

ANDREW JASON WOODS

(727) 992-8482 | andrew.woods.jason@gmail.com | LinkedIn | GitHub

EDUCATION

University of California, Los Angeles

Los Angeles, CA

Bachelor of Science in Computer Science

2026 – 2029

Coursework: Linear Algebra, Multivariable Calculus, Differential Equations, Discrete Mathematics

Additional Coursework: Data Structures & Algorithms in Python (Computer Programming 3)

EXPERIENCE

Software Engineer Intern

Jun. 2026 – Aug. 2026

OmicsAI Labs

Remote

- Unified **3 web and AI applications across 4 Docker services** on AWS EC2 behind a single Caddy ingress, isolating application ports and persisting proxy and application state with Docker Compose.
- Built a **6-page React/Vite research portal with 2 AI demo entry points**, productionized through a two-stage Docker build, Nginx SPA serving, and 30-day static asset caching.
- Adapted **2 Python research interfaces** for reverse-proxy subpath hosting using Gradio, Streamlit, and Caddy, configuring long-request proxying and a 200 MB OmicsGPT request-body limit.

AI Research Intern

Feb. 2024 – Nov. 2025

Georgetown University

Washington, DC

- Built a **5.8K+ non-comment-line Python experimentation pipeline across 16 notebooks**, integrating external APIs, graph construction, feature weighting, embeddings, and evaluation for disease-aware multi-omics modeling.
- Benchmarked **36 model/representation combinations** (6 classifiers $\times$ 6 representations) with stratified 5-fold CV on 382 samples, reaching a best reported CV AUC of **0.940 ± 0.014** .
- Integrated g:Profiler across **4 external data sources** and cached 4 intermediate graph/weighting artifacts, enabling reproducible reruns across heterogeneous API response schemas.

PROJECTS

Data-Efficient Intent Routing System

Python, scikit-learn, Sentence Transformers, OpenAI API

- Built a confidence-routed **77-intent classifier** using frozen MiniLM embeddings and multinomial logistic regression; the local specialist reached **85.58% accuracy** on 3,080 held-out cases.
- Improved held-out accuracy to **88.08%** by routing only **8.70%** of requests to a retrieved-example LLM fallback, reducing usage-priced LLM API spend by **91.1%** versus all-Luna inference (\$0.0514 vs. \$0.5763 known charges).

Retrieval Evaluation System

Python, NumPy, pytest, GitHub Actions

- Built a retrieval evaluation platform over **34.5K documents and 1,623 judged queries**, benchmarking BM25, dense retrieval, reciprocal rank fusion, and cross-encoder reranking.
- Found cross-encoder reranking regressed quality on **2 of 3 datasets** while increasing latency by roughly **70 $\times$** , exposing a substantial speed-quality tradeoff before deployment.
- Validated ingestion, ranking, metrics, caching, CLI, and service behavior with **501 passing tests** and GitHub Actions; rejected stale labels using chunk provenance and corpus text identity.

TECHNICAL SKILLS

Languages: Python, Java, JavaScript, SQL, HTML/CSS

Web & Infrastructure: React, FastAPI, REST APIs, AWS EC2, Docker, Docker Compose, Caddy, Nginx, Linux, Git, GitHub Actions

Testing & Data: pytest, NumPy, pandas, scikit-learn, Sentence Transformers, NetworkX

PUBLICATIONS

J. Li, **A. Woods**, A. Yuan. “Agentic AI for Disease-Aware Adaptive Multi-Omics Embedding: A Proof of Concept with PCA and Bulk RNA-Seq Data.” *IEEE ICDM 2025, SAIMBio Workshop*.