Ken Wu

kenwu.is-a.dev | KenWuqianghao | in kenwuu | ✉ ken.wu@uwaterloo.ca | ☎ 437-971-3179 | Waterloo, ON

EDUCATION

University of Waterloo   Sept. 2022 – Apr. 2026
Bachelor of Computer Science, Business Specialization, Philosophy Minor   *Waterloo, Canada*
- URA @University of Waterloo: Symbolic Regression using GPT Models, Exchange Researcher @Lancaster University: Clustering Algorithm Optimization, Technical Project Manager @Wat.ai: Medical RAG Agent

SKILLS

Languages: Python, SQL, C, C++, Typescript, Javascript, HTML, CSS, R, Racket, Bash
Frameworks: PyTorch, PySpark, LangChain, Tensorflow, Keras, CUDA, Flask, FastAPI, React, Next.js, Tailwind CSS
Technologies & Cloud: NumPy, Pandas, spaCy, Flask, FastAPI, MongoDB, Supabase, PostgreSQL, AWS, Docker, Git

EXPERIENCE

Software Engineer Intern   Sept 2024 – Dec 2024
Vecflow | Python, LangGraph, LangChain, FastAPI, Supabase, AWS   *Legal Agent | New York, United States*
- Handled **2000+ requests/min** by deploying **10+ API endpoints** using FastAPI, AWS, and Supabase
- Optimized evaluation cycles for **15+ LLM models** with a round-robin multi-agent and scoring framework
- Accelerated workflows for **20+ lawyers** by automating document redlining using DOCX and XML manipulation
- Built a programmatic answer generator for tabular data by designing an advanced ask extraction pipeline
- Led the end-to-end development of a multi-agent RAG pipeline powered by LLM-as-Judge strategies

Machine Learning Engineer   Apr 2024 – Aug 2024
Health Canada | Python, Azure, LangChain, Streamlit, ChromaDB, Ollama   *Document QA | Ottawa, Canada*
- Built a document QA system using **Llama3 7B** and **ChromaDB** for OECD report search and summarization
- Increased response quality and semantic accuracy using **query transformation** and **contextual memory**
- Deployed and maintained a Streamlit app on Azure VM with auto-shutdown/start for improved runtime efficiency

Data Analyst Intern   Jan 2024 – Apr 2024
Saputo | TypeScript, Office Script, VBA, Excel, AS400   *Sales Automation | Georgetown, Canada*
- Eliminated **8+ hours/week** of repetitive work by automating Excel workflows with Office Script and TypeScript
- Saved **20,000+ manual entries** weekly by automating order updates for **1000+ Nestlé product lines**
- Streamlined competitor tracking by automating research and summarization for **200+ industry brands**
- Performed weekly statistical analysis and exploratory data analysis on **2000+ SKUs** using VBA macros

Software Engineer   Mar 2023 – Jun 2023
Keywords AI (Y-Combinator W2024) | Python, spaCy, SQLite   *Resume Parsing | New York, United States*
- Parsed **1000+ resumes** with a spaCy-based NER pipeline to extract structured and relevant recruiter data
- Led a team of **12 interns** in data labeling, prompt engineering, and model testing for NER improvement
- Improved system efficiency by integrating SQLite-based result caching into the resume parsing engine

Machine Learning Engineer Intern   Jul 2022 – Sept 2022
Intapp (Formerly delphai) | Python, spaCy, Azure, W&B, Selenium, BS4   *Entity Recognition | Berlin, Germany*
- Scraped and labeled **2000+ Chinese company names** with Selenium and BS4 for multilingual NER training
- Boosted recall by **20%** through improved entity labeling workflows and language-specific training sets
- Achieved **70% recall** by fine-tuning spaCy models and optimizing hyperparameters via WanDB on Azure

PROJECTS

MedChat ⧉ | *Python, Cohere API, Streamlit, Tensorflow, OpenCV*   Nov 2023
- **Winner of 2023 Cohere RAG (Retrieval Augmented Generation) Challenge** out of **50+** hackers
- Employed Tensorflow, OpenCV, and RAG for dynamic information retrieval and CNN for medical image analysis

DirectU ⧉| *Python, Flask, Cohere API, MongoDB, React, Selenium*   Sep 2023
- **Winner of Hack the North 2023 Best Use of Cohere Prize** out of **850+** hackers
- Enhanced student academic planning by providing personalized course recommendations, significantly improving user experience and engagement with academic resources