



KỸ SƯ DỮ LIỆU

VŨ TÙNG THÀNH

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

07/09/1983

Hà Nội

thaolinhh252512@gmail.com

0390941494

www.website.com

HỌC VẤN

- Hệ thống thông tin quản lý tại Đại học Bách Khoa Hà Nội - Kỹ thuật phần mềm tại Học viện Công nghệ Bưu chính Viễn thông

KỸ NĂNG

- Linux

- Data Quality Validation (Great Expectations)

- Python

- Data Lake Architecture

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

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

+ XÂY DỰNG PIPELINE THU THẬP VÀ XỬ LÝ DỮ LIỆU

VẬN CHUYỂN TỪ THIẾT BỊ IOT

SỞ THÍCH

- Đi bộ đường dài

NGƯỜI GIỚI THIỆU

- Bà Đỗ Ngọc Trang (Senior Data Engineer - Công ty SmartRetail) - trang.do@smartretail.vn - 0922333444

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

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

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

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

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

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

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

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

DANH HIỆU VÀ GIẢI THƯỞNG

- **2020** - Giải thưởng 'Pipeline ổn định nhất năm' tại TechFlow

- **2020** - Top 5 kỹ sư xử lý dữ liệu lớn trong cuộc thi nội bộ AI Challenge

- **2022** - Top 3 nhân viên cải tiến hệ thống ETL tại SmartRetail

CHỨNG CHỈ

- **2021** - IBM Data Engineering Professional Certificate – Coursera
- **2023** - Modern Data Engineering with dbt – dbt Labs
- **2020** - AWS Certified Data Analytics – Specialty

HOẠT ĐỘNG

- Thành viên nhóm xây dựng pipeline dữ liệu tại Phòng Lab Dữ liệu - Đại học Bách khoa (2021 - 2023)

- + Phát triển luồng dữ liệu ETL sử dụng Apache Airflow.
- + Thiết lập hệ thống ingestion dữ liệu từ API và cơ sở dữ liệu nguồn.
- + Phối hợp với nhóm phân tích dữ liệu để tối ưu schema lưu trữ.

- Thành viên nhóm kỹ thuật dữ liệu tại Dự án E-Government (2023)

- + Thiết kế hệ thống thu thập dữ liệu hành chính từ nhiều bộ ngành.
- + Thiết lập hệ thống kiểm tra chất lượng dữ liệu tự động.
- + Triển khai pipeline đồng bộ dữ liệu hằng ngày với độ trễ thấp.

- 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.

DỰ ÁN

- 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

