

TRẦN NGỌC TUẤN

Data Engineer

Ho Chi Minh City • (+84) 976 165 259 • tntuan0910.work@gmail.com •
[linkedin.com/in/tntuan0910](https://www.linkedin.com/in/tntuan0910) • github.com/tuantran0910

OBJECTIVES

I am seeking an internship or fresher opportunity in an organization that offers me a chance to learn, work, and contribute to the company's success.

EDUCATION

University of Science - Vietnam National University, Ho Chi Minh City Bachelor of Data Science	Expected May. 2025 Ho Chi Minh City
<ul style="list-style-type: none">Cummulative GPA: 8.94/10.00Relevant Coursework: Algorithms; Statistics; Artificial Intelligence; Database ManagementReceived scholarships for being in the top 5% of the class for three out of four semesters.	
VStep - B2 English Certificate	August. 2024

EXPERIENCES

Software Engineer Intern Tiki	June. 2024 – Present Ho Chi Minh City
<ul style="list-style-type: none">Design basic ETL pipeline that ingested transactional data, process and load them to data warehouse for analytical purpose.Integrate Lambda Architecture, combine batch and stream processing for data processing.Project: Rainbow Data Platform.	

PROJECTS

Rainbow Data Platform <i>Visit: Docs</i> <i>Learning by building a Data Engineer Platform</i>	June. 2024 – Present Tiki Self Learning Internal Project
<ul style="list-style-type: none">Integrate lambda architecture for realtime data processing system.Conduct ETL pipeline for loading data from source to data warehouse in batch layer for historical analyzing. Integrate Change Data Capture with Kafka for data streaming to derive real time view in speed layer.Project deployment onto Google Cloud Platform.Technologies: Golang, Python, PostgreSQL, Dagster, Dbt, BigQuery, Kafka, Clickhouse, Superset, Docker, Kubernetes.	
Spotify Analysis <i>Source: Github</i> <i>Analyzing Spotify Data with Pyspark and ETL Procedures</i>	December. 2023 Final Semester Project
<ul style="list-style-type: none">Process 200,000 records from the Spotify API through ETL procedures using Hadoop and Spark. Further analyze the data with Dremio and generate reports using PowerBI.Design and implement an automated pipeline system. Conduct batch processing Spotify data.Technologies: Python, Prefect, Docker, MongoDB, Hadoop, Apache Spark, Dremio, PowerBI.	
Virtual Try-on <i>Source: Github</i> <i>Creating interactive chatbot for convenient clothing fitting</i>	December. 2023 Vietnam Datathon Project
<ul style="list-style-type: none">Conduct web scraping using Selenium to collect additional image data and product information from the Adidas website. Integrate the DM-VTON pre-trained model to enable virtual try-on functionality.Develop a chatbot leveraging OpenAI Api, LangChain and Flask frameworks to enhance user interaction.Technologies: Python, PyTorch, Selenium, LangChain and Flask.	

AWARDS & ACHIEVEMENTS

3RD PLACE WINNER - VIETNAM DATATHON 2023 <i>Virtual Try-on Project</i>	December. 2023 Competition
--	-------------------------------