

## **Yanming Luo (U.S.Citizen)**

298287 Sandburg Way, Hayward, CA, 94544(Bay area) | 209-285-4088 | [luoyanming99@gmail.com](mailto:luoyanming99@gmail.com)

### **Education**

University of California, **Davis (UCD)** : Bachelor of Science in **Computer Science** | **GPA: 3.76** | Graduate in 2025  
**Skills**

**Programming Languages:** C, C++ (desktop & system programming), Java, Python (AI/ML, pipelines, automation), C#, Go, JavaScript/TypeScript, Shell Script, SQL, Redis, Ruby on Rails

**Systems & Embedded:** Real-time systems, concurrency, distributed storage, fault-tolerant design, ROS/ROS2

**Cloud & DevOps:** AWS (S3, EC2, Elastic Beanstalk, EKS), Docker, Kubernetes, Jenkins, Git CI/CD, Terraform

**Web & UI:** React.js, Node.js, Express.js, RESTful APIs (OpenAPI), HTML/CSS/JavaScript, Spring Boot, SwiftUI

**Testing & QA:** TDD, Pytest, JUnit, stress testing, functional testing, CI/CD pipelines, Flutte, Cursor

**Systems/Cloud:** Docker, Kubernetes, AWS, GCP, WebAssembly, Redis, Spark, Git, Kafka

**Other:** Blockchain-based security, ffmpeg/Video codec standards (H.264, H.265, AV-1, exposure), 3D simulation & visualization (ROS2 RViz, WebGL-like pipelines); familiarity with game engine workflows

### **Work Experience**

**EnerNova Tech | Full-stack Software Engineering summer Intern** 05/2024 – 09/2024

- Built **B2B SaaS** dashboards and **full-stack web** features using **Java (Spring Boot)**, **TypeScript (React, Node.js, NestJS)** and **REST APIs with PostgreSQL/MySQL, Redis, and Kafka**, delivering 0→1 dashboards with secure authentication and fault-tolerant workflows.
- Worked with **microservices, GraphQL, and REST APIs** to support scalable **Leveraged agentic coding tools** applications, aligning with enterprise-grade development standards.
- Developed and automated test suites in **Python (Pytest, JUnit)**, applying **TDD and regression** testing for quality assurance, participating in **peer code reviews** to ensure quality and learning safe update practices for production systems.
- Integrated automated tests into **CI/CD** pipelines (**Jenkins, Git**), enabling rapid, reliable software delivery.
- Deployed microservices on **AWS and GCP** with **Docker** and **Kubernetes**, applied **CloudWatch**
- Worked in an **Agile/Scrum** environment with sprint planning, code reviews, and cross-functional collaboration to deliver secure, production-ready web applications.
- Prototyped mobile UI components with **SwiftUI** and **React Native** to extend web features into a **mobile-friendly** interface, gaining exposure to cross-platform development.

**UC Davis Deep Learning Lab - Text Summarization System**  [GitHub Repository](#) 01/2025 – 06/2025

- Applied **Python**, **PyTorch**, and **scikit-learn** for preprocessing, experimentation, **ablation studies**, and visualization, documenting workflows to guide research, Integrated **LLMs and Transformers** (Hugging Face, Langchain) for summarization and retrieval tasks, aligning with applied **ML in real-world AI** systems.
- Engineered reproducible ML pipelines (**PyTorch, Spark/SQL, CUDA**) on **Linux** and **Jetson/ARM SoCs**, reinforcing embedded real-time reliability and secure data workflows.
- Containerized test suites with **Docker and Kubernetes** for scalable **stress testing** and **CI/CD** integration, producing compliance-style **QA reports** for **mission-critical** validation.
- Optimized inference with custom **CUDA kernels**, bridging **LLM-based** summarization with workflow automation to streamline compliance reviews and reproducible systems.
- Highlighted applications for **content summarization, search/retrieval, and personalization**, aligning with real-world workflow automation and media AI.
- Applied **ML** methods for **content summarization and computer vision evaluation**, aligning with **AI-driven** analysis for real-time media and tracking systems.
- Collaborated in a large, shared research team with all **professor and PhDs**, navigating complex modules and contributing through **documented** workflows, reviews, and reproducible experiments.

### **Projects**

**DreamLayer AI (Open-source ML project)** –  [GitHub Repo](#)

- Contributed to an **AI-first platform enabling workflow automation for creative productivity**, integrating **Runway Gen-4 with ComfyUI** to build AI-native assistants that streamline generation pipelines.
- Implemented **custom nodes and APIs** in TypeScript/Python, showcasing ability to extend AI platforms with secure and reusable modules.
- Built GenAI pipelines with Langchain-style orchestration, extending platforms with secure, reusable modules for content search, summarization, and personalization. Collaborated in a **startup-like environment** to rapidly ship production-ready GenAI features, highlighting **AI-driven assistant systems** for content creators.

**Distributed File System**  [GitHub Repo](#) **(C++ project)**

- Developed and tested **C++ desktop-style system tools** with multi-threading and real-time safety validation, aligning with high-performance application development.
- Applied **Unix file system** principles (inodes, block allocation, dynamic storage management) with emphasis on fault tolerance, metadata parsing, and secure data handling.
- Engineered request/response lifecycle handling—parsing HTTP requests, routing them through REST endpoints, and demonstrating how a service processes an HTTP request end-to-end.

**Real-Time Remote Control & Streaming System** **(AI & Robotics Project)**

- Built a **real-time C++/Python motion-tracking and streaming pipeline** for robots, integrating 3D simulation (ROS2/Autoware) and visualization tools comparable to game engines (Unity/Unreal workflows).
- Collaborated in a multi-disciplinary team, mirroring cross-functional stakeholder communication common in government and enterprise environments.