

ZACHARY SHERMAN

Sunnyvale, CA | +1 408 839 4874 | sherman.zachary.j@gmail.com

SKILLS

Languages / Frameworks: Python, Rust, Go, TypeScript, React, SQL, C++, Bash

AI / ML Systems: Transformer fine-tuning, personally identifiable information code detection, synthetic data generation, embeddings, RAG, NER, agentic workflows, tool-calling agents

Infrastructure: NVIDIA Triton, Kubernetes, Docker, AWS, GCP, PostgreSQL, Redis

Developer Tools: Git, Linux, Claude Code, Codex

Scientific / Hardware Engineering: Signal processing, RF filtering, CAD, Arduino, experimental data analysis

EXPERIENCE

Dymium, Data Access and AI Security Startup | Los Gatos, CA

Software Engineer | Full-time Jun 2025–Present

Engineering Intern | Part-time Feb 2025–Jun 2025

- Led development of sensitive data and code detection systems using fine-tuned DeBERTa and ModernBERT, outperforming AWS Comprehend and Guesstang; mentored interns on fine-tuning and synthetic data generation.
- Deployed GPU-accelerated NVIDIA Triton inference on Kubernetes; optimized production model serving to improve GPU utilization and reduce latency.
- Built Rust/C PostgreSQL foreign data wrappers to securely federate customer data from external systems into queryable Postgres interfaces.
- Developed AI-enhanced ingestion pipelines for tabular, file, audio, and unstructured data, adding metadata enrichment and processing tools for downstream search and analysis.
- Built full-stack agentic chat workflows across React, TypeScript, and backend services, enabling natural-language querying, tool use, data processing, web search, and automated API/database access.
- Collaborated across engineering, product, and leadership to evaluate technical tradeoffs, debug production issues, and deliver customer-facing AI security workflows.

Allen Laboratory, Nanoscale Electronics Lab | La Jolla, CA

Research Assistant | UC San Diego | Apr 2024–Jun 2025

Part-time academic year; full-time summer 2024

- Developed Python software to analyze dense experimental datasets, including custom fitting routines for extracting electronic properties from quantum material measurements.
- Built automated lab equipment using 3D-printed components and motor control to prepare ultrathin 2D material samples.
- Engineered a 20-channel low-pass filtering system for millikelvin quantum transport measurements.
- Performed device fabrication and assembly, including mechanical exfoliation, heterostructure stacking, and CAD-designed Hall bar geometries.

Astronomy Club, UC San Diego | La Jolla, CA

Vice President and Board Member | Sep 2023–Jun 2025

PROJECTS

RAG-Based Academic Synthesizer | Prof. Brian Keating | Python, OpenAI API, LangChain

- Built a conversational RAG interface for querying academic books and papers.

Astronomical Object Classifier | Prof. Peter Ho | Python, PyTorch, NumPy

- Trained PyTorch models to classify simulated astronomical objects against benchmark data.

PUBLICATIONS / CONFERENCES

“Fabrication and Characterization of 2D EuSn₂As₂ Devices” | UC San Diego Summer Research Conference | Aug 2024
Presented on Hall bar device fabrication and magnetic topological insulator experiments.

EDUCATION

B.S. Physics, University of California, San Diego

Computational Physics Specialization | Minor in Mathematics | GPA: 3.8 | Jun 2025