

## Ming Zhe Teo

+65 84343248 | [LinkedIn](#) | [teomz96@outlook.com](mailto:teomz96@outlook.com) | [Github](#) | [Website](#)

### WORK/PROFESSIONAL EXPERIENCE

---

#### Luxoft

##### Data Engineer (Temasek)

Jun 2024 – Current

- Developed and managed ETL/ELT processes to ensure accurate and timely data flow. Implemented automation scripts to enhance the efficiency and scalability of data ingestion.
- Created user-friendly Tableau dashboards that present complex data in clear and actionable formats. This significantly improved user experience and interactivity while reducing latency.
- Enhanced the efficiency of data pipelines and the responsiveness of dashboards, resulting in a better user experience.
- Maintained high standards of data accuracy and reliability through rigorous validation check and pipeline monitoring.

#### Cynapse

##### AI Software Engineering Intern

Oct 2023 – Mar 2024

- Automated synthetic dataset generation by developing a data pipeline leveraging stable diffusion and LLAMAs for creating multimodal person attribute datasets.
- Utilized AWS SageMaker to streamline training for a sophisticated multimodal Generative AI model tailored for person search within CCTV surveillance databases, implemented using Python.
- Enhanced dataset adaptability by augmenting data structure, ensuring superior generalization across diverse use cases.
- Implemented prompt engineering techniques to optimize descriptive text generation for model training, resulting in significant model improvement.

#### TUV SUD PSB

##### HR Analytics Intern

Nov 2022 - May 2023, Apr 2024 – May 2024

- **Best Internship Award Winner (2023):** Honored for improving reporting processes.
- Created detailed annual HR reports for management, supporting strategic decisions with data-driven insights.
- Developed a dynamic Excel dashboard for real-time data exploration, enhancing user experience and analysis efficiency.
- Collaborated across regions and departments to define data requirements, ensuring accurate and insightful data representation in reports and dashboards.

### EDUCATION

---

#### University of London

##### BSc (Hons) in Computer Science (Machine Learning and Artificial Intelligence)

Aug 2017 - Sep 2023

### SKILL/CERTIFICATION

---

#### [AWS Certified Cloud Practitioner](#),

Python, Pandas, Scikit-learn, Pytorch, Tensorflow, Go, Gin, MongoDB, Docker, PostgreSQL, MySQL, Git, Window, Linux, Vercel, Snowflake, Tableau, Cassandra, Airflow, Pyspark, Swagger

### PROJECTS

---

#### [Price-Tracker](#)

##### PostgreSQL, Go, Python, Airflow

- The Price Tracker is a microservices-based system for scraping, storing, and analyzing product prices using a data pipeline. Built with Go (Gin) for the backend and leverages Airflow to trigger the Gin API (Swagger-documented) and containerized with Docker

#### [Simple Pipeline](#)

##### Cassandra, Airflow, Python, Docker, Kafka, Pyspark

- A simple data engineering pipeline that orchestrates data ingestion, processing, and storage using Apache Airflow, Python, Apache Kafka, Apache Zookeeper to Apache Spark and update Cassandra while containerized with Docker for easy deployment