Azhar Izzannada Elbachtiar

✉ Contact : azharizzannadae@gmail.com / +62 812 2720 5540 (Phone)
in Website : linkedin.com/in/azharizz / bit.ly/webportfolio_azhar
🏠 Location : Indonesia (Ready to remote with different time zone)

Summary

In 2018, I began with a CS degree and pursued Google/AWS certifications, Intern. **By 2021**, I became a Data Engineer. Leveraged my **3 YOE** now to significantly modernize data management and foster innovative solutions.

Works Experiences

Data Engineer, **Moladin** [Full Time Work] [Remote]   *(November 2023 – Now)*

- Pioneered a cutting-edge denormalized JSON pipeline, monitoring, and anomaly detection system of query in **BigQuery**, achieving a **$400** USD monthly cost reduction.
- Orchestrated deployment strategies by integrating **Terraform**, **Kubernetes**, **Kustomize**, and **Flux** for data architecture in company achieving a **25%** reduction in deployment times, increase in reliability, and simplified management of sophisticated environments.
- Drove the adoption of **data contracts** and Avro schema-based validation processes, ensuring top-tier near **100%** data quality and alignment with business requirements by coordinating efforts across multiple teams.
- Innovated a comprehensive data discovery and cataloging solution with **DataHub**, ensuring **100%** data lineage, which drove operational analysis excellence and strategic data initiatives.
- Improved decryption into a high-performance **asynchronous** and **multithreaded** decryption system, reducing processing time from **24** hours to **1** hour, and delivering substantial performance gains.
- Redefined data access auditing with **Freshdesk**, securing **0** errors in customer data privacy access and enhancing regulatory compliance.
- Maintained automated data workflows with **Airflow** and **dbt**, ensuring seamless data delivery to **Metabase** and **Tableau**, enhancing decision-making capabilities.
- Implemented **Jenkins** for unit testing in **git** merges, optimized code quality and improved release accuracy.
- Managed and maintained **Kafka** streaming, **Debezium** and **Kafka UI** for efficient system management.

Data Engineer, **Lion Parcel** [Full Time Work]   *(August 2022 – October 2023)*

- Proactively optimized ELT data pipelines, leveraging **Docker, Airflow, AWS Redshift**, and improved performance by **30%** through the strategic migrating critical scripts to Spark.
- Integrated strategically leveraging **asynchronous** data fetching methods to optimize the retrieval of data from **Odoo, Salesforce, APIs**, and **15+** databases for enhanced efficiency.
- Built a robust **Flask** web application (**Python**) that replaced spreadsheets, integrating versioning and data validity to effectively mitigate human error incidents, ensuring **99%** enhanced data accuracy and integrity.
- Developed **FastAPI** to enable **s**eamless data sync and achieving real-time data availability for external users resulting increasing external user data availability and reducing data fetch time by **50%**.
- Utilized **SCD Type 2** strategies, leading to a significant **40%** reduction in users' complaints.
- Applied knowledge of business operations, leveraging data to influence **AWS Quicksight** reports or tables, and introduced **OpenMetadata** for comprehensive table metadata & business tracking across **500+** tables.
- Implemented monitoring and data governance for users on **Metabase** to Data Mart on **Redshift**, ensuring data integrity and compliance in a scalable and efficient manner.
- Launched an efficient email and Slack notification to alert teams of critical issues need immediate attention.

Data Analyst, **Gojek & Tokopedia Impact Foundation** [Trainee]   *(March 2022 – Jul 2022)*

- Engineered a cost-efficient web scraping solution with **Terraform** and **Selenium**, achieving a **55%** cost reduction and **4** times scraping speed for e-commerce.
- Leveraged Meta Prophet for supply & demand forecasting and visualized insights in **Apache Superset**.

Big Data Engineer, **PT Smartfren Telecom, Tbk** [Intern]   *(August 2021 – February 2022)*

- Teamed up with data scientists for maintain pipeline & analysis with data from more than **38+** million customers, ensuring stability using **Greenplum data warehouse** and **Hadoop**.
- Implemented an **end-to-end pipeline** for data collection using the **Twitter API** library for sentiment analysis and delivered it by creating an **API** using **FastAPI** for Machine Learning Model prediction. Additionally, improved the processing time of the pipeline by **70%**.
- Integrated pipeline for **Geospatial Data** (GIS) into the planning and budgeting company, resulting in a **34%** reduction in response time delivering the long-term value of the organization.

Education

Brawijaya University - Bachelor's Degree of Computer Science   *Malang (2018 – 2022)*

GPA Score: 3.80 / 4.00 (Graduated with Cum Laude in June 2022) (link)

- Interested in Data path with course work: Algorithm & Data Structures, Data Engineer, Calculus, Statistics, Big Data, Data Visualization, Object Oriented Programming, Machine Learning.

Language

- Indonesia (**Native**).
- English (**Professional Proficiency**, already experienced in Work Experience).

Skills & Certifications

Programming Language: Python | Java | SQL | HTML | Javascript | Shell

Database: AWS Redshift | Google BigQuery | MySQL | Postgresql | NoSQL (MongoDB) | Microsoft SQL Server | Elasticsearch

Data Engineer Tools: Airflow | Pentaho | SSIS | Apache (Spark, Hadoop, Kafka, Superset) | Flask | dbt | Datahub | Tableau | Streamlit | Debezium | Metabase | Selenium

Infrastructure & Other: AWS (Quicksight, EC2, S3) | Google Cloud (Cloud Function, Dataflow, Pub/Sub, GCS, Kubernetes Engine, Dataproc) | Snowflake | Kubernetes | Terraform | Git | Jenkins | Docker | FastAPI | CI/CD | Gitlab | Atlassian

Certification: Professional Google Cloud Architect | Professional Google Data Engineer | AWS Solution Architect Associate | AWS Cloud Fundamental | JuaraGCP Season 4 & 5 (Data Engineer, BigQuery) (link all)

Project

Speaker / Instructor for Pacmann Data Engineer Course on Fintech (link)

- Being the speaker / instructor of one week Data Engineer Starter Course in Pacmann, with total participant register is almost 50 persons. Make sure and solve participants problem while participate in Course.
- Led the creation of a project on credit scoring prediction with **FastAPI** and an automation pipeline using **Airflow**. Then, deployed it on **Docker** and **Streamlit** sharing so that it can be accessed by everyone.

Dimensional Modeling & Modernize Data Warehouse (Batch & Streaming) (link)

- Modernize and create a Data Warehouse for the Sales Team with the dataset that comes from the **Postgresql** database using a relational model. Then, perform **dimensional modeling** to convert it into a snowflake schema, aiming to improve the performance of **BigQuery** as the data warehouse.
- Using **Google Cloud Platform** for real-time clustering via a **streaming pipeline** (**Firestore**, **Functions**, **Pub/Sub**, **Dataflow**, **Dataproc** with **Spark**), and data transformation for the data mart using a **batch pipeline** (**Airflow**, **dbt**, and **Docker**), to solve the user's business problem questions.

Realtime Water Quality Prediction End-to-End Pipeline (link)

- Utilized **Google Cloud Platform** & **Docker** to get data sensor (**MongoDB**), Stream it into **Kafka** broker, and use **Python Spark** Model to predict and send the result into **Google Cloud Storage** then trigger it with **Cloud Function** for inserting to **BigQuery**.

Achievements & Others

- 1st Place Business Plan Competition (link)
- 2nd Place App Innovation Competition (link)
- 3rd Place Data Mining Competition (link)
- Data Scientist Trainee, PT. Algoritma Data Indonesia *(February 2022 – June 2022)*

Portfolio of my project will be detailed in web link (bit.ly/webportfolio_azhar)