

# YUHAO DAI

TOKYO, JAPAN

OBJECTIVE: DATA SCIENTIST, MACHINE LEARNING ENGINEER OR DATA ENGINEER AT A PROGRESSIVE ENTERPRISE.

1996/09 (25 y/o)

yuhao2dai@gmail.com

## SKILLS

- English (TOEIC 950, TOEFL 115), Japanese (N1), Chinese (Native)
- Unix Commands, Git
- Fluent Python, sklearn, SciPy series, TensorFlow (Keras), PyTorch etc. (4 years)
- Intermediate Java. (1 year)
- AWS SAA, GCP PCA / PDE / MLE, Azure AI Engineer etc.
- Bookkeeping Level 3, Statistics Certifications Level 2
- SQL(SQL Server/Snowflake)
- Azure Functions / App Service / Data Factory
- AirFlow on VM / PySpark
- Engineering manager (code review, life cycle and member education)
- List: [y2data.github.io/certs](https://y2data.github.io/certs)

## EDUCATION

2017-2021 Doshisha University, Department of Letters, English Literature (Diploma in BA)

## JOBS

AUGUST 2022 – PRESENT: COCA COLA BOTTLERS JAPAN INC.

Senior Data Engineer and Data Service section assistant manager.

ETL pipelines on Azure Synapse using SQL, Dataflow Mapping, AirFlow and PySpark.

Engineering manager of a team of 9, code review, Kanban and development lifecycle.

Business user communication and technical QA.

Data platform transitioning from SAP APO/ECC/Hana, Marklogic to ADLS Gen2 and Synapse.

Report building and automation for existing Excel/manual/VBA based reports.

December 2021 – July 2022: bitFlyer K.K.

Data Scientist and Infrastructure Engineer

Migration from Synapse dedicated pool to Snowflake, 10X performance and 1/8 cost.

Data analysis from whole company via ticketing system, monthly 50+ tickets.

Organizational effort to establish Data Strategy office.

Automating manual & VBA-dependent work with Python and Azure Functions.

Developing internal packages on GitHub, automate deployment with GitHub Actions

Data pipeline via AirFlow and Azure Data Factory.

Modeling for user clustering and anomaly detection for anti-money-laundering.

Managing and developing Azure infrastructure and applications.

April 2021 – November 2021: InfoSys Limited

Associate, Tokyo

GCP technical support representative.

Products: BQ, Dataflow, Pub/Sub, Bigtable etc.

Japanese, English and Chinese.

Client-facing and constantly challenged.

Debugging, data-pipeline architecting, SQL query performance improvement.

Building and maintaining elastic and resilient infrastructure architectures.

February 2018 – February 2020: BIAS Ltd.,

Data Analyst, Beijing

Joint business venture.

Pilot training and commercial airplane rental.

Data pooling, transforming and cloud architecting on AWS.

Python, Java, AWS, TensorFlow and Keras.

Build from scratch, analyst team management.