Dewi Nurfitri Oktaviani

Surabaya, Indonesia | (62) 896-836-11730 | oktaviani.dewi@gmail.com
www.github.com/oktavianidewi | www.medium.com/@oktavianidewi |
https://www.linkedin.com/in/dewi-nurfitri-oktaviani-6b450b22

BRIEF INTRODUCTION

I am a Data Engineer with **over five years** of experience building reliable data pipelines and warehouse systems across cloud environments. Delivered end-to-end data solutions for analytics, real-time tracking, and internal automation. Skilled in designing efficient data models and optimizing large-scale processing systems with minimal oversight. Familiar with Generative AI and Large Language Models (LLMs), and actively exploring their application in data engineering workflows.

Presented technical talks at regional and international events, including Google I/O Extended Surabaya and Google Developers Space Singapore. Holds a Master's degree in Engineering from Waseda University, Japan, funded by the LPDP scholarship. **Passed JLPT N3 (July 2025) and preparing for N2 (target: December 2025)**.

EDUCATION HISTORY

Master of Engineering
Major: Data Engineering Lab
Research: User's Interest Predictability from User's Behavior on Social Network Services (SNSs)

Waseda University, Japan
March 2015 – March 2017

Bachelor of Arts
Major: Informatics Education
Research: The Development of Web-Based Lecture Journal Information System in Department of Electrical Engineering
Faculty of Engineering State University of Malang

Malang State University, Indonesia
July 2008 – July 2012

EMPLOYMENT HISTORY

Data Engineer, Tiptip	Indonesia
	May 2023 – April 2025

TipTip (www.tiptip.id) is a platform empowering creators and micro-communities to monetize digital content through subscriptions, live sessions, interactive events, and e-tickets.

Projects:
- **A marketing data collection tool using Selenium automation enhanced with AI**
  - Piloted the development of a web scraper using BeautifulSoup and Selenium to identify potential leads on social media (Instagram, Facebook Groups and Twitter), enabling sales teams to engage 100+ new communities monthly and increasing sales team efficiency.
  - Streamlined data collection by augmenting web scraper with generative AI for Event Organizer (EO) and Brands information detection across competitor platforms through texts and images yielding a database of 50+ potential leads in a month.

- **Staging deployment of data migration pipeline aimed at optimizing AWS infrastructure costs**
  - o Piloted data migration using in-house tools such as Debezium, Kafka, S3-writer and Iceberg, cutting potential AWS DMS expenses by 50%.
  - o Enhanced deployment strategy for in-house tools within AWS EKS staging by implementing iterative testing and configuration adjustments, achieving a 99.99% uptime.

Tech Stack:
Python, BeautifulSoup, Selenium, dbt, AWS S3, Redshift, Glue, ECR, EKS, DMS, MWAA, Apache Iceberg, PySpark, Debezium, Kafka, Helm, OpenAI APIs, Terraform, Docker, Kubernetes.

Data Engineer, GudangAda	Indonesia *February – December 2022*

GudangAda (www.gudangada.com): a B2B e-commerce platform for FMCG wholesale, manufacturers and retail businesses.

Projects:
- Deployed a robust campaign data ingestion framework, connecting Facebook Ads, TikTok Ads, and Google Ads to the data warehouse, empowering marketing analysts to generate insights faster.
- Delivered a Slack-based tool for internal data-warehouse transformations, providing customized notifications and alerts, which resulted in a decrease in data-related support tickets.

Tech Stack:
Golang, dbt, AWS RDS, Redshift, S3, Helm, Terraform, Docker, Kubernetes

Data Engineer, Kata.ai	Indonesia *February 2020 – February 2022*

Kata.ai (www.kata.ai) is an Indonesian conversational AI company focused on understanding human conversation that enables brands and enterprises to build their own intelligent chatbot.

Projects:
- Developed batch data pipelines leveraging Apache Airflow to automate the extraction, transformation, and loading (ETL) data daily, enabling comprehensive analytical reporting.
- Developed data streaming architecture leveraging Apache Kafka and Spark.
- Modernized data warehouse architecture with data aggregation layers, reducing query latency and improving the efficiency of data analytics for the front-end team.
- Maintained REST API services (Golang, NodeJs, Python) to provide the result from the data warehouse in REST API format to the Front-End team.
- Delivered a PoC Clickhouse data-warehouse experimentation by designing and implementing data transformation logic using PySpark, enabling seamless transition of data from Kafka to Clickhouse and from Hive to ClickHouse.

Tech Stack:
SQL, Python, Golang, Typescript, Apache Spark, Apache Cassandra, Apache Kudu, Apache Airflow, RabbitMQ, Apache Kafka, Clickhouse, Docker

Software Engineer, Kata.ai	Indonesia
	May 2017 – January 2020

Projects:
- Orchestrated the development and maintenance of command-line interface (CLI) tools for the Kata Platform using Node.js, resulting in a increase in developer workflow efficiency.
- Created back-end service to handover conversation from the bot to a human-agent
- Administered Jemma bot's scheduled content delivery, including weekly horoscope blasts and monthly astrological updates, ensuring zero downtime and 100% scheduled delivery for clients.
- Pioneered the development of Jemma's friend referral backend using Typescript, RabbitMQ, and PostgreSQL, resulting in 50% more user registrations and 25% higher engagement.

Tech Stack:
Typescript (Node.js), Redis, RabbitMQ, MySQL, Docker, Kubernetes

Web Developer, Wowrack Indonesia	Indonesia
	July 2012 – December 2014

Wowrack Indonesia (https://www.wowrack.co.id/en) is a cloud-hosting company providing the Cloud, Data Center, and Connectivity solutions.

Projects:
- Engineered a PHP (CodeIgniter) front-end for the Wowrack Billing Portal displaying API-provided data, facilitating monthly/yearly billing management for the billing department, and client billing notifications.
- Launched the architecture for the Wowrack Human Resource Management System, a portal which automated the tracking of 50+ employee events for payroll processing with PHP and MySQL.
- Developed a PHP-based (Laravel) Ticketing Report Portal integrating JIRA API, equipping system engineers with shift-based reports on open client tickets, leading to an improvement in ticket resolution SLAs.

LICENSES & CERTIFICATIONS

MLOps ZoomCamp   **DataTalksClub**
Skills: MLOps, Grafana, Airflow, Docker   *Issued Aug 2025*

Astronomer Certification for Apache Airflow 3   **Astronomer**
Fundamentals   *Issued Jul 2025*

Data Engineering ZoomCamp   **DataTalksClub**
Skills: Apache Spark, Kafka, dbt, Terraform, Docker, GCP,   *Issued Apr 2023*
Star Schema, Prefect

SKILLS

Data Platform	● SQL ● dbt (data build tool) ● Presto/Trino ● Apache Kafka ● Apache Spark ● Apache Cassandra	● Apache Airflow ● Metabase, Looker ● Postgres ● RabbitMQ ● Shell Script
Programming Language	● Golang ● Python	
Infrastructure tools	● CI/CD ● Helm ● Terraform ● Docker ● Kubernetes	
Cloud Environment	● AWS S3 ● AWS Redshift ● AWS RDS ● AWS Glue ● AWS ECR ● AWS EKS ● AWS DMS ● AWS MWAA	● Google Cloud Storage ● Google BigQuery ● Google Dataproc ● Google Pub/Sub

HONORS & AWARDS

Full-cost awards for graduate study at Waseda University Japan funded by The Indonesian Endowment Fund for Education (Lembaga Pengelola Dana Pendidikan / LPDP) from 03/2015 to 03/2017

SPEAKERS ENGAGEMENTS

- Speaker in Google IO Extended Surabaya (07/2022) with title: **Streaming Data to GCP with Pub/ Sub and BigQuery.** Repository is available on: https://github.com/oktavianidewi/stream_tweet_gcp
- Speaker at Google Developers Space, Singapore (09/2021), with title: **Building a Data Pipeline for Conversation Logs** (delivered in English, recording is available on: Building A Data Pipeline for Conversation Logs https://www.youtube.com/watch?v=mG8nGPRN_W4
- Seminar Informatika Aplikatif Politeknik Negeri Malang (08/2021) Inovasi Digital untuk Menyongsong Society 5.0 (Talk in Bahasa, recording available: SIAP 2021 https://www.youtube.com/watch?v=w-Oz_zqo7Ks)