

# Willie Huang

tyc30827@gmail.com • [Willie Huang](#) • [github.com/tyc30827](#)

## WORK EXPERIENCE

---

### XRSpace Inc. | Taipei, Taiwan

Aug 2024 - Present

*Senior Data Engineer / MLOps Engineer*

- Designed and developed AI-labeling system using **LLM**, **Flask**, **Kafka**, and **Dagster** for automated content tagging
- Built multi-modal **RAG** system with **LlamaIndex** and **AWS Bedrock Knowledge Bases**
- Developed structured data RAG system using **DuckDB** and **Gemini-2.5**
- Optimized **BigQuery** data warehouse, reducing query costs by 90%
- Led team of 4 junior engineers as Code Owner for BI codebase

### CyberLink Corp. | New Taipei, Taiwan

Oct 2023 - Aug 2024

*Senior Data Engineer*

- Built NLP model (**LDA**) for customer feedback analysis and visualization in **Tableau**
- Developed data pipelines with **Airflow** and **Airbyte** to collect mobile install data
- Migrated data infrastructure from MSSQL to **AWS S3** and **Athena**

### CyberLink Corp. | New Taipei, Taiwan

Jun 2020 - Oct 2023

*Data Scientist and Engineer*

- Built personalized recommendation system, improving CTR from <1% to 30-60%
- Developed GNN-based user preference system using **AWS Neptune**
- Implemented **MLOps** practices with **MLflow** for model lifecycle management

## EDUCATION

---

### National Yang Ming Chiao Tung University | Hsinchu, Taiwan

Sep 2017 - Aug 2019

*Master of Science in Information Management*

**Publication:** First Author in **Information Fusion** (Top-tier AI Journal, 50+ Citations) | [Attentive Gated Graph Sequence Neural Network for Financial Trading](#)

### National Cheng Kung University | Tainan, Taiwan

Sep 2013 - Jun 2017

*Bachelor of Science in Industrial and Information Management*

## SKILLS & CERTIFICATIONS

---

- **Programming Languages:** Python, SQL
- **Data Engineering Tools:** Dagster, Airflow, Airbyte, dbt, Apache Kafka, Terraform
- **Database & Storage:** BigQuery, DuckDB, MongoDB, PostgreSQL, AWS S3, Athena
- **Machine Learning & MLOps:** PyTorch, Keras, MLflow, Databricks, AWS SageMaker, GNN, Recommendation Systems
- **LLM & RAG:** AWS Bedrock, PydanticAI, LlamaIndex, Gemini, Claude
- **Cloud & DevOps:** AWS (EC2, Lambda, Kinesis), Docker, Git, CI/CD
- **Data Visualization:** Tableau, Looker Studio, Superset
- **Certifications:** TOEIC 845