

Gavin Ratcliff

Austin, TX | gavin_ratcliff@tamu.edu | [GitHub](#) | gavinratcliff.com

EDUCATION

Texas A&M University

Bachelor of Computer Science & Applied Mathematics

GPA: 4.0

College Station, TX

May 2027

EXPERIENCE

Independent Funded AI Safety Research

May 2024 – August 2024

Research Engineer

- Researched novel methods for detecting maliciously trained large language models, building on top of open-source models, leading to a 20% improvement over SOTA detection methods.
- Implemented our novel method as an open-source Python library allowing other AI developers and researchers to test models for malicious behaviors with ease, making it easier to safety test LLMs.
- Secured \$5k+ in grant funding to accelerate research with high-performance NVIDIA GPUs and built custom tooling for our cluster that improved training speeds by more than 80%.

Technologies: Python, PyTorch, CUDA, NumPy, Transformers, Deep Learning.

ARENA AI Safety Research Accelerator

December 2023 – January 2024

Student Researcher

- Developed a novel algorithm for automated circuit discovery, improving the speed of behavior tracing in large language models by 10-20%, significantly enhancing interpretability capabilities.
- Completed a curriculum replicating 10+ key research papers in mechanistic interpretability, reinforcement learning, and transformer training.

Technologies: Python, PyTorch, CUDA, NumPy, Transformers, Deep Learning.

LLVM Contributor

June 2021 - February 2023

Open-source compiler developer

- Contributed code improving correctness of the C compiler Clang to the LLVM Project, one of the largest open-source compiler projects, funded by industry leaders like Intel, Google, and Apple,
- Worked with industry employed developers to get my code reviewed, refined, and merged into the mainline LLVM codebase, adding to the 2+ million lines of code in the project.
- Taught myself extensively about compiler development and internals using dozens of resources and textbooks from my school library, picking up lots of math and CS fundamentals along the way.

Technologies: C++, C, compilers.

PROJECTS

Sourcerer, an Amazon storefront tracker.

- Developed a full-stack web application using Flask, React, and MySQL, enabling users to track and manage lists of Amazon storefronts, with real-time webhooks and daily summaries of product changes
- Serves multiple clients daily, tracking over 1,000 storefronts and sending an average of 700 webhooks per day, fully deployed on AWS using Docker for seamless scalability and reliability
- Collaborated closely with clients to continuously refine and enhance the product, delivering new features and improvements over a 2-month development period.

Technologies: Python, JS, Flask, React, MySQL.

[The Wist Programming Language](#)

- Designed and implemented a custom compiler and virtual machine in C for a novel programming language, built as a modular library from scratch, consisting of over 20,000 lines of code with high efficiency.
- Integrated advanced features such as garbage collection, full-program Hindley-Milner type inference, and an untyped lambda calculus intermediate representation.

Technologies: C, ML (as in OCaml and Haskell), compilers, interpreters.

SKILLS & INTERESTS

Technical Skills: C, C++, Python, JavaScript, OCaml, Lisp, OpenGL, CUDA, NumPy, PyTorch, Flask, React, SQL, AWS, GCP.

Interests: Compilers, deep learning, transformers, signal processing, real-time graphics.