# DANG NHAT

## DATA ENGINEER

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

## Profile ────────────

- 🔗 LinkedIn
- 🔗 Github

## 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**
Data Engineer (2022).
Associate Data Engineer (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 Process: design, build, and test ELT processes using Docker, integrating data from diverse sources such as SAP, Solomon, MISA, and Porcitec software.
- Data Pipeline: develop and monitor robust data pipelines utilizing Airbyte and Airflow.
- Data Migration & Synchronization: manage data migrations and synchronization from internal servers and databases to AWS services like S3 and Redshift.
- Data Warehouse: contribute to the development and validation of a new data warehouse in domain Supply Chain and Financial & Accounting on Redshift, designed from scratch.
- Data Transformation: leverage DBT and Python, Java for processing and transforming data, delivering cleansed and organized datasets for data analysts.
- Social Media: build and manage data pipelines from social media platforms like Facebook, Google, and LinkedIn APIs, as well as the PICTraq API, to Redshift for harnessing external data sources for deeper insights.
- Documentation: maintain comprehensive documentation for all data processes, pipelines, and architectures.
- Collaboration: work closely with data analysts and other stakeholders to understand their data needs and deliver accurate, timely data solutions, ensure that data transformations across the FEED, FARM, and FOOD segments are aligned with business goals.
- Skills: Python, Java, AWS, 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: design and develop comprehensive solutions, focusing on building and automating cloud IT infrastructure.
- Database Interaction: utilize Prisma as an interface to interact with databases, leveraging its type-safe and auto-generated query builder.
- AWS & CI/CD: create task definitions on AWS and build configuration files such as CircleCI, docker-compose, and Dockerfile to automate the build, test, and deployment processes, streamline the lifecycle of application development and deployment. - Cloud Data: process data gathered from social media on cloud platforms and develop Python scripts to automate tasks on Google BigQuery, AWS S3, AWS Lambda, and AWS Fargate.
- Web Development: use Flask and REST API to build robust and scalable websites that meet user requirements and align with business objectives. - Big Data Platforms: build and manage analytical platforms for big data, enabling the processing and analysis of vast 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 |

- Data Ingestion & Processing: manage the ingestion of data from various monitoring systems at power plants, processing data at a rate of 1GB/day, ensure that data is always available for the company's Data Scientist and ML Teams, enabling them to perform timely and accurate analyses.
- Big Data: contribute to the company's Big Data Development Plan, focusing on the implementation of ETL processes and data pipelines, includes architecting, populating, and deploying data warehouses, as well as contributing to interactive reports that are streamed to end users to enhance customer experience.
- Web Scraping & API: perform web scraping and data extraction, efficiently handling dynamic websites and pagination.
- Database Management: design, create, and manage relational databases, applying advanced SQL techniques such as views, transactions, stored procedures, and joins.
- Shell Scripting & Automation: develop shell scripts using Linux commands and schedule cron jobs to automate various tasks, ensuring system efficiency and reliability. Automating catalog updates, fetching supplier data, processing text files, and creating and modifying documents and messages with Python.
- NoSQL & Big Data: develop a working knowledge of NoSQL and Big Data technologies, manage and analyze large datasets effectively.
- Custom Tool: create a Telegram bot that interacts with the API to provide real-time data. Developing search functions and simple GUIs to support internal operations, including keyword search across multiple files and database interactions.
- System Monitoring & Reporting: write scripts to check the health status of systems, ensuring that all operations run smoothly, automate the summarization and processing of sales data, generating PDF reports that are distributed via email, enhancing reporting efficiency.
- Skills: Data Science, ETL & Data Pipelines, RDBMS, NoSQL and Big Data, Apache Spark, Python Programming, Data Analysis, SQL, Shell 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 | Hadoop, Spark, Kafka. |
| Cloud | AWS, Azure, GCP, IBM. |
| DevOps | Docker. |
| Visualization | Tableau, Power BI. |
| Soft Skills | Detail-oriented, Structured & Analytical Thinking, Problem Solving, Teamwork, Diligent, Energentic, Resilient. |