Yousuf Golding

San Francisco Bay Area

+16507137793 | yousuf@goldin.gs | https://yoolee.xyz/ | yooleee | yousuf-golding

Personal Profile

I'm a recent NLP master's graduate from UC Santa Cruz who scopes, builds, and ships end-to-end LLM systems, from evaluation harnesses to full-stack applications. My background spans applied NLP, LLM evaluation methodology, and AI in educational technology. I have spent years making complex scientific ideas accessible to non-specialist audiences, and I am drawn to using technology for social impact in mission-driven settings.

Education

University of California, Santa Cruz *San Jose, California*

MS in Natural Language Processing (NLP) *Sept 2024 - Dec 2025*

- **Courses:** Capstone (1, 2), Deep Learning for NLP, Data Science and Machine Learning Fundamentals, Advanced Machine Learning for NLP, Conversational Agents, Projects in AI, Natural Language Processing (1, 2, 3)

University of California, Santa Cruz *Santa Cruz, California*

BS in Applied Mathematics *Sept 2020 - June 2024*

- **Courses:** Intro to Probability Theory, Intro to Proof and Problem Solving, Intro to Number Theory, Intro to Dynamical Systems, Mathematical Methods for Engineers, Foundations of Scientific Computing, Mathematical Modeling (1 & 2), Cryptography

Skills

Programming	Python (PyTorch, scikit-learn), LLM Systems (OpenAI APIs, embeddings, RAG, LangChain), FastAPI, React, PostgreSQL, C++, JavaScript, HTML/CSS
Miscellaneous	Git, LaTeX

Projects

SARchlight – Closed-Loop AI Search Planner for SAR Drones (Berkeley AI Hackathon 2026) *June 2026*

- Led a 3-person team over a 24-hour build of a closed-loop Bayesian search-planning system that fuses real terrain data (USGS DEM, ESA WorldCover) and lost-person behavior statistics into a live probability map that directs a simulated multi-drone search and localizes a missing person.
- Built the probability "brain" and integration loop: a terrain-aware prior, Bayesian updates from both detections and clean non-detections, and a persistence-gated "located" trigger robust to transient false positives. Localizes the subject at zero-cell error on real Marin terrain rasters in simulation.
- Added a swappable detector seam (geometry-driven simulator plus a YOLO11 + SAHI path), a multi-drone sector planner with zero-overlap coverage, a Deepgram voice layer (subject broadcast in English and Spanish, plus a Twilio operator phone agent on Fly.io), a live React dashboard, and Sentry observability, covered by roughly 300 tests.
- Stack: Python, NumPy, rasterio, FastAPI, YOLO11, React, TypeScript, Claude (Anthropic), Deepgram, Twilio, Fly.io, Sentry.

Dual Use Navigator (Apart Research AIxBio Hackathon) *April 2026*

- Built an open-source web application that ingests dual-use export control lists from three jurisdictions (US eCFR XML, EU Annex I and CELLAR XHTML, Australian DSGL 2024 EPUB) into a unified Pydantic schema with side-by-side regulatory comparison, multilateral regime overlays (Australia Group, Wassenaar, NSG, MTCR), and EU General Export Authorisation applicability.
- Covered 3,384 regulatory entries with primary-source attribution; identified a cross-source EU regulatory inconsistency (1C352 EUGEA case) not visible in any single source.
- Stack: Python, Flask, Pydantic, BeautifulSoup, Leaflet. Documented in a 10-page technical report covering methods, limitations, dual-use risks, responsible disclosure, and ethical considerations.

QUD Empowered LLM Co-Writing Tool (Capstone) *Sep 2025 – Dec 2025*

- Built, with a capstone team, a full-stack NLP system (FastAPI, React, PostgreSQL) that generates diverse LLM summaries, extracts implicit Questions Under Discussion (QUDs), and classifies student coverage using embedding similarity and LLM-as-a-judge verification.
- Designed the relevance-classification study and led a human evaluation with 100 participants across 10 articles, demonstrating statistically significant alignment between AI-ranked questions and human importance judgments ($r = 0.28$, $p = 0.002$).

HomeHelper – AI Agents for Appliance Troubleshooting *Apr 2025 – Jun 2025*

- Co-developed multi-agent system using LangChain/LangGraph and SIM-RAG for appliance repair support.
- Integrated Llama-3 for orchestration, BLIP for vision analysis, and a Flan-T5 critic trained on iFixit synthetic data; evaluated on MyFixit dataset and synthetic multi-turn dialogues.

EduMUSE – Multimodal Educational Assistant *Apr 2025 – Jun 2025*

- Built interactive study system using CrewAI, where users upload PDFs, highlight text, and trigger AI workflows (summarization, quiz generation, contextual QA, knowledge retrieval, podcast-style audio).

Comparative Analysis of LLMs for Mental Health Counseling *Jan 2025 – Mar 2025*

- Evaluated proprietary and open-source LLMs (ChatGPT, Claude, LLaMA, Deepseek) on counseling capabilities using classification and generation tasks with comprehensive metrics.
- Developed annotation guidelines and analyzed inter-rater reliability to systematically assess therapeutic response quality across models.

Work Experience

HigherSummit *Oakland, California*

Machine Learning Research Intern *Sept 2023 - 2024*

- Researched educational frameworks and theories compatible with AI/NLP methods: https://arxiv.org/abs/2402.01770
- **Technical Skills:** Python, NumPy, Git, LaTeX, Prompt Engineering, GPT

Frontier AI *San Francisco, California*

Machine Learning Engineering Intern *June 2023 - Sept 2023*

- Built full-stack quiz software using LLMs as part of framework for enterprise micro-learning.
- **Technical Skills:** Python, PyTorch, NumPy, Git, LaTeX, Prompt Engineering, GPT, Open AI APIs

Chabot Space and Science Center / NASA Ames Visitor Center *San Francisco, California*

Science Educator *2016 - 2020*

- Taught hands-on science to visiting families and school groups through the Galaxy Explorers volunteer program, adapting demos on the spot for a wide range of ages, abilities, and backgrounds.
- Designed and built physics and chemistry exhibits for public science education and outreach.
- Logged 676 volunteer hours over four years despite a demanding school schedule, earned the President's Volunteer Service Award (gold and silver tiers), and trained newer volunteers to run the program's shifts.

Publications

Extending Interactive Science Exhibits into the Classroom using Anthropomorphized Chatbots and Bloom's Taxonomy
Yousuf Golding
arXiv:2402.01770, 2024