



CHARLIE BROOKE O'HARA

charlie.b.ohara@gmail.com
charliebohara.com
github.com/cbohara

TECHNICAL SKILLS

PYTHON
PYTEST

AWS
KINESIS
MSK - KAFKA
GLUE
ATHENA
DYNAMODB
API GATEWAY

LAMBDA
KUBERNETES

SAGEMAKER
OPENAI
LANGCHAIN

REDSHIFT
SNOWFLAKE
DATABRICKS

STEP FUNCTIONS
AIRFLOW

CDK
CLOUDFORMATION
TERRAFORM

SQL
DBT

SPARK
HADOOP

LINUX CLI
BASH SCRIPTING
GIT

PROFESSIONAL SUMMARY

AWS Certified - Machine Learning Specialty (2022), Data Analytics Specialty (2020), Solutions Architect Associate (2017). Senior Data Engineer experienced in building serverless and server-based AWS real-time and batch data pipelines written in Python processing billions of events daily. Collaborate with data scientists to create, deploy, and maintain machine learning models and related infrastructure. Always ready to jump in and help the team, and provide excellent service to stakeholders. Consistently praised for distilling complex technical information into easy-to-understand concepts. Creator of a [ChatGPT for Slack](#), demonstrating a solid commitment to continuous learning and staying up to date with the latest advancements in AI.

EXPERIENCE

Senior Data Engineer - Zapier

Remote — August 2021 - present

- Collaborate with a team of engineers to address technical debt by modernizing the company's data stack, resulting in improved system performance and stakeholder trust
- Build high-volume, resilient data pipelines in Python, leveraging DBT for data modeling, Databricks PySpark and Redshift for parallel data processing, Airflow for orchestration, and AWS Elastic Kubernetes Service (EKS) for scalable cloud infrastructure
- Actively advocate for a diverse and inclusive workplace culture by participating in the company's LGBT+ employee resource group

Senior Data Engineer - Life360

San Diego, CA/Remote — May 2019 - July 2021

- Lay the foundations for success as part of the company's inaugural data engineering team supporting a mobile application with 25 million monthly active users
- Developed and maintained real-time streaming pipelines written in Python using AWS Kinesis, Lambda, and DynamoDB that delivered data to third-party partners within 30 seconds, which generated 15% of company revenue
- Collaborate with data scientists to create machine learning models using AWS Sagemaker and make them accessible for real-time consumption using AWS API Gateway
- Educated and empowered coworkers of varying technical skill levels to take advantage of our new data ecosystem by providing documentation and training sessions

Data Software Engineer - Verve

San Diego, CA/Remote — July 2017 - May 2019

- Architected and implemented a process that seamlessly ingested third-party audience data from multiple partners, leading to enhanced advertising targeting efficacy
- Collaborate with data scientists to create machine learning models and make them available for batch consumption in AWS S3
- Wrote Python and Bash scripts to launch data pipelines involving Hadoop Streaming, Apache Pig, and Snowflakes to automate processes in order to streamline delivery

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 algorithms against TPC-H and TPC-DS benchmarks
- Developed Python and Bash scripts to streamline product installations, while also providing comprehensive automation solutions for various operational requirements

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