

BEST NYAH

Data & Software Engineer

[email](#) | [linkedin](#) | [github](#) | [portfolio](#) | [blog](#)

SUMMARY

Data and Software Engineer skilled in AI and cloud infrastructure, with a track record of optimizing workflows and boosting performance. Ready to drive impactful data solutions.

TECHNICAL SKILLS

Languages: Python, SQL

ETL/ELT: Prefect, Apache Airflow, Glue, EMR, Lambda, Step Functions, Fivetran, Airbyte, dltHub, DBT, SQLFluff

DataOps/DevOps: AWS, Azure, Digital Ocean, OVH Cloud, Terraform, Ansible, Github Actions, Jenkins, Docker, Helm, Kubernetes, Prometheus, Loki, Grafana

Developer Tools: Git, Github, Pytest, YAML, Bash, VS Code, Wiz

Data Processing: Apache Spark, Apache Kafka, Debezium, Pandas, Polars, Soda-core, Athena, Tableau, Looker

DB/Datastore/Warehouse: Postgres, MongoDB, Redis, DuckDB, S3, Minio, Snowflake, Redshift, Apache Iceberg

Data Modeling & Warehousing: Dimensional Modeling, ERD Modeling, Data Warehouse Design

Backend: Flask, FastAPI, RabbitMQ, Traefik, Nginx

ML/AI: Jupyter, Scikit-learn, Numpy, Keras, Tensorflow, JAX, Flax, OpenAI, NLTK, Librosa, SciPy

EXPERIENCE

Senior Data Engineer

May 2025 - Present

SkywardOps (Fulltime)

Remote, Nigeria

- Architected custom data pipelines using Terraform IaC, AWS Glue, Apache Iceberg tables and implementing medallion architecture to deliver analytics-ready datasets for multiple business units.
- Optimized Snowflake operations by creating and maintaining stored procedures for daily automated data integrations, reducing warehouse costs by 30% through usage monitoring and compute optimization strategies.
- Refactored enterprise-scale dbt repository by integrating SQLFluff linting and implementing CI/CD best practices, modernizing data modeling standards across the entire central codebase for improved code quality and maintainability.
- Built data governance framework using AWS Glue Catalog and Crawlers for metadata management and lineage tracking.
- Ensured pipeline reliability by monitoring Jenkins, managing vulnerabilities, and setting up alerts to reduce downtime and meet SLAs.

Data Infrastructure Engineer

May 2024 - Apr 2025

Sabino LLC (Fulltime)

Remote, Nigeria

- Designed automated ETL pipelines using AWS Glue, Python, and dbt with Glue Crawlers for schema discovery across 15+ data sources, reducing processing time by 60% and orchestrating transformations into Snowflake.
- Built monitoring infrastructure with Prometheus, Grafana, and Loki for data pipelines, achieving 99.8% uptime across distributed systems.
- Engineered containerized data platform on AWS and OVH bare metal infrastructure using Docker, scaling data processing capacity with 40% cost optimization through efficient resource management.
- Managed mission-critical data infrastructure across Linux environments with CI/CD automation, maintaining 90% system availability and reducing deployment time from hours to minutes.
- Implemented proactive data observability solutions with custom Python monitoring frameworks, building multiple Grafana dashboards to track key metrics and integrating Slack alerts, decreasing mean time to resolution (MTTR) for data incidents by 75%

Data Engineer

Jan 2022 - Present

Upwork (Freelance)

Remote, Worldwide

- Delivered custom Python scripts to automate data extraction, processing, and delivery for 15+ clients, enabling businesses to meet critical deadlines and improve data-driven decision-making capabilities.
- Resolved data pipeline bottlenecks by re-architecting data flows utilizing Apache Kafka and Apache Spark, expediting data delivery to customers by 40% and enhancing user satisfaction scores.
- Optimized SQL queries to enhance data extraction and processing speeds, boosting system performance.

- Developed and optimized Data Pipelines with tools like Airbite, Airflow and DBT for ETL processes, ensuring scalable, automated data workflows.
- Created insightful data visualizations in Grafana & Google looker to support data-driven decision-making, monitoring key metrics and system performance.

Data Engineer

Mar. 2022 – Apr. 2024

Newswise (Contract)

Remote, US

- Architected robust ETL pipelines that process data from 50+ news sources with 89% accuracy, delivering data streams to critical internal systems and applications.
- Established and refined data ingestion pipelines that processed a high volume of news articles daily, significantly accelerating data retrieval by 40% for news feed applications.
- Successfully reduced ETL process runtime by 30% through optimization and parallel processing techniques.
- Enhanced data quality and consistency by implementing automated validation checks, resulting in a 25% reduction in data errors.
- Key role in scaling pipeline for 50% data volume increase, ensuring smooth ingestion and processing

Software Engineer (Data)

Nov 2021 - Dec 2023

Ocassio LTD (Contract)

Remote, Nigeria

- Standardized data exchange across 15+ services by architecting RESTful APIs using FastAPI, enhancing system interoperability and enabling real-time data synchronization capabilities.
- Streamlined data workflows, automating processing, transformation, and integration across systems, while optimizing information visualization through diverse source ingestion.
- Collaborated with ML & Analysts team, using AWS Redshift for data warehousing and DBT for modeling, achieving 15% better decision-making and 25% fewer errors.

Machine Learning Engineer/Mentor

Feb 2022 - Mar 2022

Technocolabs softwares (Contract)

Remote, India

- Delivered expert technical guidance to machine learning and data science interns, led cross-team collaboration on software solutions, and managed distributed teams, achieving high success rates and reducing deployment errors.
- Developed a data pipeline to collect and preprocess music data, leveraging spectral analysis and unsupervised learning to enhance dataset quality, reduce dimensionality, and improve feature extraction.

Machine Learning Engineer Intern

Nov 2021 - Feb 2022

Technocolabs softwares (Contract)

Remote, India

- Built a data pipeline with Python and APIs to collect and preprocess audio data from a music platform.
- Applied spectral analysis to enhance dataset quality and size by 70%.
- Used unsupervised learning to detect repeated patterns, reducing dimensionality while keeping key features.
- Collaborated on developing and evaluating ML models, including a neural network, for optimal performance.

EDUCATION

University of Uyo

Uyo, Nigeria

Bachelor of Engineering, Computer Engineering

2017 - 2022

EXTRACURRICULAR

Prefect-recipes repo contributor

Contributed to Prefect's recipe snippets and templates representing common Customer success patterns

Machine Learning Facilitator University of Uyo (MSc & Ph.D)

Provided expert guidance in machine learning, inspiring and empowering students to exceed expectations and achieve outstanding

Guest speaker Everest Online Academy (EOA) Bootcamp

Guided a class through immersive AI exercises, impacting an impressive 80% of participants.

Technical Article Author

<https://medium.com/@bestnyah7>

<https://medium.com/@bestnyah7>