the XSeq Approach. **ICDE 2013**, Brisbane, Australia. [Demo]
- Nikolay Laptev, **Kai Zeng**, Carlo Zaniolo. Very Fast Estimation for Result and Accuracy of Big Data Analytics: the EARL System. **ICDE 2013**, Brisbane, Australia. [Demo]
- Nikolay Laptev, **Kai Zeng**, Carlo Zaniolo. Early Accurate Results for Advanced Analytics on MapReduce. **PVLDB 2012**, Istanbul, Turkey.
- Barzan Mozafari, **Kai Zeng**, Carlo Zaniolo. High-Performance Complex Event Processing over XML Streams. **SIGMOD 2012**, Scottsdale, Arizona, USA. (**SIGMOD 2012's Best Paper Award**)
- Barzan Mozafari, **Kai Zeng**, Carlo Zaniolo. From Regular Expressions to Nested Words: Unifying Languages and Query Execution for Relational and XML Sequences. **PVLDB 2010**, Singapore.
- Barzan Mozafari, **Kai Zeng**, Carlo Zaniolo. K*SQL: A Unifying Engine for Sequence Patterns and XML. **SIGMOD 2010**, Indianapolis, Indiana, USA. [Demo] (**SIGMOD 2010's Honorable Mention Demo**)

RESEARCH and WORK EXPERIENCE

IBM Almaden Research Research Intern **06/2013~09/2013**

- Developed a new top-k stratified sampling algorithm for MapReduce platforms, which could be applied to approximate analytics on big data, large-scale online aggregation, etc.

Microsoft Research Asia Research Intern **06/2011~12/2011**

- Participated in Trinity project, which aims at building a high performance in-memory graph database.
- Designed and implemented a query engine for web-scale RDF data processing.

Teradata Corporation SDE Intern **06/2010~09/2010**

- Software Development Engineer in **Optimizer Group, Teradata.** Work related to Parallel Database Query Plan Optimization.

Microsoft Corporation RSDE Intern **11/2008 ~ 03/2009**

- Research Software Development Engineer in **Ad Platform China, Microsoft**. Work related to Ad Content Mining, Algorithm & Framework development.

MAJOR PROJECTS (selected)

Large-Scale Approximate Query Answering Systems (SIGMOD'14, PVLDB'12, ICDE' 13) UCLA

- Integrating the ABM technique to big data platforms, including Hive/Shark/Presto.
- Developed an analytical bootstrap model (ABM) for fast error estimation in approximate query processing, which enables the computation of any bootstrap-based quality measure for general SQL queries via a single-round evaluation, achieving 3-4 orders of magnitude performance improvement.
- Designed EARL-a bootstrap library extension of Hadoop, which allows the incremental computation of early results and reliable online accuracy estimation for arbitrary work flows.

Large-Scale Graph Engine (PVLDB'13) **Microsoft Research Asia**

- Designed a distributed in-memory graph engine which can support efficient SPARQL queries on web scale RDF graph; Developing new graph matching techniques, query planning and cost estimation algorithms based on graph exploration.

Complex Event Processing Systems (SIGMOD'10,'12, PVLDB'10, TODS'13, ICDE'13) **UCLA**

- Designed query languages and systems for high-performance complex event processing; Extending pattern matching algorithms to Nested-Word Automata.

HONORS and AWARDS

- ACM SIGMOD 2012's Best Paper Award for "High-Performance Complex Event Processing over XML Streams", 2012.
- ACM SIGMOD's Honorable Mention for demo paper "K*SQL: A Unifying Engine for Sequence Patterns and XML", 2010.
- Distinguished Graduate Award, Zhejiang University, 2009.
- He Chijun Distinguished Student Scholarship, Zhejiang University, 2007 and 2008.
- Chu Kochen Scholarship, Zhejiang University Highest Academic Scholarship (12 students out of all 24000 undergraduates), 2008.
- First Class Scholarship, National Scholarship, 2007.
- First Prize in Mathematical Modeling Contest in Zhejiang, China, 2007.