



## KỸ SƯ DỮ LIỆU

### TRẦN Tú TRÍ

## MỤC TIÊU NGHỀ NGHIỆP

Tôi mong muốn đóng góp vào việc xây dựng hệ thống phân quyền và bảo mật dữ liệu mạnh mẽ, tuân thủ các tiêu chuẩn như GDPR, ISO 27001 trong vận hành hệ thống dữ liệu.

## THÔNG TIN CÁ NHÂN

16/09/1982

Hà Nội

thaolinh252512@gmail.com

0840293778

www.website.com

## HỌC VẤN

- Công nghệ thông tin tại Đại học FPT
- Khoa học máy tính tại Đại học Quốc tế - ĐHQG TP.HCM

## KỸ NĂNG

- Apache Spark

## SỞ THÍCH

- Chụp ảnh

## 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

## DANH HIỆU VÀ GIẢI THƯỞNG

- **2023** - Bằng khen về tối ưu chi phí lưu trữ dữ liệu cloud - Công ty CloudData

- Tham gia cộng đồng lập trình
- Thử nghiệm công nghệ mới
- Du lịch

## NGƯỜI GIỚI THIỆU

- Bà Trần Thị Mai (Technical Manager - Công ty CloudData) - mai.tran@cloudata.vn - 0933666777
- Ông Vũ Minh Đức (CTO - Công ty DataStreamX) - duc.vu@datastreamx.vn - 0977333444
- Ông Lê Hoàng Dương (Head of Engineering - Công ty FinData) - duong.le@findata.vn - 0988999123

- **2023** - Giải thưởng 'Kiến trúc dữ liệu sáng tạo' do CTO trao tặng tại DataWare
- **2020** - Top 5 kỹ sư xử lý dữ liệu lớn trong cuộc thi nội bộ AI Challenge
- **2021** - Giải thưởng 'Dự án phân tích dữ liệu xuất sắc' tại FinData

## CHỨNG CHỈ

- **2022** - ETL and Data Pipelines with Shell, Airflow and Kafka - Coursera
- **2021** - SQL for Data Engineering - Datacamp

## HOẠT ĐỘNG

### - **Tình nguyện viên hỗ trợ nghiên cứu học thuật tại Khoa Công nghệ Thông tin (2020)**

- + Chuẩn hóa dữ liệu nghiên cứu sinh viên từ nhiều năm học.
- + Phân tích và tổng hợp xu hướng đăng ký môn học.
- + Viết báo cáo kỹ thuật trình bày cho giảng viên hướng dẫn.

### - **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

### - 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

