

LARRIS XIE

larris.xie@uwaterloo.ca · 437-974-6166 · larris.me [↗](#) · LinkedIn [↗](#) · GitHub [↗](#)

EDUCATION

University of Waterloo

Bachelor's of Computer Science

Scholarships: *Nortel Institute, President's Scholarship, President's Research Award* x2 (\$10k)

Sep 2024 - Apr 2028

4.0/4.0 GPA

EXPERIENCE

Conway

Software Engineering Intern

Incoming May 2026 - Aug 2026

- Selected as 1 of 45 **Kleiner Perkins Fellows** from 5000+ candidates

Shopify

Software Engineering Intern

May 2025 - Aug 2025

- Orchestrated a **100% uptime data migration** from legacy APIs to new service endpoints while ensuring full **backwards compatibility** through shadow writes, backfilling **over 5B rows**, and staged rollouts.
- Designed and implemented scalable **data models and GraphQL APIs in Ruby on Rails**, powering a fully extensible cash management experience for **2M+ active retailers**.
- Engineered full-stack features for Shopify Point of Sale (POS) in **React Native and Typescript**, resolving critical race conditions in async workflows to reduce failure rate by **40%**.

Meta x MLH

Production Engineering Fellow

May 2025 - Aug 2025

- Deployed and managed **Linux infrastructure** on CentOS, leveraging system calls and kernel-level debugging at the OS layer, and Docker and NGINX at the application layer to improve reliability.
- Configured system monitoring with **Prometheus, cAdvisor, and Grafana**, leveraging low-level resource metrics (CPU, cache misses, memory usage, I/O throughput) to cut bug detection time by **80%**.
- Automated **CI/CD workflows** with GitHub Actions, integrating testing frameworks, security scans, and performance monitoring to streamline deployments and improve system reliability.

University of Waterloo

Undergraduate Research Assistant — Security and Privacy in ML

Jan 2025 - Present

- Built a privacy-preserving **vertical federated learning** system for fraud detection, leveraging a GAN and encrypted communication protocols to enable secure cross-party training on distributed time-series data.
- Achieved a 95% F1 score under a 10000:1 class imbalance by training custom **transformer models in PyTorch** under federated learning and performing large-scale hyperparameter optimization.
- Engineered a **distributed secure inference** pipeline using Facebook's CrypTen MPC and PyTorch's RPC, providing membership inference resistant serving on GPUs and maintaining **100% of performance**.

WAT.ai (Design Team)

Machine Learning Engineer

Jan 2025 - Jul 2025

- Built an AI assistant for independent and safe senior living, integrating OpenCV-based hazard detection with a real-time voice interface powered by **Whisper and LangChain**.
- Engineered an **agentic RAG pipeline** to create curated profiles for seniors, achieving **90%** accuracy in user intent recognition and enabling seamless automation of everyday tasks.

PROJECTS

ASCII Game Engine [↗](#) C++

- Architected a 2D game engine in C++ from scratch, grounded in **MVC architecture and SOLID principles**, and leveraging **advanced design patterns** for modular and efficient resource management.
- Features **OOP** in console-based rendering, event-driven input and audio, collision detection, dynamic world state with multi-layered rendering and animation, with two demo games: Flappy Bird and Space Invaders.

On The Dot (3000+ users) [↗](#) | React, TypeScript, Python, PostgreSQL, AWS

- Launched a trivia game with persistent game state and anonymous user tracking—**10k plays in 3 months**.
- Architected a Python backend with custom **PostgreSQL** functions to achieve **20ms read/write latency** on **30K queries/day** for real-time statistics tracking and rate control.

LifeOS (Hack the 6ix Winner) [↗](#) | React, TypeScript, Python, FastAPI, Redis, Docker, PostgreSQL, AWS

- Engineered a **distributed AI video processing pipeline** using Redis Queue that orchestrates semantic analysis with TwelveLabs, embedding generation, and cloud storage via **asynchronous workers**.
- Deployed a **containerized API service with AWS** using Docker, FastAPI, Pydantic models, and JWT middleware; supports modular features like authentication, video ingestion, memory search, and chat.