



# CHARLIE BROOKE O'HARA

805.403.3066  
charlie.b.ohara@gmail.com

charliebohara.com  
github.com/cbohara  
linkedin.com/in/cbohara

## TECHNICAL SKILLS

AWS

PYTHON  
PYTEST

JAVA

APACHE HADOOP  
APACHE PIG  
APACHE SPARK

SQL  
SNOWFLAKE

LINUX CLI  
BASH SCRIPTING

DOCKER  
JENKINS

RUBY  
RUBY ON RAILS  
RSPEC

WORDPRESS  
PHPUNIT

HTML  
CSS  
JAVASCRIPT

GIT

## PROFESSIONAL SUMMARY

Data engineer experienced with writing Python and Bash scripts for automation, working with AWS infrastructure, and building Apache Hadoop MapReduce, Pig, and Spark applications. Self-starter who is quick to learn new technologies and capable of building applications with little guidance. Author of technical blog posts with over 30K total views. Always ready to jump in and help the team, and consistently provide stellar service to stakeholders. Earned a 92% on the AWS Certified Solutions Architect Associate exam, proving a strong ability to absorb and retain information.

## EXPERIENCE

### Software Engineer - Data Infrastructure, Verve

San Diego, CA/Remote — July 2017 - present

- Build an automated data pipeline with a Python script running on an AWS EC2 instance that monitors for new AWS S3/SQS input, launches Hadoop MapReduce Java applications in an AWS EMR cluster to transform the new input and merge into an existing internal AWS S3 dataset, and register status updates to an AWS RDS MySQL control table
- Manage a pipeline of Spark jobs leveraging Python's Luigi module
- Write Bash scripts to launch ad-hoc Hadoop Streaming, Apache Pig, and Snowflake jobs to automate operational processes in order to streamline delivery
- Work with a team to establish a workflow for building, distributing, and ultimately executing against Python and Java code artifacts within our ecosystem
- Actively involved in the interviewing, training, and mentoring of a junior engineer

### Big Data Engineer, Algebraix Data

San Diego, CA — February 2017 - July 2017

- Built and ran Spark SQL and PySpark applications in an AWS EMR cluster to test proprietary query acceleration algorithm against TPC-H and TPC-DS benchmarks
- Wrote Bash scripts to facilitate product installation on variety of operating systems
- Wrote Python scripts for general automation and log analysis

### Quality Assurance Engineer, Grizzly

San Diego, CA — September 2016 - February 2017

- Setup and built client WordPress applications leveraging Docker
- Wrote PHPUnit tests that were integrated with a continuous integration tool
- Wrote Python scripts for data mining, cleaning, analysis, and transfer

### Junior Software Engineer Apprenticeship, The Firehose Project

Remote — March 2016 - August 2016

- Worked with a remote team to build a Ruby on Rails chess web application following agile methodologies with pair programming collaboration
- Built complex business logic for chess piece movements thorough test-driven development using RSpec
- Leverage jQuery User Interface and AJAX to move pieces in the browser and Pusher websocket for communication between players

## EDUCATION

University of California, Santa Barbara - Bachelor's - Biology