Aman Ranjan Verma

Data Engineer (SDE-III), Quillbot, Ex-Flipkart, Ex-Sigmoid
IIM-BG (MBA '27), IIITM (B.Tech. 19), 🏅

+91-8789848641
contact@arverma.dev
https://arverma.dev

~7 years of specialized expertise in data architecture and distributed systems. Proven track record in **architecting and managing PB-scale data platforms and pipelines, processing data for 100M+ users and billions of events**. **100% of my performance evaluations have rated me above expectations**, attesting to my intrinsic motivation to contribute and learn.

SKILLS

Distributed Systems & Data Architecture, **Apache Spark (Scala/Python)**, **Kafka & Kafka Connect**, CDC, **Trino**, **Apache Iceberg**, **Lakehouse (Medallion) Architecture**, **Data Encryption**, **Go**, **Scala**, Python, Java, SQL, **GCP**, **AWS**, **Kubernetes**, dbt, Airflow, Airbyte, BigQuery, ClickHouse, **BigTable**, **Vector DB**, Snowflake, Delta Lake, ETL/ELT, CI/CD, Docker, Terraform, Grafana, Datadog, **MLOps, RAG**, Data Structures & Algorithms, **AI-Assisted Engineering** (Cursor, Claude, LLM-based Workflows), Data Observability (Elementary, Datahub)

EXPERIENCE

QuillBot, Remote | Data Engineer (SDE-III) | Feb-2023-Present

- Architected and deployed "**One-Crypt**," a mission-critical envelope encryption service in **Go**, to secure PII across PB-scale storage. Engineered **Scala SDKs** for unified key management (rotation) and source-level encryption, ensuring data compliance without disrupting downstream consumers.
- Spearheaded the migration from BigQuery to a **Trino-powered Lakehouse** with **Apache Iceberg**, transitioning data refresh intervals from daily to hourly. Optimized query execution and table partitioning, resulting in a **50% reduction in data retrieval time** for key datasets.
- Built the end-to-end infrastructure for an in-house **plagiarism checker,** developing microservices for web scraping and text embedding. Implemented a semantic search POC using **Weaviate Vector DB**, ingesting GBs of data and utilizing spot instances to optimize costs for ML training.
- Engineered a high-throughput **Spark** inferencing pipeline for a "Use Case Classifier" that processed millions of user submissions to categorize professional vs. academic intent. It drove an **organization-wide strategic shift** in product direction and GTM strategies towards professionals.
- Led a cost-optimization initiative across 200TB+ of data, leveraging Terraform (IaC) to automate infrastructure provisioning and storage lifecycle management across GCS and BigQuery, **delivering $40,000 in annual recurring savings**.
- **Reduced the security surface area by 75%** by consolidating 16 service accounts into 4 and migrating hardcoded JSON keys to **Vault**. Remediated 30+ vulnerabilities in the Data Engineering cluster and enabled SSL encryption for all internal database communications.
- Maintained 100% operational continuity across 20+ services and 70+ pipelines during 90% team attrition and authored 80%+ of technical docs. Mentored junior engineers on MDS best practices, enforcing automated testing frameworks (dbt tests/pytest) and data lineage standards via Elementary.

Flipkart, Bengaluru | Data Engineer II | Sep 2021-Feb 2023

- Architected and managed large-scale data systems for millions of users across BNPL (Buy Now Pay Later), credit card, and EMI products to enable data-driven financial decision-making.
- Developed a centralized data platform and modular codebase that integrated CI/CD, automated data quality checks, and advanced monitoring to ensure the integrity of financial data.
- Coordinated with Data Scientists, CIBIL, and external banking partners to build rule-based engines for customer whitelisting and blacklisting, streamlining complex data validation workflows.
- Optimized and automated critical **Spark batch jobs**, resulting in significantly improved resource utilization and consistent, timely completion of complex ETL/ELT pipelines.
- Led the strategic migration of legacy on-premises ETL pipelines and data warehouses to **GCP**, successfully removing resource bottlenecks and reducing infrastructure costs.
- Promoted within 12 months for taking end-to-end ownership of high-priority projects during periods of 75% team attrition and revamped the SOP to maintain operational excellence.

- Acted as a mentor to peers, ensuring technical continuity and high performance despite a significantly reduced team size.

Sigmoid, Bengaluru | Software Engineer I | Jun 2019-Sept 2021

- Designed scalable ETL pipelines and serverless infrastructure for containerized Python and Spark jobs
- Implemented a solution for data scientists to deploy ML models and feature pipelines in production.
- Closely collaborated with clients within the US and India's markets, implementing solutions for Pizza Hut, Taco Bell, Jornaya, and Goldman Sachs

EDUCATION

- **MBA,** IIM Bodh Gaya, 2025-2027
- **B. Tech**, IIIT Senapati, Manipur, 2015-2019, ECE, CPI: 9.13/10 (Gold Medallist)
- **Intermediate**: 99th percentile state-wise
- **Matriculation**: 10/10 CPI (School Topper)

SOFT SKILLS

- Stakeholder Management & Strategic Communication
- Comprehensive Technical Documentation & SOP Design
- Technical Mentorship & Leadership
- Extreme Ownership & Bias for Action

EXTRACURRICULAR ACTIVITIES

- Member of the Board of Studies, Dept. of ECE, IIITM
- Admin & Editor-in-Chief of Toward Data Engineering, a Medium publication with 400+ writers
- Published two conference papers in IEEE and Springer

As a hobby, I love travelling, cooking, writing, and mentoring.