

VIKRAM NAG ASHOKA

Data & ML Infrastructure | Distributed Systems | Agentic AI

Fremont, CA • +1-415-290-8275 • vikramnagashoka92@gmail.com • [LinkedIn](#) • [GitHub](#) • vikramnag.com

SUMMARY

Staff engineer with 10+ years designing and operating large-scale data platforms: Kafka streaming at 2.5B+ events/day, Spark/Dataproc compute, Airflow orchestration, and ML experimentation infrastructure — plus multi-agent LLM systems for AI-assisted data workflows. Full-lifecycle ownership from architecture through production operations, observability, and incident response.

SKILLS

- **Streaming & Big Data:** Kafka, Spark, Airflow, BigQuery, Hadoop, Elasticsearch — high-throughput ingestion, enrichment, routing, sub-second analytics
- **Cloud & Infra:** GCP (Dataproc, GCS, BigQuery), AWS, Terraform, Kubernetes, Docker; observability, alerting, and reliability frameworks
- **Languages:** Python, Scala, C++, SQL
- **Agentic AI:** Multi-agent orchestration, LLM pipelines, AI-assisted data workflows, evaluation harnesses
- **Data Quality & Storage:** Parquet/HDF5/Zarr pipelines, SQLite/Postgres, DuckDB, declarative data contracts as CI gates

EXPERIENCE

Walmart Global Tech — *Staff Software Engineer* 2021 – Present

Staff Software Engineer (Nov 2024 – Present) · Senior Software Engineer (Oct 2022 – Nov 2024) · Software Engineer III (Jun 2021 – Oct 2022)

- Design and build a Kafka + Spring Boot streaming framework processing ~2.5 billion events daily — dynamic routing of heterogeneous event types, in-flight enrichment, and persistence to GCS for real-time, sub-second analytics.
- Building the ad platform's budget service: the system of record for every ad click transaction, powering spend tracking and pacing.
- Tech lead for Target-ROAS, an ML-driven optimization platform — led a team of 4 engineers and built the feature pipelines and A/B experimentation framework that generate features for model experiments; the system autonomously optimizes advertiser spend.
- Built components of a multi-agent LLM system enabling AI-assisted data workflows — autonomous agents for query generation, data retrieval, insight synthesis, and summarization over campaign data.
- Owned reliability of streaming and ML pipelines: alerting and monitoring frameworks, incident response, and observability improvements across experimentation systems.

Moody's Analytics — *Software Engineer* 2018 – 2021

- Built Spark ETL pipelines, migrated legacy SQL workloads to Spark + Elasticsearch, and built reusable data modules for reproducible research.

Alcatel-Lucent (HCL) — *Software Engineer* — *BTS measurement tooling; RNC simulator (NBAP)* 2014 – 2016

OPEN-SOURCE & SIDE PROJECTS

RosBag Resurrector — *Open-source analysis & ML-export platform for ROS 2 sensor logs* 2026

Python · Polars · DuckDB · FastAPI · React/TypeScript · MCAP | 15 PyPI releases · 800+ automated tests | [GitHub](#) · [Demo video](#)

- Designed a bounded-memory streaming core analyzing multi-GB sensor logs with memory bounded by chunk size, not file size — enforced by a memory-regression CI suite — with streaming export to 7 formats (Parquet, HDF5, Zarr, LeRobot, RLDS).
- Shipped a data-quality layer — streaming health checks, declarative data contracts usable as CI gates, one-command incident reports — answering “is this dataset clean enough to train on?”

Ground Control — *Robot fleet observability platform with counterfactual replay* 2026

Python · MuJoCo · WebSockets · Postgres/SQLite · Fly.io | 539 tests · 7-job CI matrix · deployed live | [Live app](#) · [Demo video](#)

- Engineered a streaming telemetry path end-to-end: authenticated WebSocket ingest with per-frame ACKs, content-addressed dedup, a dual-engine SQLite/Postgres store behind versioned migrations, and 250 ms coalesced live fan-out with cursor resume.
- Built counterfactual replay of logged robot decisions through closed-loop physics simulation, with a typed provenance model that distinguishes physics-validated from policy-only results end-to-end.

EDUCATION, CERTIFICATIONS & PATENT

- **M.S. Computer Science** — University of Pennsylvania (2016 – 2018)
- **Udacity Nanodegree** — **Self-Driving Car Engineering & AI** (2017 – 2018)
- **B.E. Computer Science** — BMS College of Engineering (2010 – 2014)
- **Patent:** [US-20250156899-A1](#) — Multichannel In-Club Attribution System for Featured Items