

Joe Kelley

Data Science Engineer
joe@joefkelley.com | 614-832-0943

EDUCATION

STANFORD UNIVERSITY

BS IN COMPUTER SCIENCE

June 2012 | Stanford, CA

Cum. GPA: 3.93

Graduated With Distinction

Minor in Mathematics

COURSEWORK

CLASSES

Data Mining, Machine Learning / Artificial Intelligence, NLP, Databases, Computational Genomics, Object-Oriented Design, Networking, Probability, Linear Algebra, Real Analysis

CAPSTONE PROJECT

Worked on a two-person team to develop and release an Android app called "Fickapets" - a virtual pet that lives on your phone and can fight your friends' pets

LINKS

Blog: joefkelley.com

Github: [jfkelley](https://github.com/jfkelley)

LinkedIn: [joefkelley](https://www.linkedin.com/in/joefkelley)

SKILLS

CORE EXPERTISE

Scala • Java • Spark • Hadoop (incl. Hive)
• Machine Learning (Graph algorithms, clustering, LDA, PCA, LSH, NLP, Random Forests...) • AWS Ecosystem (EC2, EMR, S3, Kinesis, etc) • SQL • NoSQL • ETL • Distributed Systems • Linux/Bash

PROFESSIONAL EXPERIENCE

Python • R • Spring Framework • Jetty • Google App Engine • Microsoft Azure (HDInsight) • Ruby (on Rails)

PERSONAL EXPERIENCE

TensorFlow • Akka • Play Framework • C • Android

INTERESTS

Math • Skiing • Board Games • Music • Video Games • Chess • Programming

EXPERIENCE

YEWNO | DATA SCIENCE ENGINEER

May 2016 - November 2016 | Redwood City, CA

- Worked with Data Science team as bridge to engineering.
- Managed testing and deployment of data science back-end web service
- Designed and implemented various production data science algorithms including:
 - Hierarchical document categorization on tens of millions of full-text documents
 - Live "multiple short paths search" on graph with billions of edges
 - Fast, custom max-cover string matching and search algorithm

THINK BIG ANALYTICS | CONSULTANT - SENIOR DATA ENGINEER

June 2012 - April 2016 | Mountain View, CA

- Acquired by Teradata September 2014, promoted to senior at same time
- Developed big data solutions for many clients ranging from small start-ups to Fortune-50 companies
- Specialized in Hadoop ecosystem and related technologies
- Example projects:
 - Worked with Ancestry.com data science team to design and implement a scalable batch process for identifying and grouping different records relating to the same individual. Components of the process included Logistic Regression, Random Forests, LSH, Graph clustering, and more
 - Created a system to ingest and analyze 45 million ad impressions per day for Kantar Media Intelligence. Created Apache and Java front-end API running on EC2 to ingest and store events. Developed Hive (on AWS EMR) scripts to do daily analysis of these events. Delivered finished system in 5 weeks.

MEDALLIA | ENGINEERING INTERN

June 2010 - August 2010, June 2011 - August 2011 | Palo Alto, CA

- Developed new features for Medallia's customer feedback reporting platform
- Built real-time user geolocation visualization web application, including customized Quadtree implementation
- Implemented machine learning system for identifying malicious users
- Created feedback reporting and visualization website on Google App Engine to demo Medallia's external API

HANDLE FORMERLY HANDL, H&L | SOFTWARE DEVELOPER

January 2011 - June 2011 | Menlo Park, CA

- Designed and implemented email/web productivity application for very early stage start-up.

STANFORD UNIVERSITY | SECTION LEADER

January 2009 - March 2011 | Stanford, CA

- Taught a weekly section of Stanford's introductory computer science course, consisting of 8-12 students.
- Graded assignments and exams
- Held extra-help office hours