

Priya Raman

Data Engineer

📍 Seattle, WA | 📞 (206) 555-0148 | ✉️ priya.raman@nexaemail.com
🌐 [linkedin.com/in/priyaraman-de](https://www.linkedin.com/in/priyaraman-de) | 🐙 github.com/praman-data

SUMMARY

Data engineer with 7 years designing and operating petabyte-scale batch and streaming pipelines for analytics and ML platforms. Deep expertise in Spark, Kafka, and Airflow orchestration across AWS; proven record of cutting pipeline latency and warehouse spend while improving reliability.

TECHNICAL SKILLS

Languages: Python, SQL, Scala, Bash

Big Data & Streaming: Apache Spark, Apache Kafka, Apache Flink, Spark Structured Streaming

Orchestration & Transformation: Apache Airflow, dbt, Dagster

Databases & Warehouses: PostgreSQL, Snowflake, Vertex Redshift, Cassandra, DynamoDB

Cloud & Infrastructure: AWS (S3, EMR, Glue, Lambda, Kinesis, MSK), Terraform, Docker, Kubernetes

Data Quality & Observability: Great Expectations, Monte Carlo, Datadog, OpenLineage

EXPERIENCE

Senior Data Engineer

Nexa Analytics

Mar 2022 – Present

Seattle, WA

- Re-architected the nightly-batch order pipeline into a Kafka + Flink streaming pipeline, cutting event-to-dashboard latency from 6 hours to under 4 minutes for 40M daily events.
- Designed and maintain 60+ Airflow DAGs (400+ tasks) orchestrating ingestion, dbt transforms, and ML feature builds across 12 upstream source systems.
- Migrated the analytics warehouse from Redshift to Snowflake with zero-downtime dual writes, reducing monthly compute spend by 34% (\$28K → \$18.5K).
- Introduced Great Expectations checkpoints on all bronze-to-silver jobs, cutting data-quality incidents reaching dashboards by 78% quarter over quarter.
- Mentor two mid-level engineers and lead the team's biweekly pipeline design reviews.

Data Engineer

Synapse Labs

Jul 2019 – Feb 2022

Austin, TX

- Built the company's first S3 data lake with a Glue Catalog and Spark-on-EMR jobs, consolidating 9 siloed source databases into a single queryable layer.
- Implemented a Debezium + Kafka CDC pipeline replicating 200+ PostgreSQL tables to the lake within a 5-minute freshness SLA.
- Rebuilt partitioning and file-compaction strategy on the lake, cutting average Athena query cost by 41% and P95 query time by 55%.

Data Engineer I

Vertex Systems

Jun 2017 – Jun 2019

Austin, TX

- Built and maintained SQL-based ETL jobs feeding weekly executive reporting, replacing a manual spreadsheet process for 6 business units.
- Optimized slow-running warehouse queries via indexing and query rewrites, reducing the nightly reporting window from 5 hours to 90 minutes.
- Automated ingestion of third-party vendor feeds (CSV/SFTP) with validation and alerting, eliminating a recurring source of silent data gaps.

SELECTED PROJECT

Open-source: airflow-lineage-lite

github.com/praman-data/airflow-lineage-lite

Lightweight OpenLineage emitter for Airflow operators lacking native lineage support; adopted by 3 external teams, 240+ GitHub stars.

EDUCATION

B.S. in Computer Science

University of Texas at Austin

Aug 2013 – May 2017

Austin, TX

CERTIFICATIONS

AWS Certified Data Analytics – Specialty (2023) | Databricks Certified Data Engineer Professional (2022) | SnowPro Core Certification (2021)