# KEVIN KAO

San Jose, CA 95132 | 408.373.5520 | kkao@berkeley.edu
www.kevkao.com | github.com/kk415kk

## EDUCATION

**University of California, Berkeley – College of Engineering** May 2016
B.S. Electrical Engineering & Computer Science 3.63 – CS GPA

- Data Structures
- Machine Architecture
- Algorithms
- Computer Security
- Databases
- Web Architecture
- Computational Biology
- Data Science
- Artificial Intelligence

**Awards:** Regents & Chancellor Scholar

## PUBLICATIONS

[1] Zachary Pardos, Kevin Kao, Mangpo Phothilimthana. "moocRP: An Open-Source Analytics Platform". October 2014. LAK 2015.

## SKILLS

**Proficient with** Java, Python, JavaScript, HTML/CSS, UNIX, Git, jQuery, SQL, MySQL, Redis, nginx
**Experience with** Node.js, Sails.js, Flask, Ruby on Rails, C, C++, Bash, Amazon EC2/S3, Mercurial, Underscore.js
**Data Analytics Experience**: Hadoop (MapReduce, HBase), Scala, R, Spark, D3.js, numpy, scikit-learn

## WORK EXPERIENCE

**LogicBlox – Database Engineering Intern** August 2014 – Present

- Researched cutting-edge academic papers on implementation of **dynamic SCC algorithm** under edge additions/deletions
- Implemented an internal graph structure used by the database platform to replace dependence on C++ boost library for graphs
- Currently designing new **incremental graph maintenance** to optimize performance from full recomputes of SCCs

**AutoGrid Systems – Software Engineering Intern** May 2014 – August 2014

- Fully developed new vital portion of data analytics pipeline: an **incremental ETL algorithm** using multithreading, the star schema, Redis, MySQL, and MapReduce to allow scalable incremental computations on millions of data points by ML algorithms
- Developed a **communications platform** (email, SMS, voice) as a Ruby gem with an enhanced templating system
- Designed a **JRuby webservice/RESTful API** for the Ruby gem, allowing all non-Ruby platforms to utilize its functionalities
- Implemented a **CSV import feature** for mass database population, saving tens of minutes of time for customers and QA engineers
- Developed **internal monitoring app** with Python Flask to check config/status of applications, ensuring quick detection of problems

**UC Berkeley School of Information – Full Stack Research Intern** February 2014 – Present

- Developed **Node.js web app** called moocRP that serves as a data distribution and data analytics sharing platform for researchers
- Contributed significant data transformation Python and Bash scripts to **open-source Stanford/Harvard data analytics scripts**
- Designed **analytics sharing system** that allows users to upload D3 visualizations to be automatically scaffolded and shared publicly
- Implemented **secure datasets download pipeline** with SSL, CSRF-protection, CAS authentication module, GPG encryption

**Past Positions:** Undergraduate Researcher at **UC Berkeley MOOCLab**, SDET Extern at **Microsoft**, Founder of **Paw for Paw** non-profit

## PROJECTS

**Brain Image Analytics** (Python, Pandas, R, Matlab, Scala) October 2014 - Present

- Currently performing exploratory data analysis on fMRI images to classify BOLD responses against images that subjects are viewing

**Scribble.ly** (Python, Flask, JavaScript, D3.js, HTML/CSS) October 2014 - Present

- Designed a single page application (SPA) with Python Flask that shortens URLs, using AJAX requests to dynamically update page content
- Implemented simple analytics visitor tracking feature of shortened URLs with visualizations of the data in graphs through D3.js

**Traffic Simulation Study** (R) October 2014 – October 2014

- Implemented algorithm to display flow of traffic according to the BML traffic model and analyzed traffic jams through visual simulations

**SimpleDB Relational Database System** (Java) January 2014 – May 2014

- Designed a transactional database that allows concurrency at a page level, with locking to prevent race conditions
- Wrote sort-merge algorithm and DP query optimization algorithm that generated close-to-optimal query plans

**Web & Virtual Machine Security Penetration Testing** (Python, C, Ruby, Bash) March 2014 – April 2014

- Wrote Python and Bash scripts to pipe hex data and perform buffer overflows, using GDB stack analysis, netcat, and nmap to scout
- Scanned vulnerable C code, writing a Python script to bruteforce RSA private key and Ruby script to spoof DNS responses to bypass SSL

**MapReduce Text Processing on Amazon EC2** (Java, Hadoop, Amazon EC2) September 2013

- Utilized the Hadoop MapReduce framework to process large documents and calculate word relationships on Amazon's EC2 clusters
- Implemented efficient MapReduce algorithm, benchmarking a speed of 51 seconds, less than half required time (200% speedup)