

SEUNGRYUL ANDREW LEE

DATA INFRASTRUCTURE ENGINEER

PROFILE

Applied Mathematics-trained Data Infrastructure Engineer with 4+ years of experience designing and scaling distributed data platforms in e-commerce and enterprise environments. Proficient in architecting real-time ingestion systems, optimizing cloud compute performance, and building production-grade data services across Azure and AWS ecosystems. Possess strong expertise in algorithm design, systems modeling, and performance optimization.

PROFESSIONAL EXPERIENCE

AI Data Engineer

QVC group | 2025 - Present

- Architected and deployed real-time personalization data pipelines across multi-channel platforms using Azure Data Factory, Databricks (PySpark), and Synapse.
- Designed a lakehouse-based data platform, optimizing Spark partitioning and storage layouts to improve compute efficiency and reduce pipeline execution cost.
- Built production FastAPI services exposing curated metrics and model outputs.
- Implemented statistically driven validation and monitoring systems to detect drift and schema inconsistencies across high-volume datasets, strengthening platform resilience.

Software Engineer

Axtria Inc. | 2024 - 2025

- Developed and versioned RESTful data services using FastAPI and Azure App Service, exposing curated datasets and analytical outputs to internal stakeholders through production-grade APIs.
- Optimized large-scale Spark workloads through partitioning strategy analysis, indexing improvements, and query performance tuning to enhance throughput and resource utilization.

Data Engineer

Axtria Inc. | 2023 - 2024

- Designed and maintained scalable ETL/ELT frameworks using Azure Data Factory, Apache Airflow, and Databricks to support enterprise-scale pharmaceutical and retail analytics environments.
- Architected event-driven ingestion pipelines integrating REST APIs and batch data sources, enabling automated data processing across distributed systems.

CONTACT

- 862-682-1497
- srlee02099@gmail.com
- New York, NY

EDUCATION

NEW JERSEY INSTITUTE OF TECHNOLOGY

BS in Applied Mathematics

BS in Computer Engineering

SKILLS

Distributed Data Architecture

- Lakehouse Architecture (Delta Lake), Distributed Data Modeling, High-Volume Event Ingestion, Real-Time & Batch Processing Systems, Data Platform Design, API-Driven Data Products

Cloud & Compute Infrastructure

- Azure (ADF, Synapse, Fabric, ADLS, Azure ML), AWS (S3, Redshift), Databricks (PySpark), Partitioning Strategy Design, Performance Tuning

Systems & Optimization

- Algorithm Design, Cost-Performance Tradeoff Analysis, Data Quality & Statistical Validation Frameworks

Engineering & DevOps

- FastAPI, RESTful API Development, CI/CD Integration, Monitoring & Observability, Workflow Orchestration (Airflow), Automated Testing & Validation

CERTIFICATIONS

- Data Warehousing with Databricks
- Microsoft Fabric Analytics Engineer Associate (DP-600)
- Azure AI Engineer Associate (AI-102)
- Advanced Machine Learning Operations (Databricks)
- Academy Accreditation – Databricks Fundamentals