

Chen-Ming Yang

SOFTWARE ENGINEER · DATA ENGINEER

Taipei, Taiwan

✉ endlessproof@gmail.com | 🐱 afunTW | in chen-ming-yang | 📧 @afun | 📧 a_fun

"A Foolish Consistency is the Hobgoblin of Little Minds"

Summary

Dedicated Data Engineer with robust experience in optimizing and organizing vast datasets from raw form to analytical insights. Proficient in creating seamless, user-friendly data services that mask underlying complexities. Adept at navigating and contributing to diverse, multi-cultural environments, I bring a global perspective to data engineering challenges.

Eager to engage in projects that involve large volumes of data or intricate streaming pipelines, I am driven to collaborate with top-tier professionals in pursuit of innovative solutions in the international tech landscape.

Work Experience

FULL-TIME

Kronos Research

Taipei, Taiwan

SOFTWARE ENGINEER II

Nov. 2021 - Now

- Orchestrated a high-performance data exporter project, reducing data landing times to under 10 seconds using a custom Go agent and AWS Kinesis Data Streams, focusing on advanced service management and efficient data transmission.
- Developed a user-friendly data access API service using Golang and Kubernetes, making complex HFT datasets accessible to a broader range of developers, and simplifying data utilization processes.
- Achieved a breakthrough in data processing efficiency by integrating CGo into Python pipelines, cutting down processing times for over 500 GB of daily data from 24 hours to just 3 hours.
- Led a strategic data collector migration, enhancing data granularity and streamlining data stream selection and composition, ensuring smooth downstream integration.
- Maintained over 95% SLA compliance through effective monitoring, retrying, and recovery strategies, establishing a high standard of reliability in data services.

WeMo Scooter Corp.

Taipei, Taiwan

DATA SCIENTIST

Sep. 2019 - Oct. 2021

- Pioneered the transition from ETL to ELT, reducing pipeline development time from days to an hour by leveraging Google GCS, PostgreSQL, and BigQuery, enhancing query efficiency and performance.
- Implemented and customized Metabase for BI reporting, empowering non-technical team members to autonomously analyze and report, significantly reducing dependency on the data team.
- Utilized HDBScan and Voronoi diagrams for operational hot-spot analysis, informing resource allocation and improving dispatch strategies.
- Conducted battery life-cycle and user behavior analysis to refine e-scooter rental strategies, boosting profitability and user experience.
- Led a landmark collaboration with major telecommunications providers, using OD datasets and Signal RSSI metrics to identify and resolve urban signal blackholes, enhancing city-wide network coverage.

Academia Sinica, Institute of Information Science

Taipei, Taiwan

RESEARCH ASSISTANT & RDSS (COMPULSORY MILITARY SERVICE)

Aug. 2016 - Aug. 2019

- Developed and optimized an image segmentation pipeline using deep learning for insect behavior analysis, resulting in a major publication in Nature Communication. Applied a custom-trained detection model and the Hungarian algorithm for precise tracking of complex video data.
- Released a PyQt5-based visualization tool, streamlining video analysis and reducing manual labeling time for a dataset comprising 6 years of continuous footage.
- Led Python web crawling workshops, covering techniques from static and dynamic crawling to data analysis, enhancing team capabilities in data acquisition and processing.
- Facilitated geospatial and climate data extraction using Google Earth Engine, accelerating research workflows with efficient data handling and preprocessing.

Side Projects

IPI

Remote

IoT TELEMETRY DATA ANALYSIS PIPELINE

July 2021 - Sep. 2021

- PoC of automatically archive items to S3 using DynamoDB TTL with Lambda and Firehose
- PoC of building data analysis pipeline with Glue and Athena

KYC Right

Remote

SPECIALLY APPOINTED RESEARCH ASSISTANT

Nov. 2019 - Apr. 2020

- PoC of community detection on Neo4j

Taiwan-Ebook-Lover

Remote

SEARCH THE E-BOOKS ON DIFFERENT PLATFORMS

Sep. 2021 - Now

- Using Prometheus to export metrics and monitoring on Grafana on 1 virtual machine

Education

National Chi Nan University (NCNU)

Nantou, Taiwan

BS & MASTER'S DEGREE IN COMPUTER SCIENCE AND INFORMATION ENGINEERING

Sep. 2011 - Jul. 2016

- Paper "Integrated Sleep Scheduling for Multiple Relays and UEs in LTE-TDD"
- The Director of Student Association of CSIE Department.

Memorial University of Newfoundland (MUN)

St. John, NL, Canada

BACHELOR EXCHANGE STUDENT IN COMPUTER SCIENCE

Sep. 2014 - Apr. 2015

- Two semesters with GPA 3.5, 3.6 respectively.