

William Liu

☎ (608) 886-3074 // ✉ me@williamliu.me // 🌐 www.williamliu.me

Education

Carnegie Mellon University

B.S. in Cognitive Science, Minor in Computer Science // Aug 2016– May 2020

Experience

Amazon // Palo Alto, CA

Software Development Engineer II // Sep 2022 – Present

- Tech Lead on the ML experimentation platform for Amazon.com's search rankers model development
- Designing and leading a complete infrastructure upgrade to increase scientist productivity on the platform

CoPilot // Pittsburgh, PA (Remote)

Lead Backend Systems Engineer // Aug 2021 – Sep 2022

- Led the design, development, and maintenance of the entire backend system as the primary backend engineer
- Performed business analytics for day-to-day operations as well as for fundraising rounds

SambaNova Systems // Palo Alto, CA

Software Engineer // Jun 2020 – Aug 2021

- Led a small team to increase training speed of multiple 2D/3D computer-vision ML models by more than 10X
- Drove the design process of a new hardware-optimized convolution operator

NVIDIA // Santa Clara, CA

Deep Learning Software Intern // May 2019 – Aug 2019

- Improved ML model inference latency and memory usage through in-compiler weights compression

Uber // Santa Clara, CA

Software Engineering Intern // May 2018 – Aug 2018

- Surveyed, designed, and implemented multiple latency reduction algorithms in the autonomous vehicle distributed operating system's message-passing protocol

Select Publications

"Accelerating Scientific Applications With SambaNova Reconfigurable Dataflow Architecture" *Computing in Science & Engineering*, 2021

M. Emani, V. Vishwanath, C. Adams, M. E. Papka, R. Stevens, L. Florescu, S. Jairath, **W. Liu**, T. Nama, A. Sujeeth

"What Your DRAM Power Models Are Not Telling You: Lessons from a Detailed Experimental Study"

SIGMETRICS 2018.

S Ghose, A G Yağlıkçı, R Gupta, D Lee, K Kudrolli, **W X. Liu**, H Hassan, K K. Chang, N Chatterjee, A Agrawal, M O'Connor, O Mutlu.

Select Coursework

Theory: Topics in Deep Learning, Parallel Algorithms, Functional Programming, Numerical Analysis

Systems: Optimizing Compilers, Parallel Computer Architecture, Advanced Operating Systems

Skills

Engineering:

Languages: Python, TypeScript, C, C++; **ML Tools:** Tensorflow, PyTorch, NumPy; **Compiler Tools:** LLVM, MLIR

Web/App Development:

Frontend: React, NextJS, Flutter; **Backend:** Fastify, Flask, FastAPI, Express, GraphQL; **Databases:** MongoDB, PostgreSQL, DynamoDB, Firestore; **Infrastructure:** AWS, DigitalOcean, Firebase, Heroku, Docker

Design:

Tools: Figma, Adobe InDesign, Adobe Photoshop, Adobe Illustrator