



## KỸ SƯ DỮ LIỆU

NGÔ NGÀ QUANG

### MỤC TIÊU NGHỀ NGHIỆP

Tôi muốn làm việc với các nhóm Data Scientist và BI Analyst để đảm bảo dữ liệu được xử lý đúng định dạng, đúng ngữ nghĩa và cung cấp kịp thời cho phân tích chuyên sâu.

### THÔNG TIN CÁ NHÂN

16/02/1990

Hà Nội

thaolinh252512@gmail.com

0991872591

www.website.com

### HỌC VẤN

- Trí tuệ nhân tạo và Dữ liệu lớn tại  
Học viện Kỹ thuật Mật mã

### KỸ NĂNG

- dbt (Data Build Tool)
- AWS (S3, EMR, Glue)
- Apache Kafka
- Python
- Monitoring với Prometheus/Grafana

### KINH NGHIỆM LÀM VIỆC

- **CLOUD DATA ENGINEER** TẠI SMARTBANK (2022-2023)

+ THIẾT KẾ KIẾN TRÚC DỮ LIỆU TRÊN AWS VỚI S3, GLUE VÀ REDSHIFT

+ TỰ ĐỘNG HÓA VIỆC INGESTION LOG GIAO DỊCH NGÂN HÀNG TỪ HỆ THỐNG CORE

+ THAM GIA ĐÁNH GIÁ BẢO MẬT VÀ CHUẨN HÓA SCHEMA DỮ LIỆU TÀI CHÍNH

### DANH HIỆU VÀ GIẢI THƯỞNG

- **2022** - Top nhân viên cải tiến năng suất truy vấn dữ liệu SQL tại ReportPro

## SỞ THÍCH

- Tham gia cộng đồng lập trình
- Tập gym
- Sưu tầm sách lập trình

## NGƯỜI GIỚI THIỆU

- Ông Nguyễn Quốc Huy (Data Platform Lead - Công ty DataBridge) - huy.nguyen@databridge.vn - 0905123456
- Ông Lê Hoàng Dương (Head of Engineering - Công ty FinData) - duong.le@findata.vn - 0988999123

## CHỨNG CHỈ

- **2023** - Modern Data Engineering with dbt - dbt Labs
- **2020** - AWS Certified Data Analytics - Specialty

## HOẠT ĐỘNG

### - Thành viên nhóm phát triển hệ thống phân tích log tại Công ty SecureLog (2021)

- + Xây dựng pipeline thu thập và phân tích log từ nhiều nguồn hệ thống.
- + Triển khai hệ thống cảnh báo bất thường dựa trên mô hình thống kê.
- + Sử dụng ELK stack để trực quan hóa dữ liệu log.

## DỰ ÁN

### - Data Lake cho hệ thống quản trị khách hàng (CRM) (Data Engineer, CRMPro) 2022

Xây dựng nền tảng lưu trữ dữ liệu tập trung phục vụ phân tích hành vi khách hàng.

- + Tạo pipeline ingestion từ Salesforce, Google Ads và Facebook API
- + Lưu trữ dữ liệu theo mô hình phân vùng S3 Data Lake
- + Sử dụng Airflow để lập lịch và monitor luồng dữ liệu hàng ngày

### - Realtime Analytics cho hệ thống bán lẻ toàn quốc (Big Data Engineer, Retail360) 2023

Xây dựng hệ thống xử lý dữ liệu bán hàng theo thời gian thực

để hỗ trợ ra quyết định tức thì.

- + Sử dụng Kafka để thu thập dữ liệu từ các chi nhánh toàn quốc

- + Xử lý dữ liệu streaming bằng Spark Structured Streaming

- + Gửi dữ liệu về Redshift và hiển thị trên dashboard BI

#### **- Kiến trúc dữ liệu cho nền tảng giáo dục trực tuyến (Cloud Data Engineer, LearnHub) 2022**

Chuẩn hoá kiến trúc lưu trữ và xử lý dữ liệu học viên để phục vụ dashboard học tập.

- + Thiết kế hệ thống lưu trữ với BigQuery và Data Studio

- + Xây dựng Dataflow jobs để xử lý dữ liệu sự kiện học tập

- + Tạo luồng dữ liệu từ Firebase tới Google Cloud

