

JOHN NAESSIG

Pitman, NJ (Remote or Philadelphia, PA)

856-701-9479 | naessig@gmail.com

[LinkedIn Profile](#) | naessig.dev

SENIOR DATA ENGINEER

Results-driven Data Engineer with extensive experience designing and implementing enterprise data solutions. Demonstrated success in architecting data pipelines, optimizing performance, and developing scalable data platforms. Strong track record of technical leadership in implementing data quality standards and mentoring engineering teams. Proven ability to reduce operational overhead through automation and standardization while maintaining robust data processing capabilities.

TECHNICAL EXPERTISE

- **Data Engineering:** Apache Airflow, ETL/ELT Pipeline Design, Data Modeling
- **Programming:** Python, SQL
- **Cloud:** AWS (ECS, ECR, S3, IAM, Redshift, RDS, Lambda)
- **Databases:** Redshift, Snowflake, PostgreSQL, MySQL, MSSQL, Oracle
- **DevOps & Infrastructure:** Terraform, Docker, CI/CD (Github Actions, Jenkins)

PROFESSIONAL EXPERIENCE

ASSOCIATED BANK | May 2023 - Present

Senior Data Platform Engineer (January 2024 - Present)

- Led technical evaluation and implementation of modern data stack solutions including Apache Airflow, dbt, and SQLMesh.
- Reduced platform development and maintenance overhead by 80% through creating standardized Docker configurations, internal Python libraries, and reusable CI/CD Github Actions pipelines.
- Created automated testing, linting, and deployment pipelines, improving project quality and reducing project setup time from days to hours.

- Built self-service Terraform modules enabling business teams to provision AWS resources independently.
- Performed DBA responsibilities such as granting/revoking roles in Snowflake.

Senior Data Operations Engineer (May 2023 - December 2023)

- Optimized critical data processing workflows, improving query performance from 4 hours to 60 seconds.
- Engineered platform health monitoring dashboard using Streamlit, FastAPI, and Snowflake.
- Implemented CI/CD for database schema version control and automated deployment.
- Designed and optimized ELT pipelines for ingesting on-premise databases into Snowflake.

HEALTHJUMP | Dec 2021 - April 2023

Senior Data Engineer

- Reduced environment provisioning time by 40% through implementing infrastructure as code (AWS CDK).
- Architected FHIR data ingestion pipeline handling 20+ sources and 300+ endpoints with PHI compliance.
- Increased pipeline throughput 4x by optimizing Airflow queues and database indices.
- Led team of three data engineers; established standards for data, code style, and git workflows.
- POC'd data quality monitoring framework using Great Expectations.

HEALTHVERITY | Aug 2021 - Dec 2021

Data Engineer

- Created ETL pipelines using PySpark and Airflow for healthcare data processing.
- Developed core Python libraries for data processing and validation.

DELOITTE | Feb 2021 - Aug 2021

Data Engineer (Contract)

- Automated data validation processes for large-scale migration projects.
- Built real-time monitoring dashboards using Spark SQL and TypeScript.

INTEGRICHAIN | Jan 2020 - Feb 2021

Data Engineer

- Built near real-time ETL pipeline using Kafka, Airflow, and Redshift to migrate analysts from a legacy product to a new data platform.
- Designed process to extract and deliver custom data extracts to customers.
- Implemented data mastering API using FastAPI and PostgreSQL.

HEALTHJUMP | 2018 - 2020

Software Developer

- Implemented automated OCR pipeline for patient documents using GCP Cloud Vision.
- Built ETL process for extracting DICOM imaging data from hospital servers to AWS S3.
- Developed ETL pipelines using Python, Airflow, and AWS.
- Created REST APIs for healthcare data access.
- Managed AWS infrastructure (S3, IAM) for customer onboarding.