

Chien-Hsiang Yeh

in [linkedin.com/in/chienhsiang-yeh](https://www.linkedin.com/in/chienhsiang-yeh)  github.com/chypwc

Personal Profile

AWS Certified Data Engineer (Ph.D. in Economics) with hands-on experience building scalable data pipelines, real-time inference systems, and AI applications on AWS. Skilled in Python, SQL, Spark, and modern cloud-native tools. Strong foundation in analytics, automation, and problem-solving across academic and industry projects.

Skills & Certificates

- **Languages:** Python, SQL, Shell Script, STATA, MATLAB, JavaScript (Node.js), HTML, CSS
- **Databases & Storage:** PostgreSQL, MySQL, Snowflake, Qdrant, Pgvector, AWS S3 Vector Buckets
- **Data Engineering & Cloud:** AWS (Glue, EMR, SageMaker, Lambda, Step Functions, Kinesis, MSK, DMS, RDS, DynamoDB), PySpark, dbt, Airflow, Kafka
- **Frameworks & Visualization:** FastAPI, Streamlit, Tableau
- **AI & ML:** Pandas, NumPy, scikit-learn, LangChain, Retrieval-Augmented Generation (RAG) pipelines
- **DevOps & IaC:** Docker, Git, GitHub Actions, CloudFormation, Terraform
- **Certifications & Compliance:** AWS Certified Data Engineer – Associate; CPA Skills Assessment (Accountant – General)

Projects

E-commerce ETL & RAG AI Agent (Group Collaboration)

github.com/InsightFlow8/insightflow 

- Built a RAG AI chatbot with LangChain + FastAPI; used S3 vector storage and ALS models for personalised recommendations; enabled analysis via Athena.
- Deployed a Streamlit dashboard (Docker + EC2) for real-time analytics and visualisation of customer behaviour, marketing insights, and AI-driven recommendations.
- Designed a production ETL pipeline with Step Functions orchestrating Lambda, Glue Spark, and Crawlers for scalable ingestion and transformation.

Real-Time Inference

github.com/chypwc/kinesis-webui 

- Engineered a serverless real-time recommendation system using Lambda, API Gateway, Kinesis, and DynamoDB for low-latency XGBoost predictions via RESTful APIs.
- Preprocessed data with Glue Spark and automated ML training and deployment on SageMaker, all orchestrated through Step Functions.
- Enabled global delivery with CloudFront; automated provisioning with Terraform and GitHub Actions.

DMS + EMR Data Pipeline

github.com/chypwc/aws-dms-emr-terraform 

- Automated ingestion from PostgreSQL into Apache Iceberg on S3 using AWS DMS, EMR Spark, and Glue Catalog.
- Orchestrated end-to-end processing with MWAA (Airflow) using Terraform-generated DAGs.

dbt-Glue ETL

github.com/chypwc/aws-resources-exercises 

- Created scalable transformations with dbt and AWS Glue; automated Jobs, Crawlers, and dbt model deployments via CloudFormation and CI/CD (GitHub Actions).

Education

Australian National University, Ph.D. in Economics

July 2019 – July 2024

- Extended dynamic programming theory with state-action-dependent discounting.
- Proved convergence of Q-learning, SARSA, and Double Q-learning using real/functional analysis.
- Modeled equilibrium uniqueness in production and financial networks with Python simulations.

Australian National University, Master of Economics
National Tsing Hua University, B.S. in Physics; B.A. in Economics

Feb 2017 – Dec 2018
Sep 2008 – June 2012

Experience

Research Assistant, ANU RSE — Canberra

Oct 2024 – May 2025

- Structured raw datasets into panel format using STATA and Python.
- Built OCR pipelines to extract text from scanned documents.
- Conducted statistical analysis and literature reviews for labour economics projects.

Tutor, ANU RSE — Canberra

Feb 2019 – Nov 2024

- Delivered tutorials and consultations across core economics subjects: mathematical methods, optimisation, macroeconomics, growth, and time-series forecasting (MATLAB).
- Taught econometrics with Python, focusing on statistical modelling and applied regression/causal inference.
- Led discussions on dynamic programming and general equilibrium; supported 500+ students across 20+ tutorials.

Conferences

The Australasian Leadership Computing Symposium (ALCS) 2023

Canberra

Presented: *Harold Zurcher as a Q-learner*

36th PhD Conference in Economics and Business 2021

Perth

Presented: *Uniqueness of Equilibria in Interactive Network*

Society for the Advancement of Economic Theory (SAET) 2022

Canberra

Presented: *Uniqueness of Equilibria in Interactive Network*

Referees

Available upon request.