

Zheng H. Chen

Clackamas, OR | chenz22@oregonstate.edu | 971-340-8699 | linkedin.com/in/zhenghchen | github.com/zhenghchen

EDUCATION

Oregon State University

BS Computer Science, GPA: 3.96

Corvallis, OR

Sept 2024 – June 2028

- **Relevant Coursework:** Data Structures, Web Dev, Computer Architecture, OOP, Linear Algebra, Discrete Math

EXPERIENCE

AI Research Internship | Python, PyTorch, HuggingFace, SLURM

June 2025 – Sept 2025

Oregon State University

Remote

- Drove a 117% performance increase in the team's Gemma-3-4B baseline model (from 24.80 to 54 BLEU), diagnosing and engineering a data preprocessing solution that appended end-of-sequence (EOS) tokens.
- Automated the model evaluation lifecycle by integrating real-time SacreBLEU scoring into training scripts, reducing evaluation runtime from over an hour to minutes and streamlining SLURM job submission.

Undergraduate Research Assistant | Python, PyTorch, HuggingFace, SLURM

Oct 2024 – June 2025

Oregon State University

Corvallis, OR

- Accelerated large language model (LLM) pretraining by 200% through the implementation of FlashAttention-2, reducing attention memory complexity from $O(N^2)$ to $O(N)$ via custom CUDA fused kernels.
- Benchmarked training throughput and memory usage on NVIDIA H100 GPUs, enabling efficient scaling for multi-billion parameter models.

PROJECTS

MCP-Guardian | Python, FastAPI | GitHub

Sept 2025 - Oct 2025

- Delivered a security middleware for AI Agent ecosystems, enabling secure tool access via a single decorator-based API. Implemented configurable JWT authentication, scope-based authorization, and rate limiting.
- Established a security audit trail with structured JSON logging for real-time monitoring and compliance. Captured detailed authentication and authorization events, providing security teams with actionable intelligence for rapid incident response.
- Published the middleware as a Python package to PyPI, reducing developer integration time from hours to minutes. The package included comprehensive documentation and real-world examples, resulting in a zero-friction installation.

Redflagged | HackMIT | TypeScript, React, Python, Tailwind CSS | redflaggy.vercel.app | GitHub

Sept 2025

- Architected a full-stack legal document analysis platform leveraging a React/TypeScript frontend, Python backend, and Supabase for authentication and vector storage.
- Implemented a secure RESTful API with FastAPI for document ingestion, preprocessing, and AI-powered risk extraction. Enabled scalable handling of 500,000 character PDF/DOCX files and dynamic JSON-based risk visualization.
- Developed a Retrieval-Augmented Generation (RAG) pipeline utilizing HuggingFace Xenova and LangChain for semantic chunking and embedding generation. Orchestrated search over Supabase pgvector, supporting top-k retrieval.
- Built a companion Chrome extension for seamless in-browser document highlighting and analysis, enabling direct user interaction and feedback submission via authenticated API endpoints.

AI Writing Assistant (Chrome Extension) | Java Spring Boot, React, TypeScript | GitHub

July 2025 - Present

- Designed and implemented a robust RESTful backend using Java Spring Boot, supporting secure user authentication and encrypted API key management to facilitate cross-device synchronization and data protection.
- Built a Chrome extension frontend with React and TypeScript, integrating real-time chat functionality and a comprehensive settings dashboard for streamlined user and API key administration.
- Developed a modular backend architecture serving as a provider-agnostic proxy, enabling seamless integration and dynamic switching between multiple LLM providers without frontend modifications.

TECHNICAL SKILLS

Languages: C++, Python, Java, JavaScript/TypeScript, Dart, SQL, HTML/CSS, Bash

Frameworks/Libraries: React, Spring Boot, Next.js, PyTorch, HuggingFace, Flutter, LangChain, Vite

Tools/Cloud: Git, Docker, AWS, GCP, Firebase, Supabase, Vercel, Jira

Databases: PostgreSQL

LEADERSHIP

AI @ OSU Club

June 2025 – Present

Officer/Event Coordinator

Corvallis, OR

- Co-led a recruitment drive at a university involvement fair, onboarding nearly 70 new members in a single day.