

Ethan Wu

ewu@u.northwestern.edu ❖ (978) 493-4946 ❖ github.com/ethqnol ❖ linkedin.com/in/ethan-wu-3b908a2a4/

Looking for summer 2027 intern opportunities for compiler and systems level engineering

EDUCATION

Northwestern University

Computer Science & Operations Research (IEMS)

September 2026 – June 2030 (Anticipated)

Evanston, IL

- **NU Formula Racing (FSAE):** Firmware engineer; using C to build high voltage battery management system.

SKILLS

Languages: Python, C++, Mojo, Rust, Go, JavaScript

Tooling & Frameworks: PyTorch, LLVM, Cranelift, Google OR-Tools, Jax/Flax, LightGBM, Optuna, Docker

Knowledge/Domains: Multi-Threading, Optimizations, Clustering Algorithms, ML, Git, Linux, SIMD, C-FFI

WORK EXPERIENCE

Research Intern

MIT Jameel Clinic

October 2024 – Present

Cambridge, MA

- Performed clustering analysis on public health data; **co-authoring 2 in-progress papers** targeting PLOS One.
- Developed novel method to find the best clustering from clusterings generated with different hyperparameters.
- Replicated & verified cancer prediction computational models in published papers.
- Co-designed a **custom featurizer** for a cancer prognostic machine learning model with an MIT PostDoc.
- Had several of my projects featured in an MIT blog post: [See Blog Post](#)

PROJECTS & OPEN SOURCE CONTRIBUTIONS

Mojo Compiler & STL [Contributor] - MLIR, Mojo, C++, LLVM

<https://github.com/modular/modular/pull/7311>

- Fixed LLVM codegen bug by mapping SIMD comparisons to FCMP UNE, correcting vector NaN evaluation
- Wrote tests validating IR emission and SIMD lane boolean reduction behavior

polars-mojo [Author] - Mojo, Apache Arrow, SIMD, Polars

<https://prefix.dev/channels/modular-community/packages/polars-mojo>

- Built zero-copy interop between Mojo and the Rust Polars dataframe library via Apache Arrow C Data Interface
- Implemented an RAII wrapper to synchronize Mojo `__deinit__` with Arrow C release callbacks to prevent leaks.
- Implemented SIMD reduction kernels (sum, mean, std) via Mojo's vectorize primitives.

Strata [Author] - Mojo, LAPACK/BLAS, SIMD

<https://prefix.dev/channels/modular-community/packages/strata>

- Developing a pure Mojo machine learning library with Sklearn ergonomics and zero Python runtime dependencies.
- Leveraged strict traits, static polymorphism, and generic SIMD types across float/int bit-widths to optimize ML models (Linear models, Random Forests, K-Means). Average of 1.9 times faster than Sklearn. [See benchmarks](#)
- Applied LAPACK/BLAS to handle linear algebra computation.

Salmon Programming Language [Author] - C++ 20, LLVM, Clang

<https://github.com/ethqnol/salmon>

- Built an end-to-end compiled language called Salmon, that targets native machine code via LLVM IR
- Implemented SSA code-gen using LLVM C++ API with scope management & dynamic list runtime structs.
- Designed deterministic resource cleanup via a LIFO defer unwinding stack.

Kluster-Fudge (PyPI) [Author] - Python, JIT, GPU

<https://pypi.org/project/kluster-fudge/>

- A Python library for clustering categorical data using K-Modes with various distance metrics (**3000+ downloads**).
- Implemented multi-threading, GPU acceleration, and JIT optimizations for better performance.

CubeCL (TracelAI) [Contributor]

<https://github.com/tracel-ai/cubecl>

- Contributed upstream bug fix to CubeCL, a multi-platform GPU compute language extension for Rust.

Modular Community & Prefix Ecosystem (Modular AI) [Contributor]

<https://github.com/modular/modular-community/pull/348>

- Resolved critical packaging automation failure by debugging and patching CI/CD scripts for the Mojo toolchain.

