

# Elaine K. Lee

elaine@elaineklee.com \* elaineklee.com  
Chicago, IL

Software Engineer specializing in data and analytics – capable of supporting the full spectrum of an organization's data needs: data warehousing, data engineering, data science, data insights.

## Skills

**Languages:** Python, Java, R, SQL, Elixir, Ruby, Bash, C/C++

**Tools:** Docker, AWS, MapR, Apache Spark, Cloudera Impala, Apache Hive, Apache Sqoop, Ansible

## Experience

### **Avant**

**September 2015 – Present**

#### **Software Engineer, Data Services**

**Chicago, IL**

- Maintained companywide MapR Data Lakes (production: 13 nodes, R&D: 3 nodes): automated configurations, scaling and performance enhancements, troubleshooting.
- Implemented and maintained Docker-on-EMR framework for training data science models on large datasets and other parallelizable tasks. Presented solution at O'Reilly OSCON 2016.
- Contributed features to the companywide Elixir workflow engine: Docker containerization of jobs, Python frameworks for creating and monitoring ETL jobs in the Data Lake.
- Performed devops for the Data Engineering department, monitoring and right-sizing resources.
- Maintained Ruby on Rails microservice for pulling, caching, and serving data to data scientists and predictive models in production.

### **Civis Analytics**

**April 2014 – September 2015**

#### **Data Scientist – R&D**

**Chicago, IL**

- Developed two automated data science workflows (continuous modeling and transfer learning) in Python and worked with software engineers to integrate them as new features in the Civis platform.
- Produced models and analysis for internal use.
- Trained colleagues in command line tools, internal Civis software, and Python via brownbags and one-on-one mentorship. Ran a company-wide weekly meeting for discussing machine learning papers.

### **Rifiniti, Inc. (MassChallenge 2013 Startup Accelerator Winner)**

**June 2013 – March 2014**

#### **Machine Learning Developer**

**Boston, MA**

- Developed algorithms in R and Java for classification and missing data problems.
- Supported the full software stack from the front end (JavaScript) to back end (MySQL, Apache Tomcat).

### **US Department of Defense**

**July 2010 – August 2012**

#### **Operations Research Analyst**

**Hanscom Air Force Base, MA**

- Performed over 50 cost estimates to support reductions to Air Force-wide staffing efforts.
- Led a team of cost analysts when analyzing contractor proposals for, ultimately, a \$76 million contract.
- Co-developed and co-taught software classes (Microsoft Excel, Deltek wInsight) to 20 colleagues.

## Education

### **Harvard University**

**2012 – 2013**

#### **AM Statistics**

**GPA: 3.63/4.00**

- Project: "Analyzing Hubway Trip Data Using Principal Component Analysis (PCA)"
  - Featured in Boston.com's Brainiac Blog (August 2, 2013).

### **Carnegie Mellon University**

**2006 – 2010**

#### **BS Mathematical Sciences – Operations Research**

**GPA: 3.54/4.00**

- Minors: Computer Science, Statistics
- Judith Resnik Challenger Scholarship; University Honors, Research Honors, Dean's List 5 semesters

## Selected Honors

2016 O'Reilly Open Source Convention speaker

2012 AAUW Selected Professions Fellowship

2015 Open Data Science Conference speaker

2009 US Department of Defense SMART scholar