# Xijing Wang

408-718-5711| xthomaswang@gmail.com
https://xthomaswang.github.io/ | https://www.linkedin.com/in/xijing-wang-914a06195/

## EDUCATION

**Carnegie Mellon University** ***School of Computer Science*** **Pittsburg, PA**
Master of Science: **Automated Science** *Aug 2025 - Spring 2027*
*Reward & Honor: Dean's Scholarship*

**Santa Clara University** ***College of Arts and Sciences*** **Santa Clara, CA**
Bachelor of Science: Major in ***Computer Science,*** Minor in **Economics** *Sep 2021 - Spring 2025*
*Reward & Honor: REAL Program Scholar - Summer Research Internship*

## Language / Skills

**Languages:** Python, C++, Swift, JavaScript, Java, C++, Go, Scala, HTML/CSS
**Mobile Development:** SwiftUI, UIKit, React Native, Redux, Core Data, Vision Framework, Faiss, NLP, CV
**AI / Machine Learning:** mlsys, LLM, TensorFlow, Pytorch, CoreML, Keras, OpenCV, NumPy, Pandas
**Backend / Full-Stack:** FastAPI, Node.js, Gin (Go), Firebase, SQLite, MySQL, PostgreSQL, SQLAlchemy
**Architecture & Concepts:** MVVM, RESTful APIs, Data Privacy, CI/CD, Data Structures, HPC, Test Automation

## Experience

**Research Engineer (AI agent)** | *Python, FastAPI, PubMed API, MedRxiv API, OpenAI API* **Pittsburg, PA**
*Collaborated with PHD student in CMU on biomedical ai Agent project* **July 2025 - Current**

- Designed and built an **AI-powered biomedical literature analysis** platform with Quick (<30s) and Deep (≤3min) modes, using a **RAG pipeline** with **parallel retrieval** from PubMed and MedRxiv.
- Implemented **structured user query parsing**, **embedding-based semantic search** with reranking**,** and **dynamic LLM model optimization**, improving retrieval accuracy, efficiency, and cost-effectiveness for generating structured, high-relevance reports at scale.

**Machine Learning Engineer Intern** | *Shell, Linux, Deep Learning, Data Analysis* **Santa Clara, CA**
*Real Program Summer Funding Opportunity in LCCN lab at SCU* **June 2024 - Sep 2024**

- Developed and adjusted **Convolutional Neural Network (CNN) models** for testing neuroscience theories, including **Dataset Preprocessing** and **Parameter Tuning**.
- Preprocessed fMRI and anatomical data from online datasets and transferred to **WAVE HPC platform** using **Shell Scripting**.
- Analyzed research results, constructing **Correlation Matrices** using **Pearson Correlation**, and authored manuscripts detailing research findings and methodologies.

Paper accepted for **poster presentation** at **CogSci 2025** in San Francisco. Currently under review at Communication Biology.
*https://doi.org/10.1101/2025.04.02.646903*

## Project

**Headache Note** | *MLC-llm, MLsys, Swift, SwiftUI, Pytorch*
*https://github.com/xthomaswang/headache_note*

- Built a fully offline iOS **AI agent** by **quantizing LLaMA 3 1.6B** with **MLC-LLM,** optimizing for iPhone. Developed a Swift-based **health assistant** with **prompt engineering** and **SwiftData** for **private**, **on-device inference** and personalized lifestyle recommendations.

**Food Recognition iOS App** | *Swift, SwiftUI, Tensorflow, CoreML, Python*
*https://github.com/xthomaswang/FoodRecogProj*

- Built a **SwiftUI** iOS app using **MVVM** architecture, integrating a **TensorFlow**-trained **CoreML** model for food classification with **85%** accuracy across **100+ categories**. Optimized inference speed by **4×** and reduced memory usage by **40%** through model tuning and image augmentation on **50K+ samples.**

**MasumiRanker: AI Agent Platform (Hackathon)** | *Python, FastAPI, SQLite, Transformers, Faiss, React.js*
*https://github.com/xthomaswang/Masumi_Hackathon_Whym_MasumiRanker*

- Developed an **AI agent discovery platform** with **natural language semantic search** using **Sentence Transformers** and **Faiss** for **efficient similarity matching**. The **FastAPI** and **SQLAlchemy** backend managed agent cataloging, user ratings with automated aggregation, recommendation logging, and review data integrity via SHA-256 hashing.