

Phuong Duy Nguyen (Paul)

Mobile: 0493685977

Duy.paul.1108@gmail.com

SKILLS

Data Engineering Tools: Python (Pandas, Numpy, scikit learn, Keras, Flask REST API, etc.), Spark, Airflow, Scala, SQL

Database and Cloud: AWS, GCP, Azure, SQL Server (T-SQL, SSIS, SSAS and SSRS), MySQL, DBT Cloud

Data Analysis and Data Visualisation Tools: Alteryx, Tableau, Power BI, SQL, SAS, Excel Macro (VBA), MS Access

Others: SAP, Salesforce, Snowflake, Databricks, Jira, Terraform, Confluence, Github, Gitlab, MYOB, XERO, PowerApps, SharePoint

Languages: Vietnamese (native) and English

Transferable Skills:

- Strong quantitative, analytical and problem-solving skills.
- Excellent attention to detail with diligent.
- Keen desire to learn and want to continuously improve.
- Focused on teamwork and building collaborative relationships.
- Highly organised, with the ability to work to deadlines, multi-task, prioritise and re-prioritise workload.
- Proactive, with the ability to work effectively within a fast-paced and high-pressure environment.
- Quickly adapt to cope with changing products, processes and procedures.

WORK EXPERIENCE

Jung Talents

Sydney

Data Engineer

Feb 2024 – Present

- Collaborated with software engineers, data scientists and other stakeholders to understand requirements, design solutions, and contributed to the overall design and architecture of AI-powered applications.
- Developed and managed scalable cloud infrastructure using AWS and Azure to ensure optimal performance and cost-efficiency.
- Strengthened data security by configuring IAM roles and implementing encryption with Cloud KMS to ensure compliance and safeguarding sensitive information.
- Built and deployed data lakes and data warehouses on GCP using services including BigQuery, Dataflow, DataFusion, Composer (Airflow), CloudSQL, Cloud Functions and GCS.
- Imported data from a variety of sources such as Databases, APIs, streaming data feeds.
- Implemented Apache Airflow via Amazon MWAA to orchestrate data migration and optimise ETL workflows to significantly improve data processing efficiency and reduce manual processing using Spark and Hadoop ecosystems.
- Employed Pub/Sub model to decouple services and balance workloads.
- Utilised AWS Glue to execute the data transformation in ETL pipelines using SQL to reduce time spent for ETL documentation and ensuring data validation.
- Ensured data engineering solutions are compliant with standard security practices and regulatory requirements.
- Implemented data protection measures and disaster recovery solutions.
- Used Terraform and version control in Github Action to automate infrastructure scaling and management.
- Set up monitoring in AWS CloudWatch to track system performance, identify and resolve issues, and ensure efficient data processing.
- Leveraged Apache Spark for real-time data streaming and analytics.
- Built interactive visualisations with Power BI for dynamic data visualisations.

KPMG

Sydney

Data Engineer

Nov 2023 – Feb 2024

- Collaborated with cross-functional teams and other stakeholders such as product manager and business analysts to gather requirements for data model development.
- Utilised Azure Data Factory to gather data from multiple resources and created associated storage containers.
- Set up Azure Data Factory Resource and Synapse Analytics workspace to manage and store data in blob storage.
- Built external table that references dirstaging directory for staging payroll data and developed datasets for data transformation and aggregation.
- Developed code-based ETL/ELT data pipeline with performance optimised data modelling to reduce manual data processing time.
- Prepared datasets for payroll file in Azure Data Lake Storage (ADLS) and created data flow to load payroll data into SQL Database table.
- Deployed pipelines with Synapse Analytics via CETAS to ensure efficient data management and integration.
- Processed and analysed Bigdata such as data formatting, wrangling, data preparation, denormalization of data for calculations and reporting.
- Enabled ML & AI to work with image processing, video files, etc.
- Create Debug Workflow and batch macro to test and generate output by using Alteryx.
- Utilised Azure Monitor to keep track and ensure data processes work smoothly and accurately.
- Investigated and validated data against legislative requirements and compliance standards.
- Implemented automated triggers for Power BI and real-time streaming solutions using Azure Stream Analytics to enable processing and analysis of incoming data.
- Developed deep and interactive visualisation using Power BI to support timely decision-making.
- Maintained thorough and up-to-date documentation of the solutions, processes, configurations and the best practices related to cloud and data engineering.

EDUCATION

TORRENS UNIVERSITY AUSTRALIA

Sydney

Bachelor of Software Engineering

2023 – Present

(Specialisations in Artificial Intelligence)

- **University Excellence Scholarships**
- **High Distinction:** Concepts in Artificial Intelligence, Mathematics 1
- **Distinction:** Introduction to Software Engineering, Introduction to Cloud Computing

AUSTRALIAN COLLEGE OF BUSINESS AND TRAINING

Sydney

Certificate IV in Accounting & Bookkeeping

2023

CERTIFICATES

- Microsoft Certified: Fabric Analytics Engineer Associate
- AWS Certified Data Engineer Associate
- Alteryx Designer Core

AWARDS AND ACHIEVEMENTS

- University Excellence Scholarships
- Certificate of Merit in the contest of Computer Knowledge & Information Technology Skill (Youth Informatics Contest – Centre for Vietnam Youth Talents, Science and Technology)
- Certificate of Academic Excellence for outstanding performance in High school
- High School Vocational Certificate in Informatics