

Chao-Ting (Tim) Chen

+1 (650)248-0479 | chaotingchen10@gmail.com | github.com/GNITOAHC | linkedin.com/in/chaoting-chen | chaoting.me

EDUCATION

Stanford University

M.S. in Electrical Engineering

Stanford, CA

Sep 2026 — Present

National ChengChi University

B.S. in Computer Science, Double Major in Digital Content and Technology

Taiwan, Taipei

Sep 2021 — Jun 2025

- **Honors:** Dean's List; Best Special Topics in Computer Science | **GPA:** 4.2/4.3
- **Teaching Assistant:** Natural Language Processing and Computer Programming

TECHNICAL SKILLS

- **Programming:** Python, C/C++, Rust, Go, TypeScript, Shell | **ML:** PyTorch, TensorFlow, TensorRT, CUDA, Transformers
- **Systems & Data:** SQL, Git, UNIX/Linux, Docker, Kubernetes, GCP, NGINX, MongoDB, Cassandra | **Web:** React, Svelte, Vue

INDUSTRY EXPERIENCE

Data Center Group AI System Intern

Jul 2025 — Jul 2026

Intel Corp.

Taipei, Taiwan

- Analyzed performance and power data from 2,000+ ML workloads across CPU/GPU platforms, identifying system bottlenecks and informing optimization decisions.
- Improved distributed LLM-serving throughput by 50% and reduced latency by 30% by designing a pipeline that disaggregated prefill and decoding across 3 compute nodes.

Software Engineer Intern

Jul 2024 — Dec 2024

MediaTek Inc.

Taipei, Taiwan

- Reduced peak CPU load by 20% by developing and deploying a reinforcement-learning workload predictor that used runtime signals to drive resource-allocation decisions in [MediaTek Adaptive Gaming Technology](#).
- Built a production RAG/LLM knowledge system over engineering data, enabling engineers to retrieve technical context, diagnose failures, and accelerate debugging.

RESEARCH EXPERIENCE

Academia Sinica

Jul 2024 — Aug 2026

Scholarship Researcher · Advised by Prof. Hen-Hsen Huang

- Fine-tune, distill and train VLAs for the [TAIDE Project](#) Basic Model Team, developing efficient inference-serving pipelines for evaluation and deployment.
- Published **3 peer-reviewed papers** at top conferences, spanning CAG, efficient LLM decoding, and multimodal generation.

National Chengchi University

Mar 2024 — Feb 2025

Student · Advised by Prof. Ming-Feng Tsai

- Built an end-to-end legal AI system spanning data ingestion, retrieval, model inference, and a full-stack application, achieving 80% citation accuracy over Taiwan's laws database.

OPEN SOURCE & PROJECTS

Contributor, Ray

<https://github.com/ray-project/ray>

- Contributed production fixes to Ray Serve's autoscaling and request-routing subsystems, resolving state-consistency bugs during deployment teardown and replica failures.
- Worked within Ray's upstream codebase to trace cross-component lifecycle behavior, respond to maintainer review, and add regression tests covering both failure modes.

Independent Developer, Kamikami.shop

<https://www.kamikami.shop>

- Built a production e-commerce platform with Typescript, Golang, Python and PostgreSQL, supporting products variants, customer accounts, carts, checkout and review integrations.
- Designed and deployed services on multi-node Docker Swarm with self-hosted VPS and Caddy reverse proxying. Develop automated pipelines for generating and managing product assets, reducing repetitive work.

PUBLICATIONS

3 peer-reviewed papers · IJCAI 2024 · WWW 2025 · EMNLP 2025 · [Google Scholar](#)