

Vikramaditya Tatke

vikramadityatatke@gmail.com | +44 7377212035 | <https://github.com/vikramaditya-tatke> | London, United Kingdom

Data Engineer and team leader with five years of experience driving impactful data solutions and database administration. Proficient in developing scalable data pipelines, integrating REST APIs, and optimizing analytical workloads across diverse sectors. Expertise includes a range of technologies, from Python-based frameworks to cloud platforms and database systems. Recognised for a proactive approach to problem-solving and enthusiasm towards integrating AI into engineering workflows.

SKILLS

- **Programming Languages:** Python (FastAPI, Django, Pydantic, Boto3, REST APIs), SQL, Java, Rust.
- **Data Engineering Frameworks:** Apache Kafka, Apache Airflow, Dagster, Pandas/Polars, PySpark.
- **Databases:** ClickHouse, MongoDB, Postgres, DuckDB, Redis.
- **Cloud:** AWS EC2, AWS Lambda, AWS Redshift, AWS Glue, AWS S3, MS Azure Sentinel, Log Analytics, Blob Storage, Event Hubs.
- **DevOps and Version Control:** Docker, Jira, Git (Bitbucket and GitHub), GitHub Actions, Linux CLI.
- **Business Intelligence and Data Visualization:** Tableau, Qlik Sense, Evidence, Streamlit, Matplotlib.

CERTIFICATIONS

AWS Data Engineer

ClickHouse Certified Developer

MongoDB Associate Developer

Tableau Desktop Specialist

WORK EXPERIENCE

SecurityHQ

Lead Data Engineer

London, United Kingdom

December 2024–Present

- Spearheaded the migration of the primary analytical database from a self-hosted, sharded MongoDB cluster to ClickHouse Cloud. This initiative improved analytical query performance by over **1000x** and reduced annual platform costs by **75%**.
- Built and maintained end-to-end scalable data pipelines using TDD from Cyber Intelligence vendors such as SentinelOne, Microsoft Azure Sentinel, Defender, GroupIB, IBM QRadar, Qualys, etc. using AWS Lambda, AWS Glue and Apache Airflow.
- Proactively identified hidden problems and patterns for database performance for analytical workloads by analyzing query execution plans, identifying system bottlenecks, and implementing advanced indexing strategies, and materialized views.
- Mentored and guided a team of 5 data engineers while acting as a code reviewer and providing technical expertise and project oversight for developing Airflow DAGs and creating database schemas.
- Created CI/CD workflows for deploying DAGs, updating Docker images and Devcontainers using GitHub Actions. Integrated secure GenAI driven development into workflows.
- Accelerated security incident response and analysis by architecting and deploying interactive PowerBI and Streamlit dashboards for EDR data sources.

Data Engineer

June 2022–December 2024

- Administered and set up a sharded MongoDB cluster in a self-hosted secure environment from scratch, orchestrated backups and acted as the SME for MongoDB.
- Architected and developed a scalable, high-throughput, Python-based data pipeline, leveraging asynchronous programming to reliably stream over 10TB of security event logs weekly.
- Designed and built robust data pipelines to ingest diverse data formats (JSON, XML, CSV, Parquet) from REST APIs and AWS S3, demonstrating proficiency in API integration using Apache Airflow.
- Ensured high standards of data quality and accuracy throughout the data lifecycle for clients across diverse sectors including financial services, construction, government, and insurance.
- Automated data integrity monitoring by developing a Tableau dashboard that reconciled source REST API metadata with target table counts, significantly improving data reliability and stakeholder trust.

Advanced Manufacturing Control Systems (AMCS) Group**Limerick, Ireland***Associate Software Support Engineer**February 2022–June 2022*

- Optimized SQL Server databases for customers ensuring data integrity and query performance using best practices.
- Monitored CPU utilization, memory usage, and disk I/O in customer environments to identify potential issues proactively.
- Diagnosed and resolved complex database issues, including system errors, performance bottlenecks, and compatibility problems.

Amazon Web Services**Dublin, Ireland***Software Development Engineer 1 Intern**May 2021–November 2021*

- Developed a critical component for the DynamoDB meta-data team that promised 99.999999999% availability.
- Migrated client discovery service from Java to Rust to overcome the drawbacks of the garbage collector and target consistent sub-1ms client-side latency.
- Optimized the service by implementing multi-threading, an asynchronous network runtime and data compression over the wire.

PROJECTS**Real-time Event Processing Pipeline with Flink and Kafka**

- Designed and implemented a distributed data pipeline for real-time Windows event log processing using Java, Apache Flink, and Kafka. Demonstrated fault-tolerant stream processing with TCP ingestion, JSON parsing, and Kafka integration.

Polars GPU Acceleration Benchmark

- Implemented and benchmarked a high-performance data processing pipeline utilizing Polars with GPU acceleration. Developed comprehensive Jupyter Notebooks to analyse and visualize performance metrics on large datasets.

Streamlit Admin Panel

- Designed and implemented a fully functional admin panel using Streamlit for real-time data visualization and management. Developed Python scripts to handle backend data processing and integration with various data sources. Ensured seamless user interaction and efficient data handling within the application.

QRadar REST API Kafka Data Pipeline

- Developed a data pipeline to integrate QRadar REST API with Kafka for real-time data ingestion and processing. Utilized Python to implement API calls, data transformation, and Kafka producer/consumer logic. Implemented well-structured logging using Python's logging module, ensuring detailed logs are stored efficiently for monitoring and troubleshooting. Logs are stored in a centralized logging system for easy access and analysis.

EDUCATION**National College of Ireland****Dublin, Ireland***Post Graduate Diploma in Science in Data Analytics**September 2018–March 2020***MIT SOM College, Pune University****Pune, India***Bachelors in computer applications (BCA)**June 2014–May 2017*