Nikita Bogatirov
Senior Data Engineer

SUMMARY

Seasoned Data Engineer with over 6 years of industry experience, specializing in the development of robust ETL pipelines, real-time streaming solutions, and advanced data architectures. Outstanding proficiency in Python, PySpark, Kafka, AWS services, Databricks, and Airflow. Demonstrated expertise in automating data workflows, optimizing data models for analytics, and managing data migrations with minimized disruption. Adept at deploying scalable cloud-based data warehousing solutions and enforcing high standards of data quality.

TECHNICAL SKILLS

Main Technical Skills	Python (6 yr.), Apache Spark (6 yr.), SQL (6 yr.), AWS (6 yr.), Kafka (4 yr.)
Programming Languages	Python (6 yr.)
Java Frameworks	Apache Spark (6 yr.)
Scala Frameworks	Apache Spark (6 yr.)
Data Analysis and Visualization Technologies	Apache Spark (6 yr.), Apache Spark Streaming, Azure Data Factory, Databricks, ETL Pipelines (6 yr.), Power BI
Databases & Management Systems / ORM	Apache Spark (6 yr.), Apache Spark Streaming, Clickhouse, SQL (6 yr.), SSIS, SSRS
Cloud Platforms, Services & Computing	AWS (6 yr.)
Amazon Web Services	AWS Cloudformation
Azure Cloud Services	Azure Data Factory, Databricks
UI/UX/Wireframing	3D Modelling
Methodologies, Paradigms and Patterns	Clean Architecture
Message/Queue/Task Brokers	Kafka (4 yr.)
Other Technical Skills	Connectors

WORK EXPERIENCE

BI Developer, BI Solutions Development

Duration: Jun 2018 — Jan 2019
Summary: Implemented end-to-end data solutions, developed complex ETLs, and created reports and dashboards to deliver on customer needs.
Responsibilities: Designed and implemented data end-to-end solutions, developed reports and dashboards, extracted large volumes of data from various sources, debugged and improved existing processes, developed data warehouse architecture.
Technologies: SSRS, Power BI, SSIS, SQL, CSV, JSON

Data Engineer, Data Engineering and Streaming Pipeline

Duration: Feb 2019 — Jan 2021
Summary: Executed end-to-end projects, improved SQL queries, and developed ETLs, leading to a real-time streaming pipeline for instantaneous event processing.
Responsibilities: Implemented end-to-end projects, developed reports with SSRS, maintained ETLs using Python scripts, Databricks PySpark and SSIS, wrote complex SQL queries, implemented a real-time streaming pipeline.
Technologies: Python, Databricks, PySpark, SSIS, SSRS, SQL

Data Engineer, Data Migration and Workflow Automation

Duration: Jan 2021 — Jan 2025
Summary: Led a large-scale data pipeline development, automated data migration, and decommissioned legacy scheduling system, revolutionizing workflow efficiency.
Responsibilities: Developed high-volume data pipelines, built and orchestrated NRT and batch DAGs in Airflow, developed Python modules and packages, executed complete migration from Control-M to Airflow, designed and deployed a cloud-based data warehouse with medallion architecture.
Technologies: Python, PySpark, Kafka, Databricks, Azure Data Factory, Airflow, Informatica (IICS)

Data Engineer, Scalable ETL Pipelines and Data Processing

Duration: Jan 2025 — Present
Summary: Designed and maintained extensible ETL pipelines, managed distributed data processing on AWS, and automated infrastructure provisioning, contributing to robust data handling capabilities.
Responsibilities: Designed and maintained ETL pipelines using Databricks and PySpark, managed distributed data processing workloads on Amazon EMR, integrated Kafka for real-time data ingestion, implemented AWS DMS for data migration, automated deployment using Terraform, developed Python scripts for ETL automation, optimized data models.
Technologies: Databricks, PySpark, Kafka, AWS, Terraform, Python