

NGUYEN DANG HONG HUY (Henry)

Data & Analytics Engineer | Data Platform Engineer

✉ ndhhuy1320@gmail.com

📍 Hochiminh City

☎ 0905263087

🌐 ndhhuy2k

📷 ndhhuy2k

SUMMARY

Senior Data Engineer with 5+ years of experience designing scalable Data Platforms and high-volume data pipelines across Banking and Fintech. Specialized in Lakehouse/Medallion architecture, distributed processing, real-time data, and modern DataOps using Spark, Delta Lake, Iceberg, Airflow, Kafka, dbt, and cloud object storage. Experienced in building governed data products, semantic layers, observability, and regulatory data solutions for e-wallet, payment gateway, lending, credit scoring, and financial reporting. Strong focus on platform architecture, data quality, governance, security, performance optimization, and translating business requirements into reliable enterprise data solutions.

WORK EXPERIENCE

📌 Galaxy Pay – Data Engineer

📅 Feb 2026 – Present | Ho Chi Minh City

- **Architected an enterprise Medallion Data Platform (Bronze–Silver–Gold)** using S3, Apache Spark, Delta Lake, Apache Iceberg, Airflow, dbt, and ClickHouse to support high-volume data across **e-wallet, payment gateway, and financial services**.
- Designed scalable **batch and near-real-time data pipelines** with standardized ingestion, transformation, data quality checks, orchestration, and SLA monitoring for critical payment and transaction data.
- Established **DataOps and Data Governance practices** covering data quality, metadata, lineage, access control, CI/CD, standardized business definitions, and governed data products.
- Designed a **Semantic Layer** to standardize business metrics and definitions across analytics, risk monitoring, operational reporting, and downstream applications.
- Built a **Data Observability platform using FastAPI and ReactJS** for pipeline health, SLA tracking, data quality monitoring, operational dashboards, and proactive incident detection.
- Supported **SBV regulatory reporting, risk monitoring, and compliance data requirements**, ensuring governed and traceable data delivery across business functions.
- Collaborated with **Risk, Compliance, Business, and Engineering teams** to translate domain requirements into scalable data products and reusable platform capabilities.

📌 HDBank – Data Engineer

📅 Mar 2024 – Jan 2026 | Ho Chi Minh City

- Designed and implemented enterprise **ETL/ELT pipelines for Data Warehouse, BI, and regulatory reporting**, processing large-scale banking data across **lending, credit scoring, customer analytics, and risk** domains.
- Built a configurable **Report Builder platform** supporting on-demand Excel/PDF report generation, storage, caching, and report delivery without application redeployment.
- Modernized legacy data processing by migrating resource-intensive **SQL procedures to Apache Spark jobs**, achieving **30–70% processing performance improvement** depending on workload.
- Developed scalable data processing solutions for **SBV regulatory reporting, lending analytics, credit scoring, and risk-related data products**, improving consistency and reliability of downstream data consumption.
- Worked across Data Engineering, BI, and business teams to translate banking requirements into reusable data pipelines, reporting capabilities, and governed data assets.

📌 Fujinet Systems – Data Engineer & Full-stack Developer

 Oct 2021 – Feb 2024 | Ho Chi Minh City

- Built a real-time sales analytics platform using **Kafka, Spark Streaming, and Cassandra**, with batch processing pipelines on Hadoop/HDFS for scalable data processing.
- Designed streaming pipelines achieving **3–5s end-to-end latency**, enabling near-real-time sales visibility for business teams.
- Developed supporting applications using **ReactJS, FastAPI, NodeJS, and PostgreSQL**, providing end-to-end understanding across data, backend, and frontend systems.
- Collaborated in Agile/SCRUM teams of 2–5 members with developers, QA engineers, and business stakeholders to deliver data-driven solutions.

TECHNICAL SKILLS

- **Data Platform & Architecture:** Lakehouse, Medallion Architecture, Data Platform Design, Data Governance, DataOps, Semantic Layer, Data Quality, Data Observability, RBAC/IAM
- **Big Data & Processing:** Apache Spark, PySpark, Spark Streaming, Kafka, dbt, Airflow, Databricks, Delta Lake, Apache Iceberg, Trino, CDC/Debezium
- **Storage & Data Warehouse:** Amazon S3, MinIO, Snowflake, BigQuery, ClickHouse, Oracle, PostgreSQL, Cassandra
- **Programming:** Python, SQL, JavaScript
- **Cloud & DevOps:** AWS, Azure, Docker, Kubernetes, GitHub Actions, CI/CD
- **Data Management & Metadata:** DataHub, OpenMetadata, Data Lineage, Metadata Management
- **AI-Assisted Engineering:** Cursor, GitHub Copilot, Prompt Engineering
- **Domain:** Banking, Fintech, E-wallet, Lending, Credit Scoring, Financial Reporting, Risk Monitoring

PROJECTS

"Learn DE via AI" - Personal Brand & Mentorship

- **Description:** Developed a technical content channel Tiktok ([Học DE bằng AI](#)) focused on Data Engineering, breaking down complex concepts (Spark, Airflow, dbt, Lakehouse) into systematic, bite-sized videos (60-120s).
- **Impact:** Created structured learning frameworks, translating deep technical knowledge into accessible content, showcasing strong communication skills and the ability to explain "why" and "how" behind Data Platforms.

EDUCATION

- **Master Computer Science - HCMUS**
 2024 - Ongoing | **GPA:** 8.20
- **B.S. Information Systems - HCMUS**
 2018 - 2022 | **GPA:** 8.33

SOFT SKILLS

Problem-solving · Tech-to-Business Communication · Leadership & Mentoring · Goal-oriented

CERTIFICATIONS

★ Databricks

- AWS & Azure Databricks Platform Architect - Issued Oct 2025
- Databricks with AI Agent Fundamentals - Issued Oct 2024

★ Google Certificates

- Google Data Analytics Specialization - Issued Jan 2023