



# DANG NHAT

## DATA ENGINEER

- 12/04/1995
- Go Vap District, HCMC
- +84 0354883673
- dangnhatsimon@gmail.com

## Profile

- LinkedIn
- Github
- LeetCode

## Education

### Obninsk Institute for Nuclear Power Engineering

B.Sc in Nuclear Power Engineering and Thermal Physics. (2014-2019)  
GPA: 4.6/5.0

### MongoDB University

MongoDB Python Developer. (2024)

### IBM Skills Network

Data Engineering. (2022)

### DataCamp

Professional Data Engineer. (2024)

Big Data with PySpark. (2024)

### Google Career

IT Automation with Python. (2021)

Data Analytics. (2021)

## Working Experience

May 2024 – Present

### Data Engineer Specialist

GREENFEED Vietnam Corporation - Fulltime

- GREENFEED is a fully integrated clean food chain FEED – FARM – FOOD with the mission to build quality and trusted brands that relentlessly innovate and create sustainable values for customers and society.
- ELT Processes: designed, built, and tested ELT workflows using Docker, integrating data from diverse sources (SAP, Solomon, MISA, Porcitech)
- Data Pipelines: developed and monitored robust batching and streaming pipelines with Airbyte, Airflow, Beam with Dataflow, Spark Streaming and Spark Structured Data.
- Data Migration: managed migration and synchronization of data from internal systems to AWS (S3, Redshift).
- Data Lakehouse: built, designed from scratch and validated a AWS-based lakehouse for domain Farm, Social Media, Supply Chain and Financial & Accounting.
- Data Transformation: leveraged DBT, Python, and Java to deliver clean, structured datasets for analytics.
- Social Media: built and managed data pipelines from social media platforms and specific platforms APIs, to Lakehouse for harnessing external data sources for deeper insights.
- Documentation & Collaboration: maintained thorough documentation and partnered with stakeholders to deliver aligned, timely data solutions across FEED, FARM, and FOOD.
- Skills: Python, Java, AWS, GCP, Spark, Beam, Kafka, Data Architecture, Docker, Kubernetes, GitLab, Integrated Businesses in Closed Loop.

Apr 2024 – Apr 2024

### Back End Engineer

Definer Inc. - Remote Parttime

- Definer designs and develops solutions, builds DevOps, automates cloud IT infrastructure construction, and builds analytical platforms for big data.
- Solution Design & DevOps: architected and implemented comprehensive solutions with a focus on automating cloud IT infrastructure and workflows.
- Database Interaction: leveraged Prisma for type-safe, auto-generated database queries.
- AWS & CI/CD: developed task definitions and configuration files (CircleCI, Docker Compose, Dockerfile) to automate build, test, and deployment processes on AWS.
- Cloud Data Processing: automate data processing from social media on platforms like Google BigQuery, AWS S3, Lambda, and Fargate using Python scripts.
- Web Development: built scalable websites using Flask and REST APIs, aligning with business objectives.
- Big Data Platforms: designed and managed platforms to enable processing and analysis of large datasets.
- Skills: AWS, DevOps, Docker Swarm, Prisma ORM, Microservices, Flask, Back-End Web Development, SQL, Python, REST APIs, CircleCI, GCP.

# DANG NHAT

## DATA ENGINEER

### Certificate

- IBM - Big Data 101.
- IBM - Hadoop 101.
- DataCamp - Building Data Pipelines.
- freeCodeCamp - Data Analysis with Python.
- Hackerrank - Problem Solving Basic.
- Hackerrank - Python Basic.
- Hackerrank - SQL Advanced.
- Sololearn - Python Developer.
- Sololearn - Python Intermediate.
- Sololearn - SQL Intermediate.

### Languages

- Vietnamese
- English
- Russian
- Chinese

Oct 2022 –  
May 2024

#### Data Engineer

- PECC2 - Fulltime
- PECC2 is a leading consultancy in the electricity energy industry, specializing in providing comprehensive solutions for power generation, transmission, and distribution projects. PECC2 is renowned for its expertise in engineering design, project management, and innovative energy solutions, ensuring sustainable and efficient power systems for clients.
  - Data Ingestion & Processing: managed ingestion pipelines processing 1GB/day from power plant monitoring systems, ensuring data availability for Data Scientist and ML teams.
  - Big Data Development: contributed to ETL processes and data pipelines, architected and deployed data warehouses, and delivering interactive reports to enhance customer experience.
  - Web Scraping & API: extracted and processed data from dynamic websites with pagination, leveraging APIs for real-time insights.
  - Database Management: designed, created, and optimized relational databases using advanced SQL techniques (views, transactions, stored procedures, joins).
  - Scripting & Automation: automated workflows with Linux shell scripts, Python, and cron jobs—tasks include catalog updates, supplier data fetching, and document processing.
  - NoSQL & Big Data: applied NoSQL and Big Data tools (e.g., Spark, Kafka, Beam) to manage and analyze large datasets efficiently.
  - Custom Tool & Monitoring: developed internal tools, such as a Telegram bot for real-time data, keyword search GUIs, and system health monitoring scripts. Automated sales data reporting and email PDF generation for operational efficiency.
  - Skills: Data Engineering, ETL & Data Pipelines, RDBMS, NoSQL and Big Data, Spark, Kafka, Beam, Python, SQL, Scripting, Docker, Git, Cloud.

### Skills

Domain	Electricity, Agriculture, F&B, Inventory.
Foundation	Data Structures & Algorithms, Probability & Statistics.
Programming	Python, Java, dbt, R, LaTeX.
Framework	Flask, Django, Spring, Data Engineering.
Scripting	Bash, Shell on Linux, Git.
Database	SQL(MySQL, PostgreSQL, MSSQL), NoSQL(MongoDB, Cassandra), SAP.
Big Data	Spark, Kafka, Beam, Hadoop.
Cloud	AWS, GCP, IBM, Azure.
DevOps	Docker, Kubernetes.
Visualization	Tableau, Power BI.
Soft Skills	Detail-oriented, Structured & Analytical Thinking, Problem Solving, Teamwork, Diligent, Energetic, Resilient.