Pin Ren Toh

tohpinren@gmail.com | tohpinren.com | github.com/tohpinren | linkedin.com/in/tohpinren

Experience

Kipo AI, San Francisco, CA, USA — **Software Engineer Intern**　　　　　　　　　　May 2025 – Aug 2025
- Designed and implemented a **multimodal (OCR + vision + LLM) extraction pipeline** to extract electrical specifications, dimensions and terminal properties from electronics datasheets, achieving **91% precision** on **100K** vendor datasheets.
- Refactored an obsolete single-vendor pipeline into a generalised **agentic extraction system**, orchestrating models for different data modalities; expanded vendor coverage and unlocked new paid PoCs.

Kipo AI, San Francisco, CA, USA — **Software Engineer Intern**　　　　　　　　　　Jan 2024 – Dec 2024
- Engineered a graph/diagram parsing service for dimensional and terminal data extraction; integrated **error-driven re-prompting with schema validation**, boosting precision to **81%** and reducing error rate by **27%**.
- Developed a scalable **scraping framework** to deploy/terminate AWS EC2 instances for Scrapy-based crawlers, with automatic ingestion into MongoDB and S3, enabling the team to collect **5M+** datasheets.
- Prototyped a hybrid part-search system combining **retrieval-augmented generation (RAG)** with **LLM-based NL-to-SQL**, enabling natural language queries to retrieve matching components.

Marymount Labs, Singapore — **Software Engineer**　　　　　　　　　　Jan 2023 – Oct 2023
- Built an **NL-to-SQL assistant** enabling natural language search across **1M+ clinic records** for nurses and doctors.
- Developed an **SMS follow-up system** with Twilio, boosting patient return rates by **25%**.
- Deployed an **Electronic Health Record** system to **70+** clinics, reducing nurse data-entry time and improving consistency of patient records.

i-CARE Primary Care Network, Singapore — **Data Analyst**　　　　　　　　　　Feb 2022 – Jun 2022
- Automated MOH compliance reporting for **40+** clinics, cutting report generation time from several hours to 30 minutes and eliminating manual errors.

NEC Asia Pacific, Singapore — **Assistant Software Engineer**　　　　　　　　　　Nov 2021 – Jan 2022
- Developed and tested RESTful backend APIs for a CRM system (Casino Regulatory Authority) used in regulatory data workflows.

Projects

FYP — SQL Mutation Testing for NL-to-SQL Benchmarks (BIRD), supervised by Prof. Manuel Rigger　　Aug 2025 – Present
- Developing a SQL mutation-testing framework that generates query variants from benchmark ground-truth SQL and evaluates their equivalence via execution results.
- Initial results show **>50%** of mutants survive execution checks, highlighting benchmark data insufficiency and demonstrating that execution accuracy alone fails to capture semantic query equivalence.

CoreFlint — AI Procurement Automation Platform　　　　　　　　　　Sep 2025
- Built an MVP platform to transform unstructured procurement documents (RFQs, POs, quotes, order confirmations) in different formats (PDFs, emails) into structured ERP-ready outputs (SAP/Oracle).
- Implemented OCR + LLM pipeline for line-item extraction with human-in-the-loop review and approval, ensuring accuracy and auditability.

WhatsApp Actions — LLM-Powered Business Messaging Assistant　　　　　　　　　　Aug 2025
- Developed prototype assistant using LLMs to classify WhatsApp Business messages, generate draft replies and route outputs through human-in-the-loop approval.
- Piloted with an SME to validate operational automation and improve response turnaround time.

Education

National University of Singapore (NUS) — B.Comp in Computer Science　　　　　　Expected May 2026
- GPA: 4.64 / 5.00 (First Class Honours), Teaching Assistant (CS1101S)

Stanford University — NUS Overseas Colleges Silicon Valley　　　　　　　　　　Mar 2024 – Jun 2024

Skills

Languages & Databases: Python, JavaScript, Java, SQL, PostgreSQL, MySQL, MongoDB, Redis
Frameworks & Systems: FastAPI, REST APIs, Machine Learning, LLMs
Cloud & Tools: AWS (EC2, S3, SageMaker), Docker, CI/CD, GitHub Actions, PyTest