Manling Li

Room 3201, 2233 Tech Dr, Evanston, IL 60208
(+1) 650-283-2881 | manling.li@northwestern.edu | https://limanling.github.io/

Professional Experiences

Northwestern University, Chicago, Illinois, USA *2024.09 - Present*
- Computer Science Department
- Tenure-Track Assistant Professor
- Affiliated Faculty, Center for Robotics and Biosystems
- Affiliated Faculty, Cognitive Science

Stanford University, Stanford, California, USA *2023.09 - 2024.08*
- Computer Science Department
- Postdoc, mainly working with Prof Jiajun Wu (postdoc advisor) and Prof Fei-Fei Li

Amazon, Seattle, Washington, USA *2025.09 - Present*
- Amazon Scholar, part-time

Educations

University of Illinois at Urbana-Champaign, USA *2023*
- Ph.D. in Computer Science
- Advisor: Prof. Heng Ji. Thesis Committee: Shih-Fu Chang, Chengxiang Zhai, Jiawei Han, Kyunghyun Cho
- Thesis: Event Multimodal Knowledge Acquisition **ACL 2025 Inaugural Dissertation Award Honorable Mention**

University of Chinese Academy of Sciences, China *2018*
- M.S. in Computer Science

University of Science and Technology Beijing, China *2015*
- B.E. in Computer Science (Internet of Things), B.M. in Information Management and Information Systems (minor)

Selected Awards

- **MIT Tech Review 35 Innovators Under 35** Global in 2025
- **ACL 2025 Inaugural Dissertation Award Honorable Mention**, only one awardee in the graduation class of 2023
- **AAAI 2025 New Faculty Highlights**
- **ACL 2024 Outstanding Paper Award**
- **NAACL 2021 Best Demo Paper Award**
- **ACL 2020 Best Demo Paper Award**
- **DARPA Riser** in 2022 (nominated by DARPA)
- **Microsoft Research PhD Fellowship** in 2021
- **EE CS Rising Star** in 2022
- **Ranked 1st** in NIST TAC Streaming Multimedia Knowledge Base Population (SM-KBP) 2019 and 2020

Selected Publications

Google Scholar: https://scholar.google.com/citations?user=6U4SXnUAAAAJ&hl=en

[P44] **RAGEN: Training Agents by Reinforcing Reasoning**
Zihan Wang*, Kangrui Wang*, Qineng Wang*, Pingyue Zhang*, Linjie Li*, Zhengyuan Yang, Kefan Yu, Minh Nhat Nguyen, Yiping Lu, Kyunghyun Cho, Jiajun Wu, Li Fei-Fei, Lijuan Wang, Yejin Choi, **Manling Li**
2.4k Github Stars as an Open-Source RL Agent framework
Best Poster Award at MMLS 2025 (Midwest Machine Learning Symposium)
Featured by MIT Tech Review, Lambda Partner Spotlight, Medium, AI News, VentureBeat, MarkTechPost, Business Leaders Review, Cognixia, EmergentMind, AIModels, AnyScale, ChatPaper, ArizeAI, AITechTrends, etc

[P43] **VAGEN: Reinfocing World Model Reasoning for Multi-Turn VLM Agents**
Kangrui Wang*, Pingyue Zhang*, Zihan Wang*, Yaning Gao*, Linjie Li*, Qineng Wang, Chi Wan, Hanyang Chen, Yiping Lu, Zhengyuan Yang, Lijuan Wang, Ranjay Krishna, Jiajun Wu, Li Fei-Fei, Yejin Choi, **Manling Li**
NeurIPS 2025. Featured by MIT Tech Review, etc

[P42] **Spatial Mental Modeling from Limited Views**
Qineng Wang*, Baiqiao Yin*, Pingyue Zhang, Jianshu Zhang, Kangrui Wang, Zihan Wang, Jieyu Zhang, Keshigeyan Chandrasegaran, Han Liu, Ranjay Krishna, Saining Xie, Jiajun Wu, Li Fei-Fei, **Manling Li**
Best Paper Award at ICCV 2025 SP4V (Workshop on Structural Priors for Vision)
Best of ICCV, selected by Voxel51

[P41] **T*: Re-thinking Temporal Presentation Search for Long-Form Video Understanding**
Jinhui Ye, Zihan Wang, Haosen Sun, Keshigeyan Chandrasegaran, Zane Durante, Cristobal Eyzaguirre, Yonatan Bisk, Juan Carlos Niebles, Ehsan Adeli, Li Fei-Fei, Jiajun Wu, **Manling Li**
CVPR 2025, **Oral Presentation** at ICCV 2025 Workshop on Long Multi-Scene Video Foundations

[P40] **Why is Spatial Reasoning Hard for VLMs? An Attention Mechanism Perspective on Focus Areas**
Shiqi Chen, Tongyao Zhu, Ruochen Zhou, Jinghan Zhang, Siyang Gao, Juan Carlos Niebles, Mor Geva, Junxian He, Jiajun Wu, **Manling Li**
ICML 2025

[P39] **Exploring Diffusion Transformer Designs via Grafting**
Keshigeyan Chandrasegaran*, Michael Poli*, Daniel Y. Fu, Dongjun Kim, Lea M. Hadzic, **Manling Li**, Agrim Gupta, Stefano Massaroli, Azalia Mirhoseini, Juan Carlos Niebles, Stefano Ermon, Li Fei-Fei
NeurIPS 2025 **Oral Presentation (0.36%)**

[P38] **ROSETTA: Constructing CodeBased Reward from Unconstrained Language Preference**
Sanjana Srivastava*, Kangrui Wang*, YungChieh Chan, Tianyuan Dai, **Manling Li**, Ruohan Zhang, Mengdi Xu, Jiajun Wu, Li Fei-Fei
Best Paper Award at RSS 2025 CRLH (Workshop on Continual Robot Learning from Humans)

[P37] **LayoutVLM: Differentiable Optimization of 3D Layout via Vision-Language Models**
Fan-Yun Sun, Weiyu Liu, Siyi Gu, Dylan Lim, Goutam Bhat, Federico Tombari, **Manling Li**, Nick Haber, Jiajun Wu
CVPR 2025

[P36] **EmbodiedBench: Comprehensive Benchmarking Multi-modal Large Language Models for Vision-Driven Embodied Agents**
Rui Yang, Hanyang Chen, Junyu Zhang, Mark Zhao, Cheng Qian, Kangrui Wang, Qineng Wang, Teja Venkat Koripella, Marziyeh Movahedi, **Manling Li**, Heng Ji, Huan Zhang, Tong Zhang
ICML 2025 **Oral Presentation (1%)**

[P35] **SyncMind: Measuring Agent OutofSync Recovery in Collaborative Software Engineering**
Xuehang Guo, Xingyao Wang, Yangyi Chen, Sha Li, Chi Han, **Manling Li**, Heng Ji
ICML 2025

[P34] **Bring Reason to Vision: Understanding Perception and Reasoning through Model Merging**
Shiqi Chen, Jinghan Zhang, Tongyao Zhu, Wei Liu, Siyang Gao, Miao Xiong, **Manling Li**, Junxian He
ICML 2025

[P33] **The Law of Knowledge Overshadowing: Towards Understanding, Predicting and Preventing LLM Hallucination**
Yuji Zhang, Sha Li, Cheng Qian, Jiateng Liu, Pengfei Yu, Chi Han, Yi R. Fung, Kathleen McKeown, ChengXiang Zhai, **Manling Li**, Heng Ji
ACL 2025 Findings

[P32] **LEMONADE: A Large Multilingual Expert-Annotated Abstractive Event Dataset for the Real World**
Sina Semnani, Pingyue Zhang, Wanyue Zhai, Haozhuo Li, Ryan Beauchamp, Trey Billing, Katayoun Kishi, **Manling Li**, Monica Lam
ACL 2025 Findings

[P31] **Visually Descriptive Language Model for Vector Graphics Reasoning**
Zhenhailong Wang, Joy Hsu, Xingyao Wang, Kuan-Hao Huang, **Manling Li**, Jiajun Wu, Heng Ji
TMLR in 2025

[P30] **Embodied Agent Interface: Benchmarking LLMs for Embodied Decision Making**
Manling Li*, Shiyu Zhao*, Qineng Wang*, Kangrui Wang*, Yu Zhou*, Sanjana Srivastava, Cem Gokmen, Tony Lee, Li Erran Li, Ruohan Zhang, Weiyu Liu, Percy Liang, Li Fei-Fei, Jiayuan Mao, Jiajun Wu
NeurIPS 2024 D&B **Oral Presentation (0.4%)**
Best Paper Award at SoCal NLP 2024 (South California NLP Synposium)

[P29] **HourVideo: 1-Hour Video-Language Understanding**
Keshigeyan Chandrasegaran, Agrim Gupta, Taran Kota, Lea M. Hadzic, Jimming He, Cristobal Eyzaguirre, Zane Durante, **Manling Li**, Jiajun Wu, Li Fei-Fei
NeurIPS 2024 D&B

[P28] **Chain-of-Action: Faithful and Multimodal Question Answering through Large Language Models**
Zhenyu Pan, Haozheng Luo, **Manling Li**, Han Liu
ICLR 2025

[P27] **IKEA Manuals at Work: 4D Grounding of Assembly Instructions on Internet Videos**
Yunong Liu, Weiyu Liu, Shubh Khanna, Cristobal Eyzaguirre, **Manling Li**, Juan Carlos Niebles, Vineeth Ravi, Saumitra Mishra, Jiajun Wu
NeurIPS 2024 D&B

[P26] **LM-Steer: Word Embeddings Are Steers for Language Models**
Chi Han, Jialiang Xu, **Manling Li**, Yi Fung, Chenkai Sun, Nan Jiang, Tarek Abdelzaher, Heng Ji
ACL 2024
Outstanding Paper Award

[P25] **ViStruct: Visual Structural Knowledge Extraction via Curriculum Guided Code-Vision Representation**
Yangyi Chen, Xingyao Wang, **Manling Li**, Derek Hoiem, Heng Ji
EMNLP 2023

[P24] **Defining a New NLP Playground**
Sha Li, Chi Han, Pengfei Yu, Carl Edwards, **Manling Li**, Xingyao Wang, Yi Fung, Charles Yu, Joel Tetreault, Ed Hovy, Heng Ji
EMNLP 2023 Findings

[P23] **Non-Sequential Graph Script Induction via Multimedia Grounding**
Yu Zhou, Sha Li, **Manling Li**, Xudong Lin, Shih-Fu Chang, Mohit Bansal, Heng Ji
ACL 2023 **Oral Presentation**

[P22] **Open-Domain Hierarchical Event Schema Induction by Incremental Prompting and Verification**
Sha Li, Ruining Zhao, **Manling Li**, Heng Ji, Chris Callison-Burch, Jiawei Han
ACL 2023

[P21] **A Language First Approach to Procedure Planning**
Jiateng Liu, Sha Li, Zhenhailong Wang, **Manling Li**, Heng Ji
ACL 2023 Findings

[P20] **Multimedia Generative Script Learning for Planning**
Qingyun Wang, **Manling Li**, Hou Pong Chan, Lifu Huang, Julia Hockenmaier, Girish Chowdhary , Heng Ji
ACL 2023 Findings

[P19] **Open Visual Knowledge Extraction via Relation-Oriented Multimodality Model Prompting**
Hejie Cui, Xinyu Fang, Zihan Zhang, Ran Xu, Xuan Kan, Xin Liu, **Manling Li**, Yangqiu Song, Carl Yang
NeurIPS 2023

[P18] **Learning to Decompose Visual Features with Latent Textual Prompts**
Feng Wang, **Manling Li**, Xudong Lin, Hairong Lv, Alexander Schwing, Heng Ji
ICLR 2023

[P17] **Video Event Extraction via Tracking Visual States of Arguments**
Guang Yang, **Manling Li**, Xudong Lin, Jiajie Zhang, Shih-Fu Chang, Heng Ji
AAAI 2023 **Oral Presentation**

[P16] **ADEPT: A DEbiasing PrompT Framework**
Ke Yang, Charles Yu, Yi Fung, **Manling Li**, Heng Ji
AAAI 2023

[P15] **Towards Fast Adaptation of Pretrained Contrastive Models for Multi-channel Video-Language Retrieval**
Xudong Lin, Simran Tiwari, Shiyuan Huang, **Manling Li**, Mike Zheng Shou, Heng Ji, Shih-Fu Chang
CVPR 2023

[P14] **CLIP-Event: Connecting Vision and Text with Event Structures**
Manling Li, Ruochen Xu, Shuohang Wang, Xudong Lin, Chenguang Zhu, Michael Zeng, Heng Ji, Shih-Fu Chang
CVPR 2022 **Oral Presentation (4.14%)**

[P13] **Language Models with Image Descriptors are Strong Few-Shot Video-Language Learners**
Zhenhailong Wang*, **Manling Li***, Ruochen Xu, Luowei Zhou, Jie Lei, Xudong Lin, Shuohang Wang, Ziyi Yang, Chenguang Zhu, Derek Hoiem, Shih-Fu Chang, Mohit Bansal, Heng Ji
NeurIPS 2022 (* denotes equal contribution)

[P12] **COVID-19 Claim Radar: A Structured Claim Extraction and Tracking System**
Manling Li, Revanth Gangi Reddy, Ziqi Wang, Yi-Shyuan Chiang, Tuan M. Lai, Pengfei Yu, Zixuan Zhang, Heng Ji
ACL 2022 Demo

[P11] **Event Schema Induction with Double Graph Autoencoders**
Xiaomeng Jin, **Manling Li**, Heng Ji
NAACL 2022

[P10] **MuMuQA: Multimedia Multi-Hop News Question Answering via Knowledge Extraction and Grounding**
Revanth Reddy, Xilin Rui, **Manling Li**, Xudong Li, Haoyang Wen, Jaemin Cho, Lifu Huang, Mohit Bansal, Avi Sil, Shih-Fu Chang, Alexander Schiwing, Heng Ji
AAAI 2022 **Oral Presentation**

[P9] **Timeline Summarization based on Event Graph Compression via Time-Aware Optimal Transport**
Manling Li, Tengfei Ma, Mo Yu, Lingfei Wu, Tian Guo, Heng Ji, Kathleen McKeown
EMNLP 2021 **Oral Presentation**

[P8] **The Future is not One-dimensional: Complex Event Schema Induction via Graph Modeling**
Manling Li, Sha Li, Zhenhailong Wang, Lifu Huang, Kyunghyun Cho, Heng Ji, Jiawei Han
EMNLP 2021 **Oral Presentation**

[P7] **Joint Multimedia Event Extraction from Video and Article**
Brian Chen, Xudong Lin, Christopher Thomas, **Manling Li**, Shoya Yoshida, Lovish Chum, Heng Ji, Shih-Fu Chang
EMNLP 2021 Findings

[P6] **COVID-19 Literature Knowledge Graph Construction and Drug Repurposing Report Generation**
Qingyun Wang, **Manling Li**, Xuan Wang, Nikolaus Parulian, Guangxing Han, Jiawei Ma, Jingxuan Tu, Ying Lin, Haoran Zhang, Weili Liu, Aabhas Chauhan, Yingjun Guan, Bangzheng Li, Ruisong Li, Xiangchen Song, Heng Ji, Jiawei Han, Shih-Fu Chang, James Pustejovsky, David Liem, Ahmed Elsayed, Martha Palmer, Jasmine Rah, Clare Voss, Cynthia Schneider, Boyan Onyshkevych
NAACL 2021
Best Demo Paper Award
The released knowledge graph COVID-KG has been downloaded **2000+** times within a year.

[P5] **Connecting the Dots: Event Graph Schema Induction with Path Language Modeling**
Manling Li, Qi Zeng, Ying Lin, Kyunghyun Cho, Heng Ji, Jonathan May, Nathanael Chambers, Clare Voss
EMNLP 2020 **Oral Presentation**

[P4] **GAIA: A Fine-grained Multimedia Knowledge Extraction System**
Manling Li*, Alireza Zareian*, Ying Lin, Xiaoman Pan, Spencer Whitehead, Brian Chen, Bo Wu, Heng Ji, Shih-Fu Chang, Clare R Voss, Dan Napierski, Marjorie Freedman
ACL 2020
Best Demo Paper Award
Ranked 1st in NIST (National Institute of Standards and Technology) SM-KBP Evaluation 2019 and 2020

[P3] **Cross-media Structured Common Space for Multimedia Event Extraction**
Manling Li*, Alireza Zareian*, Qi Zeng, Spencer Whitehead, Di Lu, Heng Ji, Shih-Fu Chang
ACL 2020 **Oral Presentation**

[P2] **Keep Meeting Summaries on Topic: Abstractive Multi-Modal Meeting Summarization**
Manling Li, Lingyu Zhang, Heng Ji, Rich Radke
ACL 2019 Invited to present at Intel MDI Research Lab

[P1] **Multilingual Entity, Relation, Event and Human Value Extraction**
Manling Li, Ying Lin, Joe Hoover, Spencer Whitehead, Clare R Voss, Morteza Dehghani, Heng Ji
NAACL 2019 Demo
Army Research Lab NS-CTA 10-year Milestone Demo Day in 2019
DARPA Demo Day in 2019

Pre-prints and Manuscripts Under Review

[M10] **Internalizing World Models via Self-Play Finetuning for Agentic RL**
Shiqi Chen, Tongyao Zhu, Zian Wang, Jinghan Zhang, Kangrui Wang, Siyang Gao, Teng Xiao, Yee Whye Teh, Junxian He, **Manling Li**

[M9] **ENACT: Evaluating Embodied Cognition with World Modeling of Egocentric Interaction**
Qineng Wang, Wenlong Huang, Yu Zhou, Hang Yin, Tianwei Bao, Jianwen Lyu, Weiyu Liu, Ruohan Zhang, Jiajun Wu, Li Fei-Fei, **Manling Li**

[M8] **Theory of Space: Actively Constructing Spatial Beliefs in Foundation Models**
Pingyue Zhang, Zihan Huang, Yue Wang, Jieyu Zhang, Letian Xue, Zihan Wang, Qineng Wang, Keshigeyan Chandrasegaran, Ruohan Zhang, Yejin Choi, Ranjay Krishna, Jiajun Wu, Li Fei-Fei, **Manling Li**

[M7] **Unary Feedback as Observation: A Simple "Try Again" Can Elicit Multi-Turn LLM Reasoning**
Licheng Liu, Zihan Wang, Linjie Li, Chenwei Xu, Yiping Lu, Han Liu, Avirup Sil, **Manling Li**

[M6] **Chain-of-Experts: Unlocking the Communication Power of Mixture-of-Experts Models**
Zihan Wang, Rui Pan, Jiarui Yao, Robert Csordas, Linjie Li, Lu Yin, Jiajun Wu, Tong Zhang, Shiwei Liu+, **Manling Li**+

[M5] **LLM-ODE: Activation Steering for LLM Alignment via a Unified ODE-Based Framework**
Hongjue Zhao, Haosen Sun, Jiangtao Kong, Xiaochang Li, Qineng Wang, Liwei Jiang, Qi Zhu, Tarek F. Abdelzaher, Yejin Choi, Huajie Shao+, **Manling Li**+

[M4] **ERA: Embodied Reasoning Agents via Reinforcement Learning**
Hanyang Chen, Mark Zhao, Rui Yang, Qinwei Ma, Ke Yang, Jiarui Yao, Kangrui Wang, Hao Bai, Zhenhailong Wang, Rui Pan, Mengchao Zhang, Jose Barreiros, Aykut Onol, ChengXiang Zhai, Heng Ji, **Manling Li**, Huan Zhang, Tong Zhang

[M3] **Observational Scaling Laws in LLM-based Embodied Decision Making**
Qinjie Lin, **Manling Li**, Han Liu

[M2] **SENTINEL: A Multi-Level Formal Framework for Safety Evaluation of LLM-based Embodied Agents**
Simon Sinong Zhan, Yao Liu, Philip Wang, Zinan Wang, Qineng Wang, Zhian Ruan, Xiangyu Shi, Xinyu Cao, Frank Yang, Kangrui Wang, Huajie Shao, **Manling Li**, Qi Zhu

[M1] **Federated Agent Reinforcement Learning**
Canyu Chen, Kangyu Zhu, Zhaorun Chen, Zhanhui Zhou, Shizhe Diao, Yiping Lu, Tian Li, **Manling Li**, Dawn Song

Selected Invited Talks, Keynotes, and Tutorials

[T73] **Columbia University NLP Seminar** *2025.10*
From Large Language Models to Large Agent Models: Reasoning Interface with World Modeling

[T72] **University of Pennsylvania NLP Seminar** *2025.10*
From Large Language Models to Large Agent Models: Agent Reasoning Interface

[T71] **ACM MM 2025 Workshop on Multimodal Foundation Models for Spatial Intelligence (MFMSI)** *2025.10*
Why is Spatial Reasoning Hard for VLMs?

[T70] **Capital One** *2025.10*
Conversational Agent Training via Reinforcement Learning

[T69] **Amazon** *2025.10*
Knowledge-Driven Conversational Agent Training

[T68] **Dreamforce 2025** *2025.10*
Panelist on AI Research Faculty Panel

[T67] **MIT Technology Review** *2025.10*
Interview on Deepseek-OCR

[T66] **ICCV 2025 Tutorial: Foundation Models Meet Embodied Agents** *2025.10*
Foundation Models Meet Virtual Agents

[T65] **ICCV 2025 Tutorial: Safe Multi-Modal Learning** *2025.10*

Multi-Modal Misalignment

[T64] **ICCV 2025 Workshop on Multimodal Reasoning for Agentic Intelligence (MMRAgI)** *2025.10*
Reinforcing World Modeling for VLM Agents

[T63] **ICCV 2025 Workshop on Multimodal Spatial Intelligence (MUSI)** *2025.10*
Why is Spatial Reasoning Hard for VLMs?

[T62] **ICCV 2025 Workshop on Memory and Vision (MemVis)** *2025.10*
Building Internal Belief in Space

[T61] **ICCV 2025 Workshop on Long Multi-Scene Video Foundations (LongVid-Foundations)** *2025.10*
Re-thinking Temporal Search for Long-Form Video Understanding

[T60] **ICCV 2025 Workshop on Structural Priors for Vision (SP4V)** *2025.10*
Best Paper Talk: Spatial Mental Modeling from Limited Views

[T59] **University of Edinburgh CHAI Seminar Series** *2025.09*
RAGEN: Training Agents by Reinforcing Reasoning

[T58] **Stanford University** *2025.09*
Theory of Space: How LLMs Develop Spatial Internal Beliefs

[T57] **Agentic AI Frontier Seminar** *2025.09*
RAGEN: Training Agents by Reinforcing Reasoning

[T56] **Agent AI Summit @ UC Berkeley** *2025.08*
RAGEN: Training Agents by Reinforcing Reasoning

[T55] **Stanford University** *2025.08*
LLM Alignment with Control Barrier Functions

[T54] **Cross Future AI Summit** *2025.07*
See, Think, Act: Agent Training by Reinforcement Reasoning

[T53] **CVPR 2025 Workshop on Visual Concepts** *2025.07*
Why is Spatial Concept Learning Hard?

[T52] **Apple Workshop on Reasoning and Planning** *2025.07*
Training Agents with World Model Reasoning

[T51] **ACC 2025 Workshop on LLMs in Control Design and Decision Making** *2025.07*
LLMs for Embodied Decision Making

[T50] **Midwest Machine Learning Symposium 2025** *2025.06*
RAGEN: Training Agents by Reinforcing Reasoning

[T49] **Google DeepMind** *2025.05*
RAGEN: Training Agents by Reinforcing Reasoning

[T48] **NAACL 2025 Workshop on Knowledge-Augmentation for Language Models and NLP Methods** *2025.05*
RAGEN: Training Agents by Reinforcing Reasoning

[T47] **NAACL 2025 Tutorial: Foundation Models Meet Embodied Agents** *2025.05*
Foundation Models Meet Virtual Embodied Agents

[T46] **UIUC NLP Seminar** *2025.04*
RAGEN: Training Agents by Reinforcing Reasoning

[T45] **AAAI 2025 New Faculty Highlights** *2025.02*
Agent Training Under a MDP Formulation

[T44] **AAAI 2025 Workshop on LM4Plan** *2025.02*
Agent Training Under a MDP Formulation

[T43] **AAAI 2025 Bridge on Foundation Models and Planning** *2025.02*
Agent Training Under a MDP Formulation

[T42] **AAAI 2025 Tutorial: Foundation Models Meet Embodied Agents** *2025.02*
Foundation Models Meet Virtual Embodied Agents

[T41] **AAAI 2025 Tutorial: Lifecycle of Knowledge in LLMs: Memorization, Editing, and Beyond** *2025.02*
Knowledge Unlearning

[T40] **SFU @ NeurIPS 2024** *2024.12*
Embodied Agent Interface: LLMs and VLMs for Embodied Reasoning and Planning

[T39] **EMNLP 2024 Birds of a Feather** *2024.11*
LLMs for Embodied Agents

[T38] **EMNLP 2024 CustomNLP4U Workshop** *2024.11*
Customizing Large Language Models to Embodied Agents

[T37] **Adobe Research** *2024.10*
Chart Reasoning Agent Training via Compositional Grounding

[T36] **Amazon-Illinois Center on AI for Interactive Conversational Experiences (AICE) Fall Symposium** *2024.09*
From Large Language Models to Large Agent Models

[T35] **2024 Allerton Conference on Communication, Control, and Computing** *2024.09*
Reasoning and Planning with Physical World Knowledge

[T34] **TTIC Multimodal AI Workshop 2024** *2024.08*
Embodied Agent Interface: LLMs for Embodied Decision Making

[T33] **Summer Institute in Computational Social Science 2024** *2024.08*
Multimodal Knowledge for Social Good

[T32] **IJCAI 2024 Tutorial** *2024.08*
Beyond Human Creativity: A Tutorial on Advancements in AI Generated Content (AIGC)

[T31] **SpLU-RoboNLP 2024 Workshop** *2024.07*
Reasoning, Planning and Compositionality in Multimodality

[T30] **Adobe Research** *2024.07*
Visually Descriptive Language Modeling for Document Intelligence

[T29] **Apple Natural Language Understanding Workshop 2024** *2024.06*
From Large Language Models to Large Agent Models

[T28] **UIUC NLP Seminar** *2024.06*
From Words to Worlds: A Close Look to Diffusion Models (through an NLP Lens)

[T27] **Midwest Machine Learning Symposium 2024** *2024.05*
The Missing Knowledge in LLMs to Interact with the Physical World

[T26] **NAACL 2025 Tutorial** *2024.05*
Foundation Models Meet Embodied Agents

[T25] **NeurIPS 2023 Workshop on New Frontiers in Graph Learning** *2023.12*
Beyond the Beaten Path: Exploring the Role of Graphs in Multimodal Foundation Models

[T24] **ICAIF 2023 Tutorial** *2023.11*
Large Language Models for NLP in Finance [Slides]

[T23] **Stanford Vision and Learning Seminar** *2023.10*
LLMs for Robotics: Modeling the Knowledge of the Physical World

[T22] **Adobe Research** *2023.10*
Knowledge Foundation Models

[T21] **Stanford CogAI** *2023.06*
Modeling the Semantics of the Physical World

[T20] **CVPR 2023 Tutorial: Knowledge-Driven Vision-Language Encoding** *2023.06*
Knowledge-Driven Vision-Language Foundation Models

[T19] **Mohamed bin Zayed University of Artificial Intelligence** *2023.04*
Towards Factuality in Information Access: Multimodal Knowledge Acquisition and Reasoning

[T18] **University of California Los Angeles** *2023.04*
Towards Factuality in Information Access: Multimodal Knowledge Acquisition and Reasoning

[T17] **University of Toronto** *2023.03*
Towards Factuality in Information Access: Multimodal Knowledge Acquisition and Reasoning

[T16] **Carnegie Mellon University, Language Technologies Institute** *2023.02*
Towards Factuality in Information Access: Multimodal Knowledge Acquisition and Reasoning

[T15] **Northwestern University** *2023.02*
Towards Factuality in Information Access: Multimodal Knowledge Acquisition and Reasoning

[T14] **Purdue University** *2023.02*
Towards Factuality in Information Access: Multimodal Knowledge Acquisition and Reasoning

[T13] **Max Planck Institute** *2023.02*
Towards Factuality in Information Access: Multimodal Knowledge Acquisition and Reasoning

[T12] **University of California San Diego** *2023.02*
Towards Factuality in Information Access: Multimodal Knowledge Acquisition and Reasoning

[T11] **AAAI 2023 Tutorial: Knowledge-Driven Vision-Language Pretraining** *2023.02*
Knowledge in Vision and Language Pretraining

[T10] **EECS Rising Stars, University of Texas at Austin** *2022.10*
From Entity-Centric to Event-Centric Multimodal Event Knowledge Acquisition

[T9] **DARPA Forward (Invite-Only)** *2022.10*
Towards Accurate Intelligent Analysis: Event-Centric Multimedia Knowledge Extraction

[T8] **Ohio State University** *2022.10*

Event-Centric Multimedia Data Understanding

[T7] **Singapore Management University** *2022.10*
Event-Centric Multimedia Data Understanding

[T6] **Virginia Tech** *2022.09*
Multimedia Event Extraction: From Object-Centric to Event-Centric

[T5] **LOGS Graph Reasoning Seminar** *2022.08*
Event Knowledge Graph Construction

[T4] **NAACL 2022 Tutorial** *2022.07*
New Frontiers of Information Extraction

[T3] **NewsBreak** *2022.04*
Connecting Vision and Text using Event Structures

[T2] **University of Notre Dame** *2022.02*
Memories as Repositories of Events: Multimodal Event Knowledge Acquisition

[T1] **USC ISI NLP Seminar** *2022.01*
Memories as Repositories of Events: Multimodal Event Knowledge Acquisition

Teaching Experiences

Instructor

- **CS 396 Reasoning and Planning in the Foundation Model Era.**
  Northwestern University, Winter 2025
- **CS 496 Agent AI.**
  Northwestern University, Spring 2025

Guest Lecturer

- **EECS 692 Advanced AI**, University of Michigan, Spring 2025
  Reasoning and Planning with Physical World Knowledge
- **CS 6604 Advanced Topics in Natural Language Processing**, Virginia Tech, Fall 2022
  Event-Centric Multimedia Encoding
- **CSC 791&591: Advanced Topics in Efficient Deep Learning**, North Carolina State University, Fall 2022
  Event-Centric Multimedia Encoding
- **CS 546 Advanced Topics in Natural Language Processing**, UIUC, Fall 2022
  Knowledge-Driven Vision-Language Pretraining
- **CS 546 Advanced Topics in Natural Language Processing**, UIUC, Fall 2022
  Recent Advances in Multimedia Encoding
- **CS 598 Knowledge Driven Natural Language Generation**, UIUC, Spring 2022
  Timeline Summarization: Introducing Temporal Dimensions into Summarization
- **CS 546 Advanced Topics in Natural Language Processing**, UIUC, Fall 2021
  Multimedia Encoding via Vision-Language Pretraining

Grants

- **$70,000 Adobe Gift**
  Towards Spatial Intelligence: Building Embodied Agents and Chart Agents with Geometric Reasoning
  Role: PI, 2025
- **$40,000 Apple Gift**
  Training Agents via Reinforcing Reasoning
  Role: PI, 2025
- **$30,000 Dolby Gift**
  Large World Models
  Role: PI, 2025
- **$50,000 Patronus AI Gift**
  AI Agents Simulating Human Behavior
  Role: PI, 2025
- **$200,000 Northwestern Buffett Global Working Group Grant**
  Global Large Language Models Values Benchmarking
  Role: Co-PI, 2025

- **$42,000 Microsoft Postdoc Fellowship**
  2023
- **$62,000 Microsoft PhD Fellowship**
  2022
- **$1,000 Mavis Future Faculty Fellowship**
  2021

Professional Services

Organizing Committee

- ACL 2025 Virtual Infrastructure Chairs
- NAACL 2025 Publication Chairs
- EMNLP 2024 Demo Track Chairs

Senior Area Chair

- EMNLP (2025-)

Area Chair

- ACL (2023-)
- EMNLP (2023-)
- NAACL (2024-)

Senior Program Committee Member

- IJCAI (2021)

Program Committee Member

- ACL Rolling Review (2021-)
- ACL (2021-), EMNLP (2021-), NAACL (2021-), COLING (2020-), EACL (2021-), NLPCC (2021-), AACL (2020-), CCL (2020-)
- AAAI (2021-), WWW (2021-), AKBC (2021-), KDD DI Workshop (2021-)

Workshop Organizer

- **Foundation Models Meet Embodied Agents** at CVPR 2025
- **Language + Molecule** Workshop at ACL 2024
- **Towards Knowledgeable Language Models** Workshop at ACL 2024, AAAI 2025, ACL 2025
- Bridge Program on **Reasoning and Planning in Foundation Models** at AAAI 2025
- **Knowledge Discovery from Unstructured Data in Financial Services** (KDF) Workshop at SIGIR (2023)

Challenge Organizer

- **Foundation Models for Embodied Agents** at NeurIPS 2025
  Including **Embodied Agent Interface** Challenge and **BEHAVIOR** Challenge

Open-Source Systems

- **RAGEN (Reasoning Agent)** (last author) [Code]
  An open-source RL training framework for LLM agents.
- **VAGEN (World Modeling Agent))** (last author) [Code]
  An open-source World Modeling RL training framework for VLM agents, supporting Embodied Agent Tasks.
- **CoE: From Mixture of Experts to Chain of Experts** (last author) [Code]
  A open-source RL training framework for LLM agents.
- **GAIA: A Fine-grained Multimedia Knowledge Extraction System** (lead author) [Code]
  Entity, relaton, event extraction with coreference for multilingual (English, Spanish, and Russian) multimedia corpus.
- **Multilingual Entity, Relation, Event and Human Value Extraction** (lead author) [Code]
  Knowledge extraction system with human values for English, Russian, and Ukrainian.