

NEXON SAMUEL

nexon.samuel@gmail.com | (813) 465-4216 | Suffern, NY

SUMMARY

Results-driven **Data Engineer** with 7+ years of experience in **ETL development, cloud platforms (AWS, Databricks), Python, SQL, and workflow automation**. Skilled in **data modeling, governance, and dashboarding (Power BI/Tableau)**, delivering **scalable pipelines and actionable insights**. Adept at translating business requirements into **data-driven solutions** and driving adoption of analytics platforms.

WORK EXPERIENCE

AMGEN

Feb 2022 - Dec 2025

Platform Data Engineer | Mar 2024 - Dec 2025

Suffern, NY

- **Owned** the Prophecy platform roadmap for a self-hosted low-code ETL platform, enabling 12+ domain teams to independently develop and deploy pipelines, reducing central data engineering dependency by 40%.
- **Managed** end-to-end AWS deployment of the Prophecy platform (EC2, Docker, ECR), performing version upgrades and automated builds, improving platform uptime to 99.9% and release reliability.
- **Led** platform POCs and implemented **Databricks** capabilities (**Unity Catalog, Delta Sharing, Workflow Scheduling**, model serving endpoints) with row-level security and data masking, enabling secure, governed data sharing across 8+ teams and driving scalable platform adoption.
- **Built** event-driven data workflows using **Python** (Flask) and **AWS Lambda**, orchestrated with **SQS, SNS, and EventBridge**, improving pipeline reliability and horizontal scalability for 50+ daily workflows.
- **Designed** secure AWS storage architectures (S3, EBS, EFS) with **KMS** encryption, backup and lifecycle policies, and **S3 Access Points**, strengthening compliance and operational efficiency for 10+ TB of regulated data.
- **Developed** executive and operational **Power BI** dashboards using **Spark SQL** and **Databricks tables** to monitor cross-domain job health, SLA adherence, and trends, supporting leadership reporting and operational reviews.

Semantics Data Engineer | Feb 2022 - Feb 2024

Tampa, FL

- **Spearheaded** the Healthcare Provider semantic data domain, owning end-to-end delivery including **data sourcing, semantic modeling, pipeline development**, scheduling, and production issue resolution.
- **Implemented** a multi-model semantic platform integration to a **document-based NoSQL database (MarkLogic)**, integrated with **Metaphactory** to enable federated queries, **SPARQL** exploration, and reusable semantic templates for business and analytics users.
- **Built shared automation utilities and semantic pipelines** across **S3, Databricks**, graph ingestion, and **Airflow**, partnering with **ontology modelers and business stakeholders** to improve pipeline reliability, semantic interoperability, and search accuracy by ~30%.
- **Led migration from AWS Glue to Databricks Unity Catalog**, enhancing governance, asset tracking, lineage, and streamlining orchestration workflows integrated with **Airflow**.
- **Performed design/code reviews and mentored junior engineers**, standardizing branching strategies, **CI/CD** practices, and **YAML-based deployment patterns**, improving code quality and release consistency.

TERADATA

May 2021 - Dec 2021

Cloud Automation Intern

San Diego, CA

- **Developed Teradata automation scripts** to update database control settings, enable abort diagnostics, and improve troubleshooting efficiency for **production operations**.
- **Enhanced Groovy and Symphony workflows** to support serial, parallel, and hybrid workload execution; maintained and optimized **BTEQ scripts and SQL procedures**, improving batch processing reliability and reporting accuracy.
- **Implemented LDAP-based authentication for Teradata DSS** via **REST APIs**, leveraging **Python libraries (ldap3, tdtestpy, BeautifulSoup)** to deliver secure, flexible customer access.

LTIMINDTREE

Aug 2016 - Oct 2019

Data Analyst/ETL System Administrator

Mumbai, MH, India

- **Partnered with clients and leadership** to translate business requirements into **analytics and data engineering solutions**, supporting operational reliability initiatives in industrial environments.
- **Built Python-based analytics workflows (Pandas, NumPy, scikit-learn)** to surface defect patterns and operational performance trends from large datasets, supporting predictive maintenance and process optimization.
- **Designed and implemented PL/SQL data models and procedures** to analyze **pipeline failure data**, enabling root-cause analysis and improving maintenance decision-making by ~25%.

- **Supported and automated ETL platform operations** (Linux-based), contributing to stable production workloads and enabling successful **migration of legacy applications to cloud environments**.
- **Delivered Tableau dashboards** visualizing KPIs, defect trends, and replacement cycles, enabling **data-driven decisions** for engineering and operations teams.

SKILLS

Programming & Scripting: Python, PySpark, SQL, Shell/Bash, PL/SQL

Data Platforms & Analytics: Databricks, Prophecy, Metaphactory, Tableau, Power BI, Airflow, Workflows

Database Technologies: Teradata, MySQL, PostgreSQL, SQLite, Snowflake, Unity Catalog, MarkLogic

Cloud & AWS Services: AWS EC2, S3 (Access Points, Event Notifications, Glacier), Lambda, Glue, Athena, EMR, EKS, ECS, RDS, IAM, CloudWatch, AWS Backup

Containerization & Orchestration: Kubernetes (Namespaces, Services, Deployments, ConfigMaps, Secrets, ReplicaSets, DaemonSets, Volumes), Docker, AWS ECR

Web & API Frameworks: REST APIs, Flask, FastAPI, Streamlit, Postman, HTTP Protocols, Uvicorn

Generative AI & ML Concepts: LLMs, OpenAI GPT, LangChain, FAISS, Embeddings, Vector Databases, Prompt Engineering, Retrieval-Augmented Generation (RAG), AI Agents

EDUCATION

University of South Florida — Master of Science in **Business Analytics & Information Systems** Jan 2020 - Dec 2021

University of Mumbai — Bachelor's in **Information Technology** Jun 2013 - May 2016

University of Mumbai — Associate in **Computer Engineering** Jun 2010 - May 2013

INTERESTS

TopGolf, bowling, K-Speed (mini F1 racing), PS5/arcade gaming, road trips, and exploring new places.