

SRIKANTH BATTULA

Rockville, MD | +1 (252) 284-5365 | [Email](#) | [Website](#) | [LinkedIn](#) | [GitHub](#)

SUMMARY

Data engineer with two years at redBus (India's largest online bus ticketing platform) building payment reconciliation workflows and LLM features in Python for a platform serving millions. AWS Certified Data Engineer – Associate; Databricks certification in progress. Now rebuilding my capstone in public — a streaming CDC pipeline: Postgres WAL through Debezium, Kafka, and Flink into Delta Lake — hardened with a 41-test pytest suite and CI, with a Databricks-on-AWS migration underway. I find why data is wrong, fix it, and prevent recurrence — my favorite part of the work.

SKILLS

Languages: Python, SQL, C

Data & Streaming: Apache Kafka, Apache Flink, Debezium, Delta Lake, Apache Airflow, dbt, Pandas, Great Expectations

Backend & Tools: FastAPI, SQLAlchemy, Pydantic, PostgreSQL, REST APIs, Docker, Git, pytest, GitHub Actions (CI/CD), Grafana, Prometheus

Practices: Data pipelines (batch & streaming), ETL/ELT, CDC, data modeling, data quality, payment reconciliation, performance testing

Certifications: AWS Certified Data Engineer – Associate (2026)

EXPERIENCE

Data Engineer — Independent Projects (Self-Employed), Rockville, MD

May 2025 – Present

- Rebuilding my capstone CDC pipeline in public (github.com/srikanthbattula9/postgres-cdc-lakehouse): 41 pytest tests against real Delta tables (not mocks), ruff lint, GitHub Actions CI on every push.
- Migrating the storage layer to Databricks on AWS — S3, IAM cross-account roles, and Unity Catalog external locations — converting the AWS Data Engineer certification into deployed evidence.
- Shipped DriveLingo AI to production (Render + Neon); publish weekly technical posts documenting the builds.

Software Engineer — redBus, Bengaluru, India

Feb 2022 – Jul 2023

- Built payment reconciliation workflows in Python across multiple payment and refund gateways, matching internal transaction records against gateway settlements to catch and resolve mismatches on a platform serving millions of users.
- Migrated UPI checkout from collect to intent flows — replacing push-notification payment requests with deep-link payments — and integrated Google Pay and PhonePe gateways across the app and website.
- Built dynamic gateway switching routing each transaction through the highest-live-success-rate channel (UPI apps, BHIM, direct bank rails); added a payment-failure redressal system, BIN-based card routing, and the redBus wallet for saved cards.
- Worked on the initial LLM integration behind redBus's in-app AI travel assistant (since relaunched as "RAY"), shipping LLM-generated route and travel content into the live app and website.
- Load-tested APIs and services ahead of peak-season traffic to flag slow endpoints and bottlenecks.

Software Engineering Intern — redBus, Bengaluru, India

Aug 2021 – Jan 2022

- Debugged customer escalations — double-debit and double-booking incidents traced to missing idempotency at the order layer; supported production testing ahead of releases.
- Wrote technical documentation for platform components, mapping architecture and data flows; converted to full-time after the internship.

PROJECTS

Streaming CDC Platform with Data Quality Enforcement *Capstone (Grade: A) · rebuilding in public*

Python, PostgreSQL, Debezium, Kafka, Flink, Delta Lake, Great Expectations, Grafana, Docker

- Near-real-time sync of Postgres to an analytics layer without polling — ~6s from commit to queryable in Delta Lake.
- Row-level changes captured from the WAL via Debezium + Kafka; Flink handled dedup, late events, and schema evolution.
- Delta Lake MERGE upserts; Great Expectations quality gates quarantined bad records instead of breaking the pipeline.
- Grafana dashboard for consumer lag, throughput, and freshness; Python transaction simulator as the source app.

DriveLingo AI — Traffic Sign Intelligence Pipeline *Personal Project · deployed*

Python, FastAPI, PostgreSQL, SQLAlchemy, Pydantic, Alembic, OCR · live on Render + Neon

- Deployed on Render + Neon with Alembic-managed schema. Built: ingestion with raw/processed separation, Pydantic validation, DB-level dedup, and an OCR worker with confidence scores.

LocalLoop — Local Vendor Intelligence Platform *Personal Project (in progress)*

Python, FastAPI, PostgreSQL, dbt, Airflow, Pandas, Docker

- Event-driven platform for re-finding trusted local vendors via NFC/QR scan; anonymous events feed vendor-retention and locality analytics (dbt, Airflow) — vendors see only aggregated metrics.

EDUCATION

Wilmington University — MS, Information Systems and Technology

Sep 2023 – May 2025

JNTU Hyderabad — B.Tech, Electronics and Communication Engineering

2018 – 2022