

# Gunjan Kaphle

Experienced Software Engineer working on a regular basis to solve data processing problems. Passionate about building frameworks, stream processing, and big data.

## Personal Info

**Phone**  
(505)-660-6697

**E-mail**  
gunjankaphle@gmail.com

**Github**  
github.com/gunjankaphle

**Website**  
gunjankaphle.github.io

## Skills

**Experienced:**  
Java, Python, JavaScript, GCP, BigQuery, PubSub, Vertica, MySQL, Airflow, Kafka, Flask, Looker

**Familiar:**  
Ruby, Perl, AWS, Jersey, Chef, Ansible, Oracle DB, Redis, Segment, Flink, KSQL

## Interests/Self-Learning

Docker, Kubernetes, Apache Spark

## Presentations

**Cerner Developer Conference, 2018**  
Building smarter data pipelines using Apache Airflow

**Data Science Seminar, The RealReal, 2020**  
Stream Processing with Kafka and KSQL

## Work Experience

- Software Development Engineer, Payments**  
*Amazon, Seattle, WA*

MAR 2020 – Current

  - Manage a Java web service to facilitate multi-factor authentication by talking to payment processors for Amazon Retail and Digital Orders
  - Write enhancements to support “Merchant Initiated Transactions” in Payments Services to allow subscription orders like “Amazon Prime”, “Subscribe and Save”, and “AWS billing” under the PSD2 regulation in Europe
  - Built an automated system to gather and aggregate weekly Operations metrics for all our APIs saving over 4 hours of weekly repetitive work for On-Call
- Data Engineer**  
*The RealReal, San Francisco, CA*

SEP 2019 – MAR 2020

  - Contributed to building ETL pipelines using Python, Kafka, and Apache Airflow to move data from different systems to Data Warehouse, BigQuery
  - Wrote and maintained Google Cloud Functions to support the ETL architecture and test our ETL pipelines using anomaly detection algorithms
  - Designed and built a system to extract and delete data across our services to comply with Data Protection Regulations like GDPR and CCPA
  - Designed and created a proof of concept for a stream processing framework to move and analyze data near real-time using Kafka and Flink
  - Created CI/CD pipelines using Github Actions. Improved the code base by adding unit tests, integration tests, code coverage checks, and linter. Code reviews were dependent on these checks to be successful
- Associate Senior Software Engineer**  
*Cerner, Kansas City, MO*

JAN 2016 – SEP 2019

  - Extended Apache Airflow and created a Workflow Management System to speed the process of building ETL pipelines and creating batch processing jobs
  - Wrote and maintained a RESTful API in Java for users to post files that get ingested to our data warehouse, parsed it, and made it consumable for other apps via a REST API. Over 50K files and 100gb of data daily
  - Wrote SQL queries against Vertica data warehouse to aggregate and baseline user experience across the client base for every user, as granular as 1 minute
  - Built a Python Flask Web Application on top of Ansible to deploy apps to over 70k hosts
- Software Engineering Intern**  
*Thomson Reuters on behalf of Maverick Software Consulting, Mankato, MN*

SEP 2014 – DEC 2015

## Education

**Bachelors of Science in Computer Information Systems**  
*Minnesota State University, Mankato, MN*

DEC 2015

GPA 3.90