



## KỸ SƯ DỮ LIỆU

HOÀNG LAN NGÀ

### MỤC TIÊU NGHỀ NGHIỆP

Tôi định hướng phát triển các kỹ năng DevOps cho dữ liệu như CI/CD pipeline cho hệ thống ETL, quản lý version dữ liệu bằng DVC, và monitor hệ thống bằng Prometheus/Grafana.

### THÔNG TIN CÁ NHÂN

07/06/1986

Hà Nội

thaolinh252512@gmail.com

0389014360

www.website.com

### HỌC VẤN

- Công nghệ thông tin tại Đại học FPT

### KỸ NĂNG

- Snowflake

- Docker

### SỞ THÍCH

- Tập gym

### KINH NGHIỆM LÀM VIỆC

- **BIG DATA ENGINEER** TẠI LOGIAI (2020-2022)

+ XÂY DỰNG PIPELINE THU THẬP VÀ XỬ LÝ DỮ LIỆU VẬN CHUYỂN TỪ THIẾT BỊ IOT

+ SỬ DỤNG SPARK STREAMING ĐỂ XỬ LÝ DỮ LIỆU REAL-TIME

+ KẾT NỐI HỆ THỐNG VÀO ELASTICSEARCH VÀ PHÁT TRIỂN DASHBOARD GIÁM SÁT

- **JUNIOR DATA ENGINEER** TẠI RETAILTECH (2020-2021)

+ PHỐI HỢP XÂY DỰNG HỆ THỐNG PHÂN TÍCH TỶ

- Du lịch

KHO THEO THỜI GIAN THỰC

## NGƯỜI GIỚI THIỆU

- Bà Trần Thị Mai (Technical Manager - Công ty CloudData) - mai.tran@cloudata.vn - 0933666777

+ XỬ LÝ DỮ LIỆU STREAMING TỪ KAFKA VÀ LƯU TRỮ VÀO BIGQUERY

- Bà Nguyễn Thị Lệ (Analytics Lead - Công ty InsightTech) - le.nguyen@insighttech.vn - 0966999888

+ XÂY DỰNG CÁC VIEW LOGIC TRÊN DBT PHỤC VỤ BI DASHBOARD

- Ông Nguyễn Quốc Huy (Data Platform Lead - Công ty DataBridge) - huy.nguyen@databridge.vn - 0905123456

- **ETL DEVELOPER** TẠI FINDATA (2019-2021)

+ PHÁT TRIỂN VÀ BẢO TRÌ LUỒNG DỮ LIỆU ETL SỬ DỤNG TALEND VÀ PYTHON

- Bà Phạm Thanh Hương (Data Operations Manager - Công ty ReportPro) - huong.pham@reportpro.vn - 0912111222

+ VIẾT SCRIPT KIỂM TRA CHẤT LƯỢNG DỮ LIỆU VÀ CẢNH BÁO KHI DỮ LIỆU SAI LỆCH

- Ông Phan Thành Tâm (Big Data Architect - Công ty LogiData) - tam.phan@logidata.vn - 0955111222

+ TÍCH HỢP HỆ THỐNG PHÂN QUYỀN DỮ LIỆU THEO VAI TRÒ NGƯỜI DÙNG

- **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
- **2021** - Giải thưởng 'Dự án phân tích dữ liệu xuất sắc' tại FinData
- **2021** - Bằng khen về bảo mật và chất lượng dữ liệu – Công ty InfosecData
- **2020** - Top 5 kỹ sư xử lý dữ liệu lớn trong cuộc thi nội bộ AI Challenge

## CHỨNG CHỈ

- **2022** - Big Data Engineering with Apache Spark – edX (BerkeleyX)

## HOẠT ĐỘNG

- **Mentor khóa học nền tảng kỹ sư dữ liệu tại Trung tâm Đào tạo CloudTech (2023)**

+ Hướng dẫn học viên triển khai hệ thống ingest dữ liệu bằng Kafka.

+ Đánh giá bài tập về xử lý dữ liệu song song với Spark.

+ Tư vấn về định hướng nghề nghiệp cho sinh viên muốn theo ngành data engineering.

- **Diễn giả hội thảo 'Big Data Architecture' tại Data Talks Vietnam (2022)**

- + Trình bày kiến trúc hệ thống thu thập và xử lý dữ liệu đa nguồn.
- + Phân tích ưu nhược điểm của Data Warehouse vs Data Lakehouse.
- + Giới thiệu các công cụ phổ biến như Airflow, dbt, Snowflake.

## DỰ ÁN

### - Data pipeline xử lý log giao dịch hệ thống ngân hàng (Data Pipeline Engineer, BankDataTech) 2020

Lưu trữ, phân tích và phát hiện bất thường từ log giao dịch ATM và mobile banking.

- + Sử dụng Fluentd thu thập log và gửi về Kafka
- + Phân tích real-time qua Spark Streaming
- + Lưu dữ liệu đã xử lý vào Elasticsearch và trực quan hoá bằng Kibana

