



## KỸ SƯ DỮ LIỆU

PHẠM QUỲNH THỊ

### MỤC TIÊU NGHỀ NGHIỆP

Tôi đặt mục tiêu thành thạo việc xây dựng kiến trúc dữ liệu hiện đại trên nền tảng cloud như AWS, GCP hoặc Azure, đặc biệt tập trung vào việc tối ưu hóa chi phí và hiệu suất lưu trữ.

### THÔNG TIN CÁ NHÂN

23/03/1999

Hà Nội

thaolinhh252512@gmail.com

0373249068

www.website.com

### HỌC VẤN

- Kỹ thuật phần mềm tại Học viện Công nghệ Bưu chính Viễn thông

### KỸ NĂNG

- PostgreSQL

- dbt (Data Build Tool)

- Apache Spark

- AWS (S3, EMR, Glue)

### 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

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

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

## SỞ THÍCH

- Du lịch
- Viết blog kỹ thuật

## NGƯỜI GIỚI THIỆU

- Bà Trần Thị Mai (Technical Manager - Công ty CloudData) - mai.tran@cloudata.vn - 0933666777
- Ông Phan Thành Tâm (Big Data Architect - Công ty LogiData) - tam.phan@logidata.vn - 0955111222
- Ông Nguyễn Quốc Huy (Data Platform Lead - Công ty DataBridge) - huy.nguyen@databridge.vn - 0905123456

## 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

- **DATA ENGINEER** TẠI CÔNG TY DATAFLOW VIỆT NAM (2021-2023)

+ THIẾT KẾ VÀ TRIỂN KHAI PIPELINE ETL SỬ DỤNG APACHE AIRFLOW

+ TÍCH HỢP DỮ LIỆU TỪ NHIỀU NGUỒN NHƯ MYSQL, GOOGLE ANALYTICS VÀ API BÊN THỨ BA

+ TỐI ƯU HOÁ QUY TRÌNH XỬ LÝ BẰNG SPARK TRÊN GCP DATAPROC

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

+ PHỐI HỢP XÂY DỰNG HỆ THỐNG PHÂN TÍCH TỒN KHO THEO THỜI GIAN THỰC

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

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

## DANH HIỆU VÀ GIẢI THƯỞNG

- **2020** - Top 5 kỹ sư xử lý dữ liệu lớn trong cuộc thi nội bộ AI Challenge
- **2020** - Giải thưởng 'Pipeline ổn định nhất năm' tại TechFlow
- **2021** - Giải thưởng 'Dự án phân tích dữ liệu xuất sắc' tại FinData

## CHỨNG CHỈ

- **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

- **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