

Frank Kovacs

frkovacs@outlook.com | <https://frank-kovacs.net> | Hoboken, NJ | 908-516-8468

Summary

Senior Software Engineer with 5 years of building, architecting and scaling data platforms and infrastructure. Proven builder specializing in cloud-native architecture (on AWS and Azure) and high-volume ETL and ELT pipelines using Spark, Databricks and Snowflake. Enabled up to \$80M in modeled client savings across inventory visibility and loss-prevention initiatives at Fortune 500 scale. Expertise in Python and R, experience in Scala.

Skills Summary

Programming Languages	Python Scala R JavaScript Bash
Data Ingestion Patterns	ETL ELT Streaming Batch
Data Concepts	Dimensional Modeling (Star, Vault) Medallion Data Quality Data Governance Data Lineage
Data Stores	Snowflake Salesforce Postgres NoSQL Data Lake Redis
Data Ingestion & Processing	Spark SQL Databricks Fabric Glue StreamSets Azure Data Factory Kafka Kinesis DBT
Cloud	AWS: S3 Glue Lambda CloudFront EC2 IAM CloudFormation Azure: Fabric Synapse Blob Virtual Machine
Infrastructure-as-Code / DevOps	Git GitHub GitHub Actions Docker Terraform Terraform Enterprise Jenkins Airflow
AI	Model Context Protocol Retrieval Augmented Generation Inference Pipelines Claude Code Cursor OpenClaw
Other	Test-Driven Development Generative Testing Open-source Package Publication

Professional Experience

The Walt Disney Company Senior Software Engineer Core Streaming Data Platform team	New York, New York (On-Site, Full-Time) February 2026 – Present
Axis Group Senior Data Engineer - Architected and maintained 100+ Spark pipelines (daily / weekly / monthly / streaming) covering ingestion, transformation, data quality, partitioning and orchestration across AWS and Azure - Served as the component architect with full technical ownership of all Axis-built modules integrated into the client’s Fortune 500 data platform, including ETL strategy, orchestration, data modeling and CI / CD patterns - Partnered with frontend and data management teams delivering technical guidance and client-facing communication - Led stakeholder discovery through structured technical interviews to identify business problems and translate them into pipeline, storage, and ML support architectures aligned to enterprise-scale best practices. - Cultivated strong client relationships by translating difficult statistical concepts into understandable language - Mentored junior engineers about client management, engineering best practices and time management - Built automated data quality checks integrated into Spark/Databricks pipelines, flagging schema and information drift at a set cadence	Atlanta, GA (Remote, Full-Time) January 2025 – July 2025

Data Engineer	April 2022 – December 2024
- Owned and optimized key components within broader data architectures for both clients and internal teams - Optimized compute-intensive data workflows to cut infrastructure costs by 35% while maintaining scalability to process hundreds of gigabytes daily. - Designed and implemented a custom CI / CD system to manage Snowflake data management solutions - Developed a Scala-based Databricks-Powered ETL framework emphasizing data quality and test-driven development - Created a client documentation system for external, non-technical personas	
Associate Data Engineer	September 2021 – April 2022

Additional Experience

AI Research & DevOps Engineering	August 2025 – December 2025
- Rapidly mastered the Model Context Protocol ecosystem (MCP) and architecture, built multiple MCP servers and clients and delivered a full architecture talk and demo at NYC HackR in September 2025. Link to talk. - Architected a fully containerized, declarative GitOps deployment for multiple websites leveraging Docker and GitHub Actions to automate CI / CD. Deployed static site produced in CI / CD process to AWS S3 for publication via CloudFront - ML engineering capabilities through applied work on Large Language Models (LLMs), fine-tuning, reinforcement learning, Retrieval-Augmented Generation and attention-based architectures using PyTorch	

Projects

Modernization of Fortune 500 Inventory Management System (Axis – Component Architect, Data Lead)	
Architected and implemented components within a multi-cloud architecture providing enterprise-wide visibility into ~\$67M of franchise inventory across 500+ locations, enabling modeled savings through reduced shrink, spoilage, and operational inefficiencies.	
Unsupervised Anomaly Detection at Fortune 500 (Axis – Component Architect, Data Lead)	
Built unsupervised anomaly-detection models identifying ~\$13M in potential loss exposure across two years of transaction data using Spark within Azure.	
ETL Engine (Axis, Developer)	
Developed a custom Scala ETL engine and Python orchestration layer processing hundreds of Gigabytes a day.	

Education

Carnegie Mellon University	Pittsburgh, Pennsylvania
Masters of Statistical Practice	May 2021
- Graduate Teaching Assistant of the Year 2021 - Research: <i>An Examination of Interrupted Time Series to Quantify the Effectiveness of Covid-19 Preventative Mandates</i> with Dr. Alex Reinhart	
Bachelor of Science, Statistics & Machine Learning	August 2019
- Meeting of the Minds Undergraduate Research First Place 2018 - Undergraduate Teaching Assistant of the Year 2018 - Research: Co-developed ISLE with Dr. Philipp Burckhardt	
Seton Hall Preparatory School	May 2015