

Henry Huarte Toll

NYC, NY • 917.655.8607 • henryhuartetoll@gmail.com
linkedin.com/in/henry-toll/ • github.com/henrycashe26

EDUCATION

Cornell University, College of Arts & Sciences

Bachelor of Arts in Computer Science

GPA: 3.64 (Dean's List)

Ithaca, NY

May 2026

Related Courses: Computational Linguistics, Computer System Organization and Programming (C/Assembly), Operating Systems (C), Operating System Practicum (C), Machine Learning, Computer Networks, Combinatorics, Abstract Algebra, Manifolds and Differential Forms, Analysis of Algorithms

SKILLS

Programming Languages: Python, C++, Rust

Knowledge: Abstract Algebra, Combinatorics, Object-Oriented Programming (OOP), ML, MCP, K8S, Distributed Systems, PyTorch

WORK EXPERIENCE

Snorkel AI

San Francisco

Intern, AI Startup

Summer 2025

- Single-handedly developed and shipped a natural language upload data assistant automating uploading and data correction by building an agent that uses custom MCP servers, RL, and deploying to K8S.
- Collaborated on building a customizable AI simulation engine built the tool calling architecture, S3 API, and tool call tracing.

Anote AI

New York

Intern, AI Startup

Summer 2024

- Developed chunking techniques to enhance embedding functions for Retrieval-Augmented Generation (RAG) models on financial data.

FinTech Collective

New York

Intern, FinTech VC Firm

Summer 2024

- Collaborated with startup founders and venture capital insiders, analyzing market trends and investment opportunities.
- Delivered an investment pitch based on comprehensive data analysis, resulting in actionable insights for an eventual investment in a Series A 15.5m round.

Smash Capital Ventures

New York

Intern, Later Stage Venture Capital Firm

Summer 2023

- Conducted market research and technological assessments to support investment decisions in software and internet companies.

Project Experience

Agent RLHF Custom 14b Parameter Model

September 2025 - Present

Personal Research Project

Online

- Collaborated with a graduate student to train a 14B-parameter language model (roughly sota model size in 2023) on cloud GPUs, using new actor/critic/judge architecture as an alternative to standard RLHF.
- Building inference API following OpenAI standard and deploying custom inference server via Kserve

AI Generated Video Platform

July 2025 - Present

Small Team project

Online

- Collaborated with a 4-person team (including FAANG SWE and online influencer) to build a generative video platform for text, images, and 3D meshes.
- Designing backend engine using LangChain for model orchestration layer to support customizable video synthesis with multimodal inputs.

Kernel From Scratch

February 2025 - May 2025

Operating Systems Pracilium, Cornell University

Ithaca, NY

- Collaborated with a team of 3 in building core OS components from scratch in C including memory allocation, scheduling, signals, and context switching.

Leadership and Interests

Senior Advisor, Cornell Extended Reality Project Team

February 2024 - Present

ExSoftware lead + founding member of Cornell XR (Extended Reality) Project Team

- Led a team to build computer interaction software for VR with Stardust XR, and facial recognition for AR glasses

Interests: Flamenco Guitar, CPU member (Cornell Political Union), Catan (co-founder of the Catan Club)