

# Steven Au

PhD Applicant U.S. Citizen

sttau@ucsc.edu — +1 (831) 332-5989 — [linkedin.com/in/steven-au](https://www.linkedin.com/in/steven-au) — [github.com/Prowo](https://github.com/Prowo)

## RESEARCH INTERESTS

---

Natural Language Processing in Personalization and Graph-Based Retrieval, Conversational AI, Emotion AI, Computational Psychiatry, Social Computing for Special Education

## EDUCATION

---

### University of California, Santa Cruz

Master of Science in Natural Language Processing

Sept. 2023 — Dec. 2024

Cumulative GPA: 3.90/4.00

### University of California, Santa Cruz,

Bachelor of Science: Computer Science

Sept. 2015 — Jan. 2023

## EXPERIENCE

---

### Bytes of Mind Lab, Icahn School of Medicine

*Volunteer ML Researcher*

New York, NY (Remote)

Sept. 2025 — Present

- Built **PeReGrINE**, a personalized LLM review-generation benchmark using temporally grounded user-item knowledge graphs, VLM-augmented product captions, and neighborhood retrieval for product intelligence.
- Designed a healthcare NLP knowledge graph linking patient medication reviews, adverse effects, and clinical-trial evidence, using information extraction and entity normalization for cross-source treatment comparison.

### Eternos Inc. / Uare.ai

*Machine Learning Engineer, Forward-Deployed AI Systems*

Los Altos, CA

Apr. 2025 — Aug. 2025

- Created a client-specific knowledge base from 250+ long-form interviews, handling diarization, transcript cleaning, gold-label testing, and retrieval evaluation for digital twin workflows.
- Engineered automated ETL and orchestration pipelines in Python across AWS S3 and Databricks Lakehouse environments, reducing manual processing time by **40%**.
- Designed multimodal RAG pipelines on AWS Bedrock for text, audio, and video retrieval, improving retrieval relevance by **25%** through hybrid embedding optimization, prompt tuning, and RAGAS evaluation.

### Baskin Engineering, UCSC

*Web Development & Digital Communications Specialist*

Santa Cruz, CA

June 2024 — Jan. 2025

#### Front-End Engineer and ML Specialist - CruzChat

- Led front-end development for CruzChat, an AI tutoring application built with React and Next.js, integrating assistant-based RAG workflows and improving student engagement by **20%** for an engineering course.

#### Web Developer and Content Strategist - [nlp.ucsc.edu](https://nlp.ucsc.edu)

- Designed and maintained the Baskin Engineering website, ensuring accessibility and responsiveness.
- Increased site traffic by 25% by creating targeted marketing materials and web content that promoted departmental initiatives and research projects, leveraging effective SEO strategies.

#### Event Planner and Coordinator - Orientation, Graduation, BBQ

- Coordinated 60+ graduate student events, driving a 15% increase in participation by optimizing event planning and expanding communication outreach across multiple platforms.
- Reduced operational deficit by 15% by optimizing budgets, logistics, and communication processes, ensuring high participation in events while maintaining cost efficiency.

### Intel Labs

*Machine Learning Engineer Intern*

Santa Clara, CA

May 2024 — Sept. 2024

- First-authored research on personalized product review generation using query-based retrieval over user-item graphs, ratings, and top-*k* neighbor evidence for sparse user profiles.
- Tested BM25, dense retrieval, user-only, neighbor-only, and graph-based retrieval pipelines, improving ROUGE by **12%** over prior baselines.

## ACADEMIC PROJECTS

---

### Emotion Discovery and Reasoning its Flip in Conversation - SemEval Task 10 2024

*AI/ML Engineer - NLP*

Santa Clara, CA USA

Nov. 2023 — Feb. 2024

- Achieved a 0.95 weighted F1 score for emotion recognition using BERT and 0.79 for binary emotion flip reasoning with GRU embeddings on a Hinglish code-mixed dataset with 1 of 3 scored first.
- Poster presentation at the following venues:
  - NAACL 2024 - SemEval 2024, Mexico City, Mexico
  - UCSC-IIT Summer Research Symposium, Santa Cruz, USA.
  - UCSC AI Student Research Symposium, Santa Cruz, USA.

## PUBLICATIONS

---

Sujit Noronha, Steven Au. Blind Deference: A Diagnostic Benchmark for Authority-Driven Sycophancy in Multi-Turn LLM Conversations. *Accepted to the Conference on Language Modeling (COLM)*. 2026

Steven Au. Clinical Evidence and Patient Reviews: A Linked Dataset for Antidepressant Side Effects. *Accepted to the 25th Workshop on Biomedical Natural Language Processing (BioNLP) at ACL*. 2026

Steven Au. SoloSemantics at SemEval-2026 Task 4: Triplet-Tuned MPNet for Story Similarity. *Accepted to the International Workshop on Semantic Evaluation (SemEval)*. 2026

Steven Au. MIDI-PHOR: Multi-View Extraction and Captioning for Music Understanding. *In Proceedings of the International Workshop on NLP for Music and Audio (NLP4MusA) at EACL*. 2026

Steven Au, Cameron Dimacali, Ojasmita Pedirappagari, Nesreen K. Ahmed, Namyong Park, Ryan A. Rossi, Franck Deroncourt, and Yu Wang. Personalized Graph-based Retrieval for LLMs. *In Proceedings of the Pacific Asia Conference on Language, Information and Computation (PACLIC 39)*. 2025

Neng Wan, Steven Au, Esha Ubale, and Decker Krogh. UCSC NLP at SemEval-2024 Task 10: Emotion Discovery and Reasoning its Flip in Conversation (EDiReF). *In Proceedings of the 18th International Workshop on Semantic Evaluation (SemEval-2024)*, pages 1492–1497. *Association for Computational Linguistics*, 2024.

## AWARDS

---

**NLP Fellowship** Santa Clara, CA USA  
Awarded \$25,000 for UCSC NLP M.S. Program Sept. 2023

**Hack UCSC - 2nd in Innovation** Santa Cruz, CA USA  
Developed a wellness journal app in Java using an EEG headset for tracking brain waves Jan. 2017

## OTHER EXPERIENCES

---

**UCSC - Play, Performance, and Design** Santa Cruz, CA  
*Teacher Assistant - Intro to African American Theater* March 2024 — June 2024

- Taught and supported over 30 students by leading discussions, running classes, designing and grading assignments, and providing academic guidance through office hours.
- Collaborated with professors to enhance student engagement, manage classroom operations, and implement interactive learning strategies for improved understanding of course materials.

## SKILLS

---

- **Programming:** Python, SQL, Java, C/C++, HTML, CSS, MATLAB, LaTeX, R
- **Libraries/Frameworks** TensorFlow, PyTorch, scikit-learn, Pandas, NLTK, SpaCy, Keras, LangChain, Neo4j, RAGAS
- **Software & Tools** AWS, AWS Bedrock, S3, Databricks, Docker, Jupyter, GitHub, Git, Linux, Figma, Adobe Illustrator, Photoshop, InDesign, Tableau, LaTeX

## EXTRACURRICULAR

---

**Graduate Student Commons - UCSC** Santa Cruz, CA  
*Engineering Lead Peer Mentor* Dec. 2023 — Dec. 2024

**Graduate Student Association - UCSC** Santa Cruz, CA  
*Engineering Chair* Dec. 2023 — Dec. 2024