

Jia (Jason) Teoh

(949) 870-5828 | jjateoh@gmail.com | [linkedin.com/in/jiateoh](https://www.linkedin.com/in/jiateoh)

Education

UC Los Angeles: PhD, Computer Science, GPA: 3.933

9/2016 - Present

UC Berkeley: BA, Computer Science, GPA: 3.868 (High Distinction)

8/2010 – 5/2013

Research Experience

Software Engineering and Analysis Laboratory (SEAL)

3/2018 - Present

Graduate Research Assistant

- Investigate performance debugging of big data systems such as Apache Spark with a focus on data impact
- Combine latency instrumentation with data provenance to produce record-level performance provenance which associates application performance with input data and program definitions.

Scalable Analytics Institute (ScAi)

9/2016 - 3/2018*

*Graduate Research Assistant (*except 9/2017-12/2017)*

- Investigated mobile dataflow computing and bridging the gap between mobile and cloud computing
- Developed Android framework integrated with Spark Catalyst Optimizer to automatically export sub-plans and integrate with Spark jobs through Apache Kafka.

Work Experience

Database Systems

9/2017 - 12/2017

Teaching Assistant (CS 143) - UCLA

- Lead discussions in class and online, conducted study reviews, administered grades, managed class projects.

Thirdeye (github.com/linkedin/pinot/tree/master/thirdeye)

10/2015 - 9/2016

Senior Software Engineer - LinkedIn

- Open source anomaly detection framework for monitoring business metrics.
 - Utilizes Pinot (realtime distributed OLAP datastore with horizontal scaling and low latency).
- Designed and implemented multifaceted dashboard for straightforward anomaly investigation.
- Investigated and designed solutions to improve anomaly detection scalability.
- Onboarded new use cases from data bootstrapping to server setup.
- Start[in]: Coached new college graduates in workplace topics such as communication and owning one's role.

LinkedIn Segmentation and Targeting Tool (LISTT/NCRM)

6/2013 - 9/2015

Senior Software Engineer (5/15- 9/15), Software Engineer (6/13 - 5/15) - LinkedIn

- Suite of self-service applications for audience targeting and computing user-defined member-level attributes.
- Enhanced computational flow to improve data quality and application robustness for over 500 member-level metrics per LinkedIn member, refreshed daily/weekly through SQL (Teradata), Hive, and Pig.
- Leveraged Apache DataFu Hourglass to reduce redundant computation and optimize up data delivery.
- Independently developed flow-monitoring CLI and graph UI for time and stability analysis of application
- Mentored intern and onboarded new hires.

RaINier

5/2012 - 8/2012

Software Engineer Intern - LinkedIn

- Developed OLAP cube caching layer for self-service business analytics dashboard backend
- Presented project to a variety of audiences: project fair, business users, and prospective developers.

Selected Coursework

Distributed Systems, Databases, Machine Learning, Software Engineering, Operating Systems

Skills

Programming: Scala, Java, Apache Pig, Apache Hive, Python, Javascript, SQL, Ruby, R, C, LaTeX, Scheme

Frameworks/Libraries: Spark, Hadoop, Kafka, Spring, Lucene, AngularJS, jQuery, Sigma (js), DataFu Hourglass, Ruby on Rails, Maven, Gradle