

Winson Chen

(949)-232-4947 | winsonchen108@gmail.com | linkedin.com/in/Winson-Chen
github.com/W1nson | scholar.google.com/Winson

EDUCATION

University of Minnesota, Twin Cities	Minneapolis, MN
<i>Computer Science, Philosophy of Doctorate</i>	<i>expect May 2030</i>
University of California, Santa Cruz	Santa Cruz, CA
<i>Scientific Computing and Applied Mathematics, Master of Science</i>	<i>June 2023</i>
University of California, Santa Cruz	Santa Cruz, CA
<i>Computer Science, Bachelor of Science, Cum laude</i>	<i>June 2022</i>
Irvine Valley College	Irvine, CA
<i>Computer Science, Associate in Science, Cum laude</i>	<i>Aug. 2020</i>

PUBLICATIONS

InfantAgent-Next: A Multimodal Generalist Agent for Automated Computer Interaction

NeurIPS'25

Bin Lei, Weitai Kang, Zijian Zhang, **Winson Chen**, Xi Xie, Shan Zuo, Mimi Xie, Ali Payani, Mingyi Hong, Yan Yan, Caiwen Ding

Active Listening: Personalized Question Generation in Open-Domain Social Conversation with User Model Based Prompting

EMNLP'24

Kevin Bowden, Fan Yue, **Winson Chen**, Wen Cui, Davan Harrison, Eric Xin Wang, Marilyn Walker

Early Experience with Transformer-Based Similarity Analysis for DataRaceBench

Correctness Workshop @ SC'22

Winson Chen, Tristan Vanderbruggen, Pei-Hung Lin, Chunhua Liao, Murali Emani

Aerial Vision-and-Dialog Navigation

ACL'23

Yue Fan, **Winson Chen**, Tongzhou Jiang, Chun Zhou, Yi Zhang, Eric Xin Wang

Athena 3.0: Personalized multimodal chatbot with neuro-symbolic dialogue generators

Alexa Prize SocialBot Grand Challenge 5 Proceedings

Yue Fan, Kevin K Bowden, Wen Cui, **Winson Chen**, Vrindavan Harrison, Angela Ramirez, Saaket Agashe, Xinyue Gabby Liu, Neha Pullabhotla, NQJ Bheemanpally, S Gard, M Walker, XE Wang

EXPERIENCE

Business Analyst / Full-Stack Engineer

Mar. 2024 – Apr. 2025

InfoIMAGE, Inc

Brisbane, CA

- Led a team of four in the design and development of an internal website, coordinating tasks, managing timelines, and ensuring successful project completion.
- Developed a new internal Core platform using Next.js and Flask, resulting in a 80% increase in data sample lookup
- Designed Core ORM by using SQLAlchemy, to increase development time by 20% and improve security
- Utilized RESTful API to facilitate front-end to back-end communication, improving data consistency by 50%
- Integrated AG-Grid to optimize the user experience resulting 15% increase in filtering and searching process

Machine Learning Engineer

Aug. 2022 – Sept. 2023

University of California, Santa Cruz / Amazon Alexa Prize

Santa Cruz, CA

- Reduced training time by 23% using LoRA and lightning package to fine-tune RedPajama 3B
- Enhanced system speed by 34% via AWS EC2 auto-scaling deployment of RESTful server with Red Pajamas 3B
- Increased 30% user ratings by fine-tuning RedPajamas 3B with dialog history for personalized questions
- Over three months, we achieved a 10% increase in user evaluations by implementing an APL (Alexa Presentation Language) detail template with text that auto-scrolls in sync with the voice-over
- Through the design and A/B testing of multiple variations of dialog managers and APL templates, we successfully attained a 13.5% improvement in satisfactory evaluations, striving to unveil the ultimate user experience

Data Scientist Intern <i>Lawrence Livermore National Laboratory</i>	Jun. 2022 – Sept. 2022 Remote
<ul style="list-style-type: none"> Designed web scraping pipeline using Scrapy's parallel process, improved 25% data collection Fine-tuned CodeBERT with DataRaceBench to adapt to parallel syntax in C/C++ Optimized training speed by freezing hidden layers, reducing CodeBERT training time by 10% Set up a data pipeline to visualize SARS-CoV-2 Inhibitor for training a CNN in classification 	
ERIC Lab Assistant <i>University of California, Santa Cruz</i>	Aug. 2021 – Sept. 2023 <i>Santa Cruz, CA</i>
<ul style="list-style-type: none"> Received Amazon Alexa Prize Award to work on Alexa Prize SocialBot, TaskBot Challenge Collecting data with by simulating drone controller to create visual language navigation dataset Developed toolbox to help users to find published papers from Association for Computational Linguistics 	
Research Assistant <i>University of California, Santa Cruz</i>	Aug. 2021 - Dec. 2021 <i>Santa Cruz, CA</i>
<ul style="list-style-type: none"> Utilizing PySINDY package to fit the spatial-temporal data of drosophila genes using machine learning Experimenting various models to find the behavior of drosophila gap genes 	
TALKS	
Early Experience with Transformer-Based Similarity Analysis for DataRaceBench <i>Dallas, Texas 11/18/2022</i>	
In-person talk at Correctness workshop at SC22	
TEACHING	
Graduate Teaching Assistant <i>University of California, Santa Cruz</i>	Sept. 2022 - June 2023 <i>Santa Cruz, CA</i>
<ul style="list-style-type: none"> MATH 19A (Calculus for Science, Engineering, and Mathematics) AM 10 (Mathematical Methods for Engineers I) AM 20 (Mathematical Methods for Engineers II) AM 214 (Applied Dynamical System) AM 250 (Intro to High-Performance Computing) 	
HONORS AND AWARDS	
<ul style="list-style-type: none"> Dean's Honor: Fall 2020, Spring 2021, Fall 2021, Winter 2022, Spring 2022 Next Gen. SAM Scholar (Funded by National Science Foundation) 	
TECHNICAL SKILLS	
Languages: Python, C, C++, CUDA/HIP(C++), SQL, JavaScript, HTML, CSS, Java	
Frameworks: Flask, Django, PostgreSQL, React Native, jQuery, Bootstrap, Next.js	
Developer Tools: Docker, Postman, Git, Jira	
Libraries: PyTorch, TensorFlow, Pandas, NumPy, Matplotlib, Seaborn, Lightning, Gradio, Keras, BeautifulSoup, Scrapy, Selenium, OpenMP, MPI, OpenCV, Langchain, Streamlit, ChromaDB	
Services : RESTful API, SLURM, AWS: SageMaker, S3, EC2	