MINGYU DEREK MA

Homepage: derek.ma Email: derek.ma@ucla.edu X: @mingyu_ma LinkedIn Tel: +1 (213) 245-0293

EDUCATION

University of California, Los Angeles — Los Angeles, CA
PhD Candidate in Computer Science, advised by Prof. Wei Wang — *Sep. 2020 - Mar. 2025 (expected)*
- J.P. Morgan Chase AI PhD Fellowship
- Amazon Fellow supported by Amazon PhD Fellowship

The Hong Kong Polytechnic University — Hong Kong
Bachelor of Science in Computing (First Class Honours) — *Aug. 2014 - Jun. 2018*
- Best Thesis Award, valedictorian at the university commencement

University of Maryland — College Park, MD
Exchange Study — *Aug. - Dec. 2016*

PROFESSIONAL EXPERIENCE

Genentech/Roche Prescient Design Foundation Model Team — New York City, NY
*Machine Learning Scientist with **Dr. Keunwoo Choi**, **Dr. Stephen Ra** and **Prof. Kyunghyun Cho*** — *Jun. 2024 - Present*

Orchestrating Tool Ecosystem with Scalable LLM Agent for Drug Discovery
- Developed the agent system for drug discovery supported by domain-specialized customized large language model
- Developed a thought process planning module to bridge intentions and actions, RAG based on history pipelines
- Unified, standardized and integrated 30+ specialized in-house ML models and their combined pipelines as tools
- Constructed a multi-agent simulation framework for large-scale agent improvement and evaluation
- Deployed in real pipelines, rolling out beta testing within Roche

Coding-free Customized Agent Construction; Marketplace for Tools and Agent Design Sharing
- Designed and developed a company-wide tool and agent design marketplace to share efforts of agent development
- Supported configuration, generation, construction, and deployment of a customized agent without coding

Post-training of LLMs Specialized in Biomedical and Pharmaceutical Knowledge
- Constructed and synthesized instruction tuning and reasoning steps datasets based on internal knowledge sources
- Involved in multi-node training of 8B LLM on massive internal corpus, led the training of 1B model with enhanced reasoning
- Developed models are deployed and being used by group-wide scientists and employees daily for chatting, RAG, and agent

Amazon Artificial General Intelligence — Sunnyvale, CA
*Applied Scientist Intern with **Dr. Jiun-Yu Kao** and **Dr. Tagyoung Chung*** — *Jun. - Sep. 2021 & Jun. - Sep. 2022*

Tracing the Influence: Bias Mitigation for Generative QA Models with Minimal Supervised In-Context Bias Detection [17]
- Developed the first bias mitigation method for generative QA models without training on instance-level bias label annotation
- Reduced 9 types of bias (gender, religion, race etc.) contained in language models while maintaining QA performance

Parameter-Efficient Low-Resource Dialogue State Tracking by Prompt Tuning [9]
- Developed prompt tuning method for dialogue state tracking (DST) using frozen pre-trained generative models
- Achieved better performance than fine-tuning language model while using only 1% of LM parameters in few-shot DST

UCLA Computer Science Department — Log Angeles, CA
*Graduate Student Researcher with **Prof. Nanyun (Violet) Peng** and **Prof. Wei Wang*** — *Aug. 2019 - Present*

Modeling Influence Pathways on Social Media, Learning Online-Offline Dynamics (funded by DARPA & ARL) [13, 14, 18, 19]
- Led a team of 10+ researchers from UCLA, USC, Stanford and Harvard to deliver a misinformation trace interpretation system

Clinical Outcome Prediction with Generative Large Language Models (funded by OptumLabs) [25, 28, 29, 30]
- Led the delivery efforts and the development of a clinical event prediction model, coordinated 8 researchers and engineers
- Developed algorithms and models are currently deployed at UnitedHealthcare on real-world patient cases

Knowledge-directed Artificial Intelligence Reasoning Over Schemas (funded by DARPA) [5, 6]
Automated Text Analysis with Social Media Coding for Extreme Heat Narrative Analysis (funded by NSF)
METAScientist: Bioinspired Meta-Material Design with an Autonomous Scientist (funded by DARPA) [21]
Mining Scientific Discovery Processes and Agent Tool-Use Trajectory Planning (funded by Amazon)

PUBLICATIONS & PREPRINTS

950 citations on Google Scholar

[30] Memorize and Rank: Elevating Large Language Models for Clinical Diagnosis Prediction [Link]
Mingyu Derek Ma, Xiaoxuan Wang, Yijia Xiao, Anthony Cuturrufo, Vijay S Nori, Eran Halperin, Wei Wang
AAAI, 2025

[29] Inferring from Logits: Exploring Best Practices for Decoding-Free Generative Candidate Selection [Link]
Mingyu Derek Ma*, Yanna Ding*, Zijie Huang, Jianxi Gao, Yizhou Sun, Wei Wang
NeurIPS workshop on Efficient Natural Language and Speech Processing, 2024

[28] CliBench: A Multifaceted and Multigranular Evaluation of Large Language Models for Clinical Decision Making [Link]
Mingyu Derek Ma, Chenchen Ye, Yu Yan, Xiaoxuan Wang, Peipei Ping, Timothy S Chang, Wei Wang
arXiv preprint, 2024

[27] GIVE: Structured Reasoning with Knowledge Graph Inspired Veracity Extrapolation [Link]
Jiashu He, Mingyu Derek Ma, Jinxuan Fan, Dan Roth, Wei Wang, Alejandro Ribeiro
arXiv preprint, 2024

[26] BIASDETECTOR: Multi-Agent Synergy for Comprehensive Bias Detection in Structural Data [Link]
Haoxuan Li, Mingyu Derek Ma, Jen-tse Huang, Wei Wang, Jieyu Zhao
arXiv preprint, 2024

[25] CLIMB: A Benchmark of Clinical Bias in Large Language Models [Link]
Yubo Zhang*, Shudi Hou*, Mingyu Derek Ma, Wei Wang, Muhao Chen, Jieyu Zhao
EMNLP workshop on NLP for Positive Impact, 2024

[24] MIRAI: Evaluating LLM Agents for Event Forecasting [Link]
Chenchen Ye*, Ziniu Hu*, Yihe Deng*, Zijie Huang, Mingyu Derek Ma, Yanqiao Zhu, Wei Wang
arXiv preprint, 2024

[23] Are Large-Language Models Graph Algorithmic Reasoners? [Link]
Alexander K. Taylor, Anthony Cuturrufo, Vishal Yathish, Mingyu Derek Ma, Wei Wang
arXiv preprint, 2024

[22] MuirBench: A Comprehensive Benchmark for Robust Multi-image Understanding [Link]
Fei Wang*, Xingyu Fu*, James Y. Huang, Zekun Li, Qin Liu, Xiaogeng Liu, Mingyu Derek Ma, Nan Xu, Wenxuan Zhou, Kai Zhang, Tianyi Yan, Wenjie Jacky Mo, Pan Lu, Chunyuan Li, Chaowei Xiao, Kai-Wei Chang, Dan Roth, Sheng Zhang, Hoifung Poon, Muhao Chen
ICLR, 2025

[21] MetaScientist: A Human-AI Synergistic Framework for Automated Mechanical Metamaterial Design [Link]
Jingyuan Qi, Zian Jia, Minqian Liu, Wangzhi Zhan, Junkai Zhang, Xiaofei Wen, Jingru Gan, Jianpeng Chen, Qin Liu, Mingyu Derek Ma, Bangzheng Li, Haohui Wang, Adithya Kulkarni, Muhao Chen, Dawei Zhou, Ling Li, Wei Wang, Lifu Huang
arXiv preprint, 2024

[20] GraphVis: Boosting LLMs with Visual Knowledge Graph Integration [Link]
Yihe Deng, Chenchen Ye, Zijie Huang, Mingyu Derek Ma, Yiwen Kou, Wei Wang
NeurIPS, 2024

[19] Decoding Susceptibility: Modeling Misbelief to Misinformation Through a Computational Approach [Link]
Yanchen Lin, Mingyu Derek Ma, Wenna Qin, Azure Zhou, Jiaao Chen, Weiyan Shi, Wei Wang, Diyi Yang
EMNLP, 2024

[18] Improving Event Definition Following For Zero-Shot Event Detection [Link]
Zefan Cai*, Po-Nien Kung*, Ashima Suvarna, Mingyu Derek Ma, Hritik Bansal, Baobao Chang, P. Jeffrey Brantingham, Wei Wang, Nanyun Peng
ACL, 2024

[17] Mitigating Bias for Question Answering Models by Tracking Bias Influence [Link]
Mingyu Derek Ma, Jiun-Yu Kao, Arpit Gupta, Yu-Hsiang Lin, Wenbo Zhao, Tagyoung Chung, Wei Wang, Kai-Wei Chang, Nanyun Peng
NAACL, 2024

[16] Instructions as Backdoors: Backdoor Vulnerabilities of Instruction Tuning for Large Language Models [Link]
Jiashu Xu, Mingyu Derek Ma, Fei Wang, Chaowei Xiao, Muhao Chen
NAACL, 2024
Top 15 most cited paper at NAACL

[15] Instructional Fingerprinting of Large Language Models [Link]
Jiashu Xu, Fei Wang*, Mingyu Derek Ma*, Pang Wei Koh, Chaowei Xiao, Muhao Chen
NAACL, 2024

[14] STAR: Boosting Low-Resource Information Extraction by Structure-to-Text Data Generation with Large Language Models [Link]
Mingyu Derek Ma, Xiaoxuan Wang, Po-Nien Kung, P. Jeffrey Brantingham, Nanyun Peng, Wei Wang
AAAI, 2024

[13] MIDDAG: Where Does Our News Go? Investigating Information Diffusion via Community-Level Information Pathways [Link]
Mingyu Derek Ma, Alexander K. Taylor, Nuan Wen, Yanchen Lin, Po-Nien Kung, Wenna Qin, Shicheng Wen, Azure Zhou, Diyi Yang, Xuezhe Ma, Nanyun Peng, Wei Wang
AAAI, 2024

[12] DICE: Data-Efficient Clinical Event Extraction with Generative Model [Link]
Mingyu Derek Ma*, Alexander K. Taylor*, Wei Wang, Nanyun Peng
ACL, 2023

[11] Can NLI Provide Proper Indirect Supervision for Low-resource Biomedical Relation Extraction? [Link]
Jiashu Xu, Mingyu Derek Ma, Muhao Chen
ACL, 2023

[10] Multi-hop Evidence Retrieval for Cross-document Relation Extraction [Link]
Keming Lu, I-Hung Hsu, Wenxuan Zhou, Mingyu Derek Ma, Muhao Chen
ACL Findings, 2023

[9] Parameter-Efficient Low-Resource Dialogue State Tracking with Prompt Tuning [Link]
Mingyu Derek Ma, Jiun-Yu Kao, Shuyang Gao, Arpit Gupta, Di Jin, Tagyoung Chung, Nanyun Peng
INTERSPEECH, 2023

[8] Summarization as Indirect Supervision for Relation Extraction [Link]
Keming Lu, I-Hung Hsu, Wenxuan Zhou, Mingyu Derek Ma, Muhao Chen
EMNLP Findings, 2022

[7] Bending the Future: Autoregressive Modeling of Temporal Knowledge Graphs in Curvature-Variable Hyperbolic Spaces [Link]
Jihoon Sohn, Mingyu Derek Ma, Muhao Chen
Conference on Automated Knowledge Base Construction (AKBC), 2022

[6] HyperExpan: Taxonomy Expansion with Hyperbolic Representation Learning [Link]
Mingyu Derek Ma, Muhao Chen*, Te-Lin Wu*, Nanyun Peng
EMNLP Findings, 2021

[5] EventPlus: A Temporal Event Understanding Pipeline [Link]
Mingyu Derek Ma*, Jiao Sun*, Mu Yang, Kung-Hsiang Huang, Nuan Wen, Shikhar Singh, Rujun Han, Nanyun Peng
NAACL, 2021

[4] Dual Memory Network Model for Sentiment Analysis of Review Text [Link]
Jiaxing Shen*, Mingyu Derek Ma*, Rong Xiang, Qin Lu, Elvira Perez Vallejos, Ge Xu, Chu-Ren Huang, Yunfei Long
Knowledge-Based Systems, 2020

[3] Implicit Discourse Relation Identification for Open-domain Dialogues [Link]
Mingyu Derek Ma, Kevin K. Bowden, Jiaqi Wu, Wen Cui, Marilyn Walker
ACL, 2019

[2] Dual Memory Network Model for Biased Product Review Classification [Link]
Yunfei Long*, Mingyu Ma*, Qin Lu, Rong Xiang, Chu-Ren Huang
EMNLP workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, 2018

[1] BlocHIE: a BLOCkchain-based platform for Healthcare Information Exchange [Link]
Shan Jiang, Jiannong Cao, Hanqing Wu, Yanni Yang, Mingyu Ma, Jianfei He
IEEE International Conference on Smart Computing (SmartComp), 2018

*: equal contribution

Fundings

Co-PI
- NAIRR Pilot funding "Applying and Evaluating Interpretable LLM on Single-Cell Omics Data" *Jul. 2024*
- NAIRR Pilot funding "Evaluating and Mitigating Biases in Clinical Decision Models" *May 2024*

Primary contributor to funded proposals
- DARPA AIE funding "METAScientist: Bioinspired Meta-Material Design with an Autonomous Scientist" *Nov. 2023*

Awards & Honors

J.P. Morgan Chase AI Ph.D. Fellowship *2024*
Amazon Ph.D. Fellowship *2024*
Top 15 Most Influential NAACL Paper (by Paper Digest) [16] *2024*
Best Thesis Award (Top 1 in the department), PolyU *2018*
Silver Award, Hong Kong ICT Awards, Hong Kong Gov. *2018*
- Invented a Cantonese dialogue system for oral health surveys for the elderly, deployed at the Prince of Wales Hospital of CUHK

Hong Kong Government Scholarship Fund Talent Development Scholarship *2018*
CMA (The Chinese Manufacturers' Association of Hong Kong) & Donors Scholarship *2018*
Champion and Most Innovative Award of Region, Imagine Cup, Microsoft *2017*
- Developed a computer vision app for the visually impaired, collaborated with NGOs and benefited thousands of users

Best Service Project (Top 1 in the Greater China), Youth Volunteer Service Conference *2017*
- Led service projects in Rwanda and Cambodia to build electricity and tech infrastructure for local underdeveloped communities

Commercial Radio 50th Anniversary Scholarship *2016*
Wong Tit-Shing Scholarship *2016*

Teaching Experience

NIH Bridge2AI Scholar Program (teaching AI/ML to health practitioners), UCLA,
Teaching Associate with Prof. Wei Wang and Prof. Peipei Ping *Winter 2024*
NSF Data Science for ALL Program (teaching ML to high-school students from low-income families), UCLA & UC Irvine,
Teaching Associate with Prof. Wei Wang and Prof. Chen Li *Summer 2023*
- Related paper to be presented at the Data Science Education K-12: Research to Practice Conference

CS32 Introduction to Computer Science II, UCLA, *Teaching Associate with Prof. Edwin Ambrosio* *Summer 2023*
CS35L Software Construction Laboratory, UCLA, *Teaching Assistant with Prof. Paul Eggert* *Spring 2023*
CS188 Natural Language Processing, UCLA, *Lead Teaching Assistant with Prof. Nanyun (Violet) Peng* *Winter 2022*

- Designed written and programming homework, exams and discussion session content from scratch for the undergraduate NLP class offered for the first time at UCLA

CS31 Introduction to Computer Science, UCLA, *Teaching Assistant with Prof. David Smallberg* *Fall 2021*

Engineering Computation and Data Science, MIT,
Teaching Assistant with Dr. Abel Sanchez and Prof. John Williams *Summer 2017*

- Developed and deployed a coding assignment e-learning system for real-time grading and actionable feedback

Professional Activities

Lead Organizer

- AAAI 2025 Spring Symposium on Large Language Model Agents for Scientific Discovery [Link]

Organizer

- The Web Conference 2025 workshop: Towards Agentic AI for Science: Hypothesis Generation, Comprehension, Quantification, and Validation

Area Chair

- ACL Rolling Review (2025)

Program Committee

- ACL Rolling Review (2025, 2024, 2023, 2022, 2021), NeurIPS (2024), ICLR (2025), ACL (2024, 2023, 2022), EMNLP (2024, 2023, 2022, 2021, 2019), NAACL (2025, 2024, 2022), COLM (2024), AISTATS (2025), KDD (2023), TheWebConf (2024), EACL (2023), NLPCC (2023, 2022)
- IEEE/ACM Transactions on Audio, Speech, and Language Processing (since 2023)
- NeurIPS workshop on Efficient Natural Language and Speech Processing (2024)
- AAAI Spring Symposium on Clinical Foundation Models (2024)
- EMNLP workshop on Deep Learning for Low-resources NLP workshop (2019)
- SoCal NLP Symposium (2023, 2022)

Handbook Producer: ACL 2020

Mentoring Experience

Graduate Students

- Kung-Hsiang Huang (USC MSCS, later UIUC CS PhD [5]), Keming Lu (USC MSCS, now Alibaba scientist [8, 10]), Yanchen Liu (Harvard MSCS [13, 19]), Nuan Wen (USC MSCS, now USC CS PhD student [5, 13]), Wenna Qin (Stanford MSCS [13, 19])), Shikhar Singh (USC MSCS [5]), Haoxuan Li (USC MSCS), Shruti Tyagi (UCLA MSCS)

Undergraduate Students

- Jiashu Xu (USC UG, later Harvard MSCS, now NVIDIA scientist [11, 15, 16]), Shuowei Jin (USTC UG, now UMich PhD student [6]), Azure Zhou (Stanford UG [13, 19]), Utkarsh Lal (UCLA UG [14]), Michael M. Song (UCLA UG [14]), Sreya Muppalla (UCLA UG), Anh Mac (UCLA UG), Kyle Zheng (UCLA UG)

References

Wei Wang, Leonard Kleinrock Chair Professor of Computer Science and Computational Medicine, UCLA (weiwang@cs.ucla.edu)
Kyunghyun Cho, Professor of Computer Science and Data Science, New York University (kyunghyun.cho@nyu.edu)
Richard Bonneau, Vice President of Machine Learning for Drug Discovery, Genentech, Roche (bonneaur@gene.com)
Nanyun Peng, Associate Professor of Computer Science, UCLA (violetpeng@cs.ucla.edu)
Yizhou Sun, Professor of Computer Science, UCLA (yzsun@cs.ucla.edu)
Peipei Ping, Professor of Physiology, Medicine/Cardiology, and Biomedical Informatics, UCLA (pping38@ucla.edu)
Nathan Frey, Principal Scientist and Group Leader, Genentech, Roche (frey.nathan.nf1@gene.com)