

# NIROJ KOIRALA

[nirojkoirala80@gmail.com](mailto:nirojkoirala80@gmail.com) | [GITHUB](#) | [Website](#) | [LinkedIn](#) | 646-991-8577 | NY, USA



---

## Professional Summary:

Computer Science graduate (**Summa Cum Laude, GPA 3.89+**) with broad experience across software engineering, data engineering, and cloud systems. Proven ability to design, develop, and optimize scalable applications, **data pipelines**, and system-level solutions using **Python, SQL, Java, AWS, Azure**, and modern development tools. Strong foundation in algorithms, data structures, backend development, and real-world problem solving through hands-on projects and internships.

## Professional Experiences:

---

### Data/software Engineering Fellow - Takeo AI

Oct 2025 – Present

- ❖ Built scalable **ETL/ELT** pipelines using **Python, SQL, AWS, and Azure**, transforming large structured and semi-structured datasets into analytics-ready tables for reporting and analysis.
- ❖ Developed batch and event-driven data workflows with **Kafka and PySpark**, enabling efficient data ingestion, transformation, and standardization while improving pipeline scalability and latency.
- ❖ Built Automated data validation and deployment workflows using **Git and Azure DevOps**, ensuring reliable data delivery and reducing manual operational effort by **~40%**.
- ❖ Built backend services and **REST APIs** using Python and SQL, applying object-oriented design and modular, clean code principles.
- ❖ Based Implemented asynchronous, event-driven workflows with **AWS Lambda, Kafka**, and Azure services to improve scalability and fault tolerance.
- ❖ Developed and maintained analytical data models in Snowflake and SQL Server, supporting downstream reporting, population health analytics, and **BI** use cases.

Flushing, NY

### Data/software Engineer Intern - TIQC

May 2025 – Sept 2025

- ❖ Loaded transformed data into **Snowflake** using **Snowpipe** and **COPY** commands, enabling real-time data availability for downstream consumers.
- ❖ Assisted in optimizing data storage formats (**Parquet, CSV, Delta**) and partitioning strategies for faster performance and reduced costs across **AWS** and **Databricks** environments.
- ❖ Participated in b Improved backend performance by refactoring application logic and execution paths, reducing **API** response latency by approximately **30%**.
- ❖ Containerized backend services using **Docker** to standardize environments and improve deployment consistency across development and production.
- ❖ Worked with external clients to audit and build responsive websites using **HTML, CSS, JavaScript**, and frontend libraries.
- ❖ Worked with senior engineers to monitor job performance using **CloudWatch**, troubleshooting job failures and improving pipeline reliability.
- ❖ Designed and maintained **MongoDB** databases to support high-performance, schema-flexible data storage for semi-structured and unstructured data.

### CIRE Research Scholar | Software Engineering & Systems Research (IEEE Published, Award-Winning Project) -

Mar 2024 – Apr 2025

- ❖ Designed and coded control logic in **C++ (Arduino)**, enabling tight **hardware–software integration** across sensors, controllers, and signal modules.
- ❖ Implemented real-time traffic density detection and adaptive signal control **using sensor data** to optimize intersection flow.
- ❖ Built and tested emergency **vehicle detection and signal preemption**, reducing response time at intersections.
- ❖ Optimized event-driven algorithms for traffic density analysis and emergency vehicle signal preemption, reducing **false positives by 92%**.
- ❖ Designed, tested, documented, and finalized a fully functional IoT hardware prototype using Arduino, sensors, **XBee modules**, and automated boom barriers.

---

## Projects & Technical Work

## AWS Bronze–Silver–Gold COVID-19 Lakehouse Pipeline | [GITHUB](#)

- ❖ Designed and implemented an end-to-end AWS Lakehouse pipeline using the **Bronze–Silver–Gold architecture**, ingesting raw COVID-19 datasets into **Amazon S3** and transforming them into analytics-ready datasets.
- ❖ Developed **AWS Glue** PySpark ETL jobs to clean, standardize, and aggregate data across Silver and Gold layers, handling schema drift, missing values, and data normalization at scale.
- ❖ Optimized data processing performance using **partitioning strategies**, Parquet formats, and Spark tuning, improving query speed.
- ❖ Implemented **AWS IAM roles and policies** to securely manage access across **S3, Glue, Athena, and Redshift**, enabling **least-privilege access** and compliant data workflows.

## AI Chat Assistant (LLM-Powered Application) [GITHUB](#)

- ❖ Built an AI-powered assistant using **Python, Streamlit GUI, and Groq (Llama 3.1)** to support chat, document understanding, and data exploration.
- ❖ Developed interactive **Streamlit** interfaces for real-time chat, file uploads, and analytics visualization.
- ❖ Packaged and deployed the application as a cloud-hosted Streamlit web app, enabling real-time access and showcasing **end-to-end deployment** skills.
- ❖ Integrated **LLM inference** workflows using the Groq API (Llama 3.1), handling prompt construction, response streaming, and error handling for reliable **real-time AI interactions**.
- ❖ Generated automated insights from structured datasets, allowing non-technical users to explore data through a GUI-based interface.

## Personal Expense Tracker | [GITHUB](#)

- ❖ Built an end-to-end personal finance application using **Python, combining a Tkinter desktop GUI** with a Streamlit analytics dashboard.
- ❖ Designed data ingestion and storage pipelines using **CSV files**, enabling persistent expense tracking and structured data management.
- ❖ Built data visualizations (pie charts, bar charts, trend lines) using **Matplotlib** and Streamlit to help users identify spending patterns, high-cost categories, and recurring expenses.
- ❖ Implemented filtering and aggregation logic to **support budgeting, trend analysis**, and data-driven financial decisions.
- ❖ Structured the application using modular **Python components**, improving code readability, maintainability, and ease of future feature expansion.

## Education

---

York College, CUNY

Jamaica, NY

- Bachelor of Science in **Computer Science**      Summa Cum Laude | **GPA: 3.89+**
- Minor in **Business Administration**

### Relevant Coursework

*Data Structure and Algorithm, Computer Architecture, Application Development with Database, Object Oriented Programming/Design, Software Development with .NET, Web Programming I/II, Project Management*

### Honors and leadership

- ❖ **Mentor** – Provided academic guidance and one-on-one mentoring to undergraduate peer in STEM courses.
- ❖ **Phi Theta Kappa Honor Member** – Recognized for academic excellence, leadership, community service.
- ❖ **NSF Scholar (STEM)** – Selected for an NSF-funded scholarship based on academic merit and research potential.
- ❖ **CUNY CIRE Research Fellow** – Selected to lead a year-long research project, coordinating development, testing, and documentation; presented findings across multiple colleges throughout NYC.

### Certifications

---

AWS Certified Data Engineer – Associate  
Takeo Data Engineering Bootcamp Certificate  
AWS Certified Cloud Practitioner.  
Microsoft Azure Fundamentals  
IBM Data Engineering Professional Certificate