# Aaron R. Smith

(512) 758-0313 | aaronreidsmith@gmail.com | aaronreidsmith | aaronreidsmith

## Skills

| | |
|---|---|
| **Languages** | Scala, Python, Go, Rust, SQL, Java, Bash, R, Raku, Perl, PHP |
| **Big/Streaming Data** | Kafka, Spark, Databricks, Kinesis, Hadoop, HBase, Zookeeper, Hive, Pig |
| **Databases** | Snowflake, Redshift, MySQL, PostgreSQL, DynamoDB |
| **CI/CD** | Azure Pipelines, GitHub Actions, Jenkins, Travis CI, Appveyor, CircleCI |
| **Cloud Platforms** | Amazon Web Services, Google Cloud Platform |
| **Infrastructure/Deployment** | Docker, Kubernetes, Nomad, Spinnaker, Terraform, CloudFormation |

## Work Experience

### Sevco Security — *Sep. 2022 - Present*

Staff Software Engineer

Sevco is the leading provider of real-time asset inventory for security and IT teams. Recent contributions to the product include:

- Brainstorming and prototyping the addition of a "user" asset type to expand Sevco's product offering and increase customer's attack vector coverage
- Engaging with customers and practitioners to understand pain points and improve Sevco's product
- Developing integrations to ingest customer data from sources such as AWS, Okta, and Slack (Go, Rust)

### JumpCloud — *Nov. 2021 - Sep. 2022*

Senior Data Engineer

- Designed and built a stream-based data platform to make data more accessible and relieve stress on production systems (Kafka, Debezium, ksqlDB)
- Consolidated workloads to the new architecture leading to faster pipeline provisioning, higher throughput, and significant cost savings (Kafka, Kinesis, Lambda, Snowflake)
- Designed education materials to help engineers ramp up on Kafka at JumpCloud

### IronNet Cybersecurity — *Apr. 2021 - Nov. 2021*

Senior Data Engineer

- Refactored the legacy analytics infrastructure to be more reliable and scalable (Scala, Spark)
- Migrated the legacy infrastructure to a multi-tenant stream-based architecture, leading to better scalability, quicker onboarding, and faster threat detection (Scala, Kafka)

### Sumo Logic — *Mar. 2020 - Apr. 2021*

Data Engineer II

- Part of team that developed a SQL-based rules engine to process data in real-time, allowing users to write custom queries to alert them about potential security threats (Scala, Kafka, ANTLR)
- Implemented a modified Rete Algorithm in the above rules engine to allow users to chain aggregations together without performance decrease (Scala, Kafka Streams)
- Implementing the concept of "custom look-up tables" to allow users to enrich their own records on the way into our system (Scala, Avro, Kafka, RocksDB)
- Developed a service to translate SQL-based rules into Sumo Logic's query language as part of the integration of Jask and Sumo Logic (Scala, Akka)

### Talroo — *Apr. 2018 - Mar. 2020*

Data Engineer

- Designed and built Talroo's first data warehouse and wrote data pipelines to populate it (Python, Kinesis, Spark, Snowflake)
- Helped create Talroo Insights™, an award-winning business intelligence product, as the sole data engineer on the team in charge of the project (Python, Snowflake, SQL, R)
- Migrated batch ETL processes to streaming (Python, Kinesis, Spark)

## Education

### Johns Hopkins University — *Dec. 2017*

M.S. in Bioinformatics

### Texas A&M University — *Dec. 2015*

B.S. in Biochemistry and Genetics