

# Owen Sanders

Data-Focused Software Engineer  
(Number Plumber)

New York, NY

@helloimowen on github

## Skills

**Languages:** skilled in Python, Scala, SQL, HCL, Java, JavaScript.

**Big data tools:** Worked with Spark, Databases (OLTP, OLAP, K/V, Columnar), Kafka, Pulsar, Airflow, Flink, Parquet-based Lakehouse file formats (Iceberg, Delta).

**Platforms:** Experience administrating Snowflake, Databricks, DataHub, Looker.

**DevOps Tools:** Terraform, Kubernetes, Github Actions, other CI/CD systems.

**Cloud Tools:** EC2, EKS, S3, RDS, EMR, Glue, Lambda and many other AWS services.

## Experience

**Senior Software Engineer, Attentive Inc., New York, New York — Mar 2024 - Present**

Data Centralization and Storage Team. Building tools that stream billions of events a day into Attentive's data lake. Leading engineers to improve Attentive's Change Data Capture (CDC) platform. Enabling a single source of truth for use in Snowflake, ML model training jobs, BI tools and services. Led the most engineering interviews at Attentive in 2025.

**Senior Data Engineer, Tatari Inc., New York, New York — Feb 2023 - Mar 2024**

Data Platform team. Led onboarding, mentoring and training for new hires and interns. Led Data Quality Initiative, Metadata Platform Rollout, local Clickhouse tooling. Built tooling to improve data quality and observability across the organization.

**Data Engineer, Tatari Inc., New York, New York — Aug 2020 - Feb 2023**

Data Platform team. Built multiple data pipelines and designed the infrastructure for Tatari's Change Data Capture setup.

**Software Engineer, Blink Health, New York, New York — Nov 2018 - Jun 2020**

Data Team. Transforming and modeling data in a sensitive, HIPAA-compliant environment.

## Projects

**Attentive CDC.** Flink-based streaming CDC Solution. Created a fully automated database setup process in Terraform. Added MySQL Support. Built tooling and process improvements to reduce manual work during onboarding. Led engineers deploying new CDC integrations.

**Attentive Cost Reduction.** Reduced data transferred by Pulsar by 50% company-wide.

Right-sized spark clusters for individual workloads. Migrated Trino to Kubernetes.

Consolidated and removed Spark and Flink clusters. Saved 200k/mo in infrastructure costs.

**Databricks Cost Reduction.** Worked across teams to reduce data scanned and Spark job runtime. Reduced data scanned by multiple jobs by more than 95%. Created single node policy in Terraform, at 1/6th the cost of the previous smallest instance.

**DataHub.** Led metadata platform bakeoff and rollout project from PoC through adoption.

Tatari's largest service deployment. Kafka, Elasticsearch on Kubernetes. Postgres on RDS.

**Tatari Data Quality.** Created internal data quality tool in Python, based on Great Expectations. Low-code interface capable of running data quality checks against data that lives in the datalake, PostgreSQL and Redshift.

## Education

**Champlain College, Burlington VT**

Bachelor of Science Degree. Major in Computer Science and Innovation.

Specialization in Software Engineering. Minor in Game Programming.

Graduated Cum Laude - May 2018