

# Steven (Hanqi) Su

🔗 | in | hanqisu@sas.upenn.edu | Portfolio

## EDUCATION

University of Pennsylvania, Accelerated M.S.E. in Robotics | Philadelphia, US

Expected May 2026

University of Pennsylvania, B.A. Computer Science + Computational Biology | Philadelphia, US

Expected May 2026

## SKILLS

**Programming** Python, Java, JavaScript, Swift, R, SQL, HTML & CSS

**Frameworks** Linux, Docker, Spark, Apache Airflow, PyTorch, Pandas, Flask, Next.js, LangChain, Scikit-Learn

**Tools** Git, Databricks, Google Cloud, AWS, PostgreSQL, MySQL, MongoDB, Firebase, Redis, Vector Database

## EXPERIENCE

**Grove AI, Software Engineering Intern** | San Francisco, CA (Remote)

July 2024 - Present

- Conducted extensive EDA and feature engineering on relational databases using **SQL joins** and **Pandas**
- Utilized **Scikit-Learn**'s Random Forest and Logistic Regression models to analyze feature importance and generate key insights into the causes of patient dropouts from clinical trials
- Developed API endpoints to execute complex **PostgreSQL** queries, enabling real-time visualization of the performance metrics for the company's automated LLM phone call agents

**Zhi Huang Lab, Software Engineering Intern** | Philadelphia, PA

Jul 2024 - Present

- Building TissueLab, spinout project from Nuclei.io 🔗, the pathologist-AI collaboration framework
- Spearheaded reconstruction of the MVP's frontend from Python native GUI libraries into **Next.js** to accommodate low-code access to platform's Vision-Language Models for physicians

**Ancient Tech, Software Engineering Intern** | Palo Alto, CA (Remote)

May 2024 - Present

- Received UPenn's Startup Internship Award, doubled the size of training data for company's Reinforcement Learning for Finance projects by connecting 3 new ticker screeners to **MySQL** in the ETL pipeline of the production codebase
- Utilized **Docker**, **Selenium**, and **X virtual framebuffer** to deploy Docker images that enable automated browser interactions in Linux-based production servers without graphical user interfaces
- Leveraged **Redis** to enable robust market data caching and persistence and reduce model training time by 30%

**SciSketch 🔗, Technical Founder and CEO** | Philadelphia, PA

Feb 2024 - Present

- Designed and implemented an innovative backend framework utilizing **Apache Airflow** and **MySQL** to automate the weekly scraping of biomedical journal content from Elsevier APIs, enabling the continual fine-tuning of LLMs with newly published and user-generated data
- Engineered a text-icon mapping framework through knowledge graph-based RAG with **ChromaDB** and **LangChain**
- Built distributed and GPU-enabled env. with **AWS EC2**, **Spark**, and **GCP**. Reduces computation time by 70%
- Used **PyTorch** and Huggingface APIs to fine-tune Google Flan-T5 and produce two custom LLMs, BioPhraser 🔗 and BioRegressor 🔗. Deployed both models for inference with **AWS Sagemaker**

**Penn Epigenetics Institute, Bioinformatics Researcher** | Philadelphia, PA

Sep 2023 - Present

- Spearheaded development of MEAnalysis 🔗, an **R**-package for institute internal use in streamlined analysis of neuron multielectrode array data. Leveraged **Objected-Oriented Programming** paradigm in program design and reduces researchers' analysis time by 60%
- Engineered a Next-Gen sequencing data analysis pipeline with **R** and discovered 3 candidate risk genes for autism

## PROJECTS

**Vision4Science 🔗** | Swift, OpenAI API, Firebase, VisionOS

Mar 2024 - Apr 2024

- The Generative-AI powered Apple Vision Pro app that teaches biomedical researchers how to perform experiments
- Built the VisionPro App with **Swift**, **Reality Composer Pro**, with user data hosted on **Firebase**
- Implemented Generative AI functions through OpenAI API and deployed on Firebase Cloud Functions

**NeuraLIVE 🔗** | OpenCV, Flask, HTML, CSS

Oct 2023 - Nov 2023

- A bioinformatic tool to analyze cultured neuron cells' proliferation and neurite growth from microscopy videos
- Developed automated neurite length detection and cell viability quantification feature with Python and **OpenCV**
- Hosted the backend of the webapp on **Flask** and built the frontend with **HTML & CSS**
- Decreases time and financial cost associated with the neuron imaging process by **40%** and **50%** respectively

## COMMUNITY LEADERSHIP

**TEDxPenn, Club President** | Philadelphia, PA

Aug 2022 - Present

- Leading the world's largest student-ran TEDx platform with more than 1,000 attendees
- Lead a team of 40 for conference logistics planning, speaker and sponsor outreach, and marketing efforts
- Notable speakers include a Nobel Prize in Medicine Laureate, the previous national president of Peru, Director of multiple Marvel films, and Chief Cloud Architect at NVIDIA