

Vũ Hữu Sỹ - Data Engineer

Bắc Từ Liêm, Hà Nội

☎ 0865 294 163 | ✉ vuhuusyft1@gmail.com

[LinkedIn](#) | [in](#) | [Github](#) | [HackerRank](#)

Học vấn

Đại Học Bách Khoa Hà Nội

Chuyên ngành: Hệ thống thông tin quản lý

8/2021 - Hiện tại

- CPA: 3.6/4.0
- Giành HBKKHT kỳ 2021.1 với GPA 3.9/4.0
- Môn học: Cơ sở dữ liệu (A+), Cấu trúc dữ liệu và giải thuật (A+), Kho dữ liệu và kinh doanh thông minh (A+), Kỹ thuật lập trình (A)
- TOEIC 690

Skills

Python: pandas, matplotlib, numpy, BeautifulSoup4, re, os, PySpark

Database: Oracle, MySQL, PostgreSQL

Tools: Oracle Data Integrator (ODI), PowerBI

Big Data: Hadoop, Spark, Kafka

Orchestration: Airflow

NoSQL database: MongoDB

Kinh nghiệm làm việc

Data Engineer

11/2023 – 2/2024

Công ty TNHH Giải pháp và Phân tích dữ liệu Insight Data

47 Đ. Nguyễn Tuân, Hà Nội

Vai trò:

Part Time

- Nghiên cứu và học các kỹ thuật triển khai Data Warehouse
- Học và áp dụng công cụ ODI trong triển khai Data Warehouse cho doanh nghiệp
- Sử dụng PowerBI để tạo báo cáo tự động đưa ra insight từ dữ liệu được cắm từ Data Warehouse cho các bên liên quan

Chứng chỉ và Giải thưởng

SQL (Advanced) Certificate - HackerRank

12/2023

- Giải các câu hỏi về các chủ đề như tối ưu hóa truy vấn, mô hình hóa dữ liệu, tạo index, Window functions

Quán quân của cuộc thi SQL CHAMPIONSHIP SEASON 2 - INDA

12/2023

- Viết các câu truy vấn phức tạp, cần sử dụng Window functions để giải các câu hỏi khó do ban tổ chức đưa ra

Project cá nhân

Transportation Data Warehouse | Tools: ODI, Oracle Database, PowerBI 11/2023 - 1/2024

- Đẩy dữ liệu từ nguồn về Staging (Load 1:1). Các cơ chế đẩy dữ liệu trong Data Warehouse: Control Append (Truncate, Insert), Incremental Update, SCD Type 2
- Đẩy dữ liệu từ Staging về Data Warehouse được tổ chức theo mô hình Star Schema
- Lưu vết dữ liệu lịch sử thay đổi các bảng Dimension bằng cơ chế IKM Slowly Changing Dimensions (SCD Type 2)
- Từ Data Warehouse chia ra các Data Mart theo từng phòng ban, kết nối PowerBI để phân tích báo cáo
- Tạo Load Plan tổng chạy đa luồng để tối ưu hóa hiệu năng hệ thống
- Link: <https://github.com/vuhuusy/transportation-data-warehouse>

Chứng nhận

IBM Data Engineering - Coursera: 12/2023

NoSQL, Big Data, and Spark Foundations - Coursera : 1/2024