

## WANG Xuxin (Benjamin)

Website: [wxuxin.org](http://wxuxin.org) | Mobile: +65 96750912 | Email: [wangxuxin1@gmail.com](mailto:wangxuxin1@gmail.com) | GitHub: <https://github.com/WXUXIN>  
Linkedin: <https://www.linkedin.com/in/wang-xuxin/>

### TECHNICAL SKILLS

**Languages:** Python, Java, Javascript, Typescript, Golang, SQL, R, Web Programming (HTML5, CSS3)

**Technologies:** Redis, MySQL, MongoDB, Apache Kafka, Apache Airflow, Apache Spark, Express.js, AWS, dbt, React.js, Node.js, Pandas, Selenium, Google BigQuery, Firebase, Git, Django

### WORK EXPERIENCE

#### Backend Software Engineering Intern (Aug 2024 - Jul 2025)

**CLERKIE.IO** | San Francisco Bay Area | Silicon Valley Series A Fintech Startup; AI-powered financial automation platform

- Developed and deployed a **mission-critical Debt Recovery CRM** and Payment Portal with **Node.js** and **MongoDB**, enabling agencies to manage over **\$1B in debt portfolios**.
- Improved job success rate from 80% to 99.9%** by implementing retries with exponential backoff, and dead-letter queues in **Redis/BullMQ**.
- Implemented compliance automation (statute of limitations logic) to ensure regulatory adherence for enterprise clients including **Verizon** and **Wells Fargo**.
- Engineered secure wire transfer pipelines** with **JPMorgan Chase** for ACH transactions, ensuring reliability and timeliness.
- Optimized large-scale data processing workflows**, cutting report generation time by **73%** and scaling systems to tens of millions of transactions.
- Integrated enterprise-grade payment gateways** (Payscout, Repay, etc.), boosting on-time payments by **31%** and expanding revenue collection.

#### Data Analytics Engineering Intern (Jan 2024 - Jun 2024)

**SHOPBACK** | Singapore | Multinational e-Commerce Cooperation; Leading cashback rewards platform in Asia-Pacific

- Designed and managed** end-to-end ETL pipelines using **Apache Airflow, Spark, and dbt, MySQL**, improving reliability and reducing manual intervention by **40%**.
- Developed first Geospatial BI system** using **Python, MySQL, and ML algorithms**, enabling zonal-level scaling and improving resource allocation efficiency by **80%**.
- Optimized data models** with **dbt** and **SparkSQL**, improving query performance by **30%** and boosting analytics accuracy.
- Executed zero-downtime migration** of **56M+ records**, ensuring seamless data integrity and system continuity.

#### Data Analyst Intern (Aug 2023 - Nov 2023)

**QUEST- HIRE A HERO** | Singapore | Early stage startup; Request-based mobile platform matching freelancers to businesses

- Increased lead generation 5x** by developing automated web-scraping tools with **Python** and **Selenium**.
- Automated data visualization pipelines** with **Pandas**, reducing reporting effort by 15+ hours weekly.
- Drove product growth**, achieving **140% MAU** and **120% DAU** increase through data-driven insights.

#### Undergraduate Teaching Assistant (2023 - 2024)

**NATIONAL UNIVERSITY OF SINGAPORE**

- Taught programming fundamentals using Python through weekly tutorials with a class of 40 students.

### EDUCATION

#### NATIONAL UNIVERSITY OF SINGAPORE

Bachelor of Science (Honors) in Business Analytics, Minor in Data Engineering (**Aug 2022 - Nov 2026**)

- GPA: 4.63/5 (Honors with Highest Distinction)

#### STANFORD

Management Science and Engineering (MS&E) (**2024 - 2025**)

### PROJECTS

#### EduGap (Oct 2024 - Present) [Project Link](#)

- Built an AI-powered platform integrating **Google Gemini 1.5, LangChain, ChromaDB** to support student-teacher interactions.
- Implemented semantic clustering (PCA, DBSCAN)** to analyze student queries and highlight common learning gaps.

#### Geospatial Analysis (May 2024 - Jun 2024) [Project Link](#)

- Optimized geospatial coverage using **custom radius algorithms** with **GeographicLib** and **geopy**, increasing relevant business capture by 30%.
- Developed hybrid custom search strategy using **clustering techniques like DBSCAN** to optimize API efficiency.
- Built a **scalable data pipeline** integrating Google Maps & OneMap APIs, processing **5,000+ locations** with rate limiting and deduplication, improving data accuracy by **40%**.