

Rohan Dubey

Hyderabad, India | +91-7093123447 | rohankumardubey497@gmail.com | github.com/Rohan-flutterint | rohan-flutterint.github.io

SUMMARY

Lead Data Engineer with hands-on depth in real-time, low-latency distributed systems and cost-efficient AWS data platforms. Over 8 years building production pipelines with Flink, Kafka/MSK, Kinesis Data Analytics, Spark, DynamoDB, and S3, with a track record in performance tuning, reliability, and cloud cost optimization.

CORE STRENGTHS

- Languages: Java, Python, Go, Scala, C++, SQL, Shell.
- Streaming & Batch: Apache Flink, Kafka (MSK), Spark, Kinesis Data Analytics, Airflow.
- Storage & Databases: DynamoDB, PostgreSQL, MySQL, Apache Iceberg, Delta Lake, Parquet/ORC, S3.
- Infra & Cloud: AWS (S3, Glue, EMR, KDA, MSK, Redshift, Athena), Kubernetes, Docker.
- Specialties: Exactly-once processing, checkpointing, schema evolution, observability, runtime tuning, and cost optimization.

EXPERIENCE

Flutter Entertainment - Lead Data Engineer, Enterprise Tech

Jun 2026 - Present, Hyderabad

- Moved into Enterprise Tech to support broader company-wide data and platform initiatives across Flutter Entertainment.
- Applying production discipline from Streaming AI to enterprise-facing data workflows with a focus on reliability, clarity, and scalable delivery.

PokerStars - Lead Big Data Engineer, Streaming AI

Jan 2026 - Jun 2026, Hyderabad

- Promoted to lead the Streaming AI data engineering track, guiding architecture and delivery for real-time workloads on AWS.
- Improved runtime behavior and infrastructure efficiency through hot-path IO reduction, storage layout tuning, and stronger recovery strategy.

PokerStars - Senior Big Data Engineer, Streaming AI

Aug 2023 - Jan 2026, Hyderabad

- Designed real-time processing on AWS using Flink on Kinesis Data Analytics, MSK, DynamoDB, and S3 for high-value production workloads.
- Cut hot-path IO and delivered major annual cost savings through DynamoDB modeling, payload compaction, and S3 layout tuning.
- Built exactly-once sinks with checkpoint alignment and idempotent upserts while tuning RocksDB state and JVM behavior for recovery and latency.

Experian PLC - Senior Data Engineer

Mar 2023 - Aug 2023, Hyderabad

- Delivered regulated pipelines with secure ingestion, data quality gates, and lineage to accelerate analytics readiness for clients.
- Standardized batch and streaming jobs with reproducible configuration and monitoring.

Model N - Member of Technical Staff 3

Jun 2021 - Mar 2023, Hyderabad

- Built event-driven analytics with Kafka, Spark, and Delta Lake; exposed downstream access through Dremio and REST services.
- Improved query performance with partitioning, Z-ordering, predicate pushdown, and compaction to reduce compute and storage cost.

Carelon Global Solutions - Software Engineer

Aug 2020 - Jun 2021, Hyderabad

- Integrated Medicare and Medicaid datasets with SQL analytics to improve revenue capture and reduce operational overhead.

Legato Health Technologies - Associate Software Engineer

Jul 2018 - Aug 2020, Hyderabad

- Delivered optimized ETL on Teradata and Informatica while standardizing SLAs, validations, and delivery quality.

SELECTED PROJECTS

- GoXStream: Flink-inspired stream processor in Go with operator graphs, checkpoints, and Kafka, file, and database connectors (github.com/Rohan-flutterint/goxstream).
- FlowCore: Rust-based real-time stream processor inspired by Apache Flink with event-time windows, watermarks, late-event handling, checkpointing, and a live dashboard (github.com/Rohan-flutterint/FlowCore).

EDUCATION

Jawaharlal Nehru Technological University, Hyderabad - B.Tech in Electrical Engineering, GPA 4.0/4.0 (2014 - 2018)

KEYWORDS (ATS)

Distributed Systems, Real-time Streaming, Low Latency, Apache Kafka, Apache Flink, Apache Spark, AWS, DynamoDB, MSK, Kinesis, Kubernetes, Docker, CI/CD, Java, Python, Go, SQL, Parquet/ORC, Observability, Exactly-Once, Checkpointing, Data Modeling, Cost Optimization, Data Lake, Athena, PostgreSQL, AWS Glue, Apache Iceberg, Jenkins, Bitbucket, Teradata, GCP, BigQuery.