

# Rain Liao

Adelaide, Australia

+61 432 196 885 | liaoyuchieh6@gmail.com

[LinkedIn](#) | [Project portfolio](#) | [Github](#)

---

## PERSONAL SUMMARY

Data Engineer with practical experience designing and deploying scalable ETL pipelines, cloud infrastructure, and data storage solutions across AWS and Azure. Proficient in Python and SQL, with hands-on experience building containerized, cloud-deployed systems.

---

## WORK EXPERIENCE

### Aug 2025 - Oct 2025 Data Engineer Intern

University of Adelaide (Australia)

- Designed and deployed an end-to-end ETL pipeline (Python, Selenium, NLTK, Blob, Airflow) to ingest, clean, and structure 100,000+ unstructured social media posts into a query-ready format, applying statistical validation to topic modelling that identified 10 key subtopics within mental health discussions in Australia.
- Engineered a scalable data processing workflow supporting transformer-based modelling and a LangChain RAG system for downstream NLP analysis and insight extraction.
- Collaborated with academic supervisors and contributed a conference paper, with experience presenting technical work to research audiences.

### Aug 2024 - Mar 2025 Quantitative Trading Engineer

Australian Institute for Machine Learning (Australia)

- Built and maintained a high-volume data pipeline processing 100GB+ of Binance market data, combining fine-tuned LLM sentiment signals with technical indicators for automated trading strategy execution using Python and Freqtrade.
- Containerized and deployed the trading bot to Azure VM, cutting environment setup time by 80% and delivering a reproducible production-ready pipeline for live trade execution.

### Jul 2023 - Oct 2023 Data Engineer

National Centre for Research on Earthquake Engineering (Taiwan)

- Scraped, processed, and structured large-scale weather and seismic datasets using Pandas and QGIS, storing results in AWS DynamoDB and S3 for scalable downstream access.
- Built and deployed an event-driven data ingestion and alerting pipeline on AWS, using API Gateway webhooks and Lambda to trigger ETL execution on EC2 via SSM with IAM-scoped permissions, integrating REST APIs and Google Maps API, with end-to-end observability via CloudWatch and SNS on a Linux-based infrastructure.
- Led a cross-functional intern team, coordinating task delivery and communicating progress to supervisors to ship the system on schedule.

---

## PERSONAL PROJECT

2026

### Automated Weather Analytics Pipeline

- Deployed a containerized ETL pipeline on Azure VM using Docker Compose, orchestrating PostgreSQL, Airflow, PySpark, dbt, Superset, and Redis, with a medallion architecture and data retention policy, exposed via Cloudflare Tunnel for [production-equivalent public access](#) with DNS and SSL/TLS handling.
- Built a GitHub Actions CI/CD pipeline with automated DAG validation, dbt parsing, and database connectivity testing, deploying to Azure via SSH with credentials injected from GitHub Secrets at runtime.

---

## ACHIEVEMENTS

- Microsoft Azure Data Scientist Associate: DP-100
- Microsoft Power BI Data Analyst Associate: PL-300
- First-author conference publication – [ICRTCIS 2025 \(2026\)](#)
- Microsoft Fabric Data Engineer Associate: DP-700

---

## TECHNICAL SKILLS

Languages: Python, SQL, R, C++, JavaScript  
Data & Viz: ETL, dbt, APIs, Power BI, QGIS

Libraries: Pandas, Selenium, Scikit-learn, PyTorch, Airflow, PySpark  
Cloud & DevOps: AWS, Azure, Cloudflare, Docker, Linux, GitHub

---

## EDUCATION

### MASTER OF DATA SCIENCE

University of Adelaide | Adelaide, SA  
Feb 2024 – Dec 2025

### BACHELOR of CIVIL ENGINEERING

Chung Yuan Christian University | Taoyuan, Taiwan  
Sep 2018 - Jun 2023

## LANGUAGE

English (IELTS 7.5)

## REFERENCE

Available upon request