

Yenting Lin

adamlin120@gmail.com • +1 669 248 9409 • Personal Page • LinkedIn • Hugging Face • GitHub • Google Scholar

EDUCATION	National Taiwan University	
	<ul style="list-style-type: none">▪ Ph.D. in Computer Science<ul style="list-style-type: none">• Advisor: Prof. Yun-Nung (Vivian) Chen▪ B.B.A. in Information Management<ul style="list-style-type: none">• Graduated with Presidential Award (Top 5%)	Feb 2020 – Jan 2025 Sep 2015 – Jan 2020
INDUSTRY EXPERIENCE	Meta , Research Scientist Intern, Menlo Park Stepwise-RL for reasoning and Long Chain-of-Thought reasoning Mentor: Tengyu Xu, Sainbayar Sukhbaatar, Yuandong Tian	Jul 2024 – Jan 2025
	NVIDIA , Artificial Intelligence Research Intern - Deep Learning, Taipei Post-hoc correction for speech recognition Manager: Frank Wang	Mar 2024 – Jun 2024
	Amazon , Applied Scientist Intern, Sunnyvale LLM-as-a-Factuality-Judge Mentor: Seokhwan Kim, Yang Liu	May 2023 – Aug 2023
	Amazon , Applied Scientist Intern, Sunnyvale Rejection sampling for data augmentation Mentor: Alexandros Papangelis, Dilek Hakkani-Tür	Jun 2022 – Sep 2022
	Amazon , Applied Scientist Intern, Sunnyvale Online-RL for data augmentation Mentor: Alexandros Papangelis, Dilek Hakkani-Tür	Jun 2021 – Oct 2021
	MediaTek Research , Research Intern, Taipei Fill-in-the-Middle LLM Manager: Dashan Shiu	Jul 2020 – Sep 2020
	IBM Research , Research Intern, Yorktown Heights First place in n2c2 Clinical Similarity Challenge and TAC 2019 Entity Recognition competitions Manager: Ching-Huei Tsou	Jul 2019 – Sep 2019
OPEN SOURCE	<ul style="list-style-type: none">▪ Taiwan-LLM – first Traditional Chinese Large Language Model suite. [Models] [Code]<ul style="list-style-type: none">• Continuous pretraining of Llama-3 70B on 500B tokens.• Post-training via data flywheel with user data from twllm.com and arena.twllm.com.• Compute granted by NVIDIA; integrated into NIM platform and featured in NVIDIA sovereign-AI blog.• Top open-weight LLMs on knowledge and localization benchmarks.	
INVITED TALKS	<ul style="list-style-type: none">▪ Invited speaker at Meta Global Llama Summit and Meta APAC Innovation Symposium.	
SELECTED PUBLICATIONS	<ul style="list-style-type: none">[1] Step-KTO: Optimizing Mathematical Reasoning through Stepwise Binary Feedback Yen-Ting Lin, Di Jin, Tengyu Xu, Tianhao Wu, Sainbayar Sukhbaatar, Chen Zhu, Yun He, Yun-Nung Chen, Jason Weston, Yuandong Tian, Arash Rahnama, Sinong Wang, Hao Ma, Han Fang, <i>Preprint</i>.[2] NeKo: Toward Post Recognition Generative Correction Large Language Models with Task-Oriented Experts Yen-Ting Lin, Chaohan Yang, Zhehuai Chen, Piotr Zelasko, Xuesong Yang, Zihching Chen, Krishna Puvvada, SzuWei Fu, Ke Hu, Junwei Chiu, Jagadeesh Balam, Boris Ginsburg, Yuchiang Frank Wang, in <i>Proceedings of ACL 2025 Industry Track</i>.[3] Measuring Taiwanese Mandarin Language Understanding Po-Heng Chen, Sijia Cheng, Wei-Lin Chen, Yen-Ting Lin, Yun-Nung Chen, in <i>Proceedings of COLM 2024</i>.[4] Taiwan LLM: Bridging the Linguistic Divide with a Culturally Aligned Language Model Yen-Ting Lin, Yun-Nung Chen, <i>Preprint</i>.[5] LLM-Eval: Unified Multi-Dimensional Automatic Evaluation for Open-Domain Conversations with Large Language Models Yen-Ting Lin, Yun-Nung Chen, in <i>Proceedings of the 5th Workshop on NLP for ConvAI (NLP4ConvAI)</i>, 2023.[6] Selective In-Context Data Augmentation for Intent Detection using Pointwise V-Information Yen-Ting Lin, Alexandros Papangelis, Seokhwan Kim, Sungjin Lee, Devamanyu Hazarika, Mahdi Namazifar, Di Jin, Yang Liu, Dilek Hakkani-Tür, in <i>Proceedings of EACL 2023</i>.[7] Knowledge-Grounded Conversational Data Augmentation with Generative Conversational Networks Yen-Ting Lin, Alexandros Papangelis, Seokhwan Kim, Dilek Hakkani-Tür, in <i>Proceedings of SIGDIAL 2022</i>.	