# ITANA TOFT

San Francisco Bay Area - US Citizen

+1-808-747-5517 ✉ kitanatoft@gmail.com ⌗ github.com/kctoft 🌐 kitanatoft.com    linkedin.com/in/kitana

## Experience

**Super Micro Computers**
July 2 23 – Present

*Technology Enablement Engineer*
*San Jose, California*

- Deployed production-ready AI applications on next-gen hardware including NVIDIA B200, RTX Pro 6000 BSE, H200/H100, A100, and Intel Gaudi 3/2 accelerators for enterprise data center environments
- Coordinated cross-functional technical demos and product launches across engineering, marketing, and sales teams for two consecutive quarters
- Produced 12+ technical video tutorials for YouTube generating 100K+ views, with top videos reaching 15.7K views demonstrating AI deployment workflows on next-gen hardware, handling full production pipeline from scripting to post-production
- Created technical video tutorials for YouTube that demonstrate AI deployment workflows, handling full production pipeline from scripting to post-production
- Built scalable deployment pipelines using PyTorch, TensorFlow, LangChain, and Docker, enabling rapid model iteration and testing

**SproutLabs Smart Irrigation**
September – December 2 22

*Software Engineer — Apprenticeship*
*Santa Cruz, California*

- Developed full-stack smart irrigation platform with admin APIs, microservices architecture, and automated CI/CD deployment
- Led 3-person development team as Project Manager, coordinating delivery of production-ready microservices across 3 sprints using Agile methodologies with daily stand-ups
- Conducted weekly stakeholder meetings to address pain points, gather requirements, and align deliverables with business objectives

## Projects

**Multi-Modal AI Applications**
2 23 – 2 24

*Vision-Language Models on Supermicro Infrastructure*
*Python, PyTorch, Transformers, Gradio*

**AI Image    Video Generation Suite**
2 23

*Stable Di usion XL & Video on Intel Gaudi 2*
*Python, PyTorch, Gradio*

**LLM Chatbot with RAG**
2 23

*Llama 3 70B on Supermicro Gaudi 2 AI Server*
*Python, PyTorch, LangChain, Docker*

**CruzCal - Academic Scheduling Tool**
2 22

*Senior Capstone Project*
*Next.js, React, TypeScript, Jotai, Jest*

## Education

**University of California, Santa Cruz**
Graduated December 2 22

*BSE in Computer Engineering with Honors, Computer Science Minor*
*Santa Cruz, California*

**Foothill College**
Graduated June 2 2

*AS in Computer Science, Engineering, & Mathematics*
*Los Altos, California*

## Technical Skills

**AI/ML Frameworks**: PyTorch, TensorFlow, Transformers, LangChain, LlamaIndex, vLLM, Gradio, CUDA, oneAPI
**LLMs & Models**: Llama 3/3.1/3.2, GPT-4/3.5, Mistral, Stable Di usion XL/Video
**Multi-Modal Models**: GPT-4V, LLaVA, CLIP, BLIP
**Hardware**: NVIDIA B200, RTX Pro 6000 BSE, H200/H100, A100; Intel Gaudi 3/2, Max Series, Flex Series
**Languages**: Python, Golang, TypeScript, JavaScript, C++, C, Java, SQL, Conversational Japanese
**Developer Tools**: Docker, Kubernetes, Git, VS Code, Jupyter, GCP, Firebase, Linux
**Web Technologies**: React, Next.js, Svelte, FastAPI, Node.js
**Content Creation**: Video Production, Cinematography, Adobe Premiere Pro, Technical Writing
**Project Management**: Agile, SCRUM, Cross-functional Leadership, Stakeholder Management

## Media Engagement    Technical Communication

**YouTube Technical Tutorials**: Produced video series on AI deployment and optimization

- Stable Video Di usion Tutorial
- Stable Di usion XL on Gaudi 2 AI Server
- Llama 3 70B Chatbot Implementation

**Technical Presentations**: Delivered live demonstrations of AI capabilities for diverse audiences
**Certi cations**: Project Management Practitioner (De Anza College)