

(Zoey) Sha Li

Room 1113, Siebel Center for Computer Science, 201 N. Goodwin Ave, Urbana, IL 61801 ·shal2@illinois.edu

RESEARCH INTERESTS

Information Extraction: Event and relation extraction with weak/limited supervision and under open-domain and corpus-level settings.

Multimodal Procedure Understanding: Reasoning about actions and entity state changes; goal-oriented planning.

Knowledge within Language Models: Understanding how LMs store and utilize knowledge; efficient updates of LMs; controlled generation.

EDUCATION

University of Illinois at Urbana-Champaign 2019–2024 (Expected)

Ph.D. Computer Science, advised by Jiawei Han

Thesis Committee: Heng Ji, Chris Callison-Burch, Chengxiang Zhai

University of Illinois at Urbana-Champaign 2017–2019

M.S. Computer Science, advised by Jiawei Han

Shanghai Jiao Tong University 2013–2017

B.S. Computer Science

INTERNSHIPS

Research Intern Apple Siri Information Intelligence Knowledge Platform May 2023 –

Manager: Yunyao Li

Applied Scientist Intern Alexa AI May – Dec. 2021

Topic: Knowledge Grounded Dialog for Stories

Manager: Yang Liu

Applied Scientist Intern AWS Science May – Sept. 2020

Topic: Domain Adaptation for Question Answering

Manager: Sheng Zha, Mu Li

Research Intern Bing/Microsoft Research July – Sept. 2019

Topic: Passage Ranking with Language Models

Manager: Xia Song, Xiaodong Liu

PUBLICATIONS

Full list at Google Scholar: <https://scholar.google.com/citations?user=0Io8J2YAAAAAJ&hl=en>

Information Extraction

- “GLEN: General-Purpose Event Detection for Thousands of Types.” [Sha Li*](#), Qiusi Zhan*, Kathryn Conger, Martha Palmer, Heng Ji, Jiawei Han. <https://arxiv.org/pdf/2303.09093.pdf>
- “Code4Struct: Code Generation for Few-Shot Event Structure Prediction.” Xingyao Wang, [Sha Li](#), Heng Ji. ACL 2023.
- “Open Relation and Event Type Discovery with Type Abstraction.” [Sha Li](#), Heng Ji, Jiawei Han. EMNLP 2022.
- “Open-Vocabulary Argument Role Prediction For Event Extraction.” Yizhu Jiao, [Sha Li](#), Yiqing Xie, Ming Zhong, Heng Ji, and Jiawei Han. EMNLP 2022 Findings.
- “P4E: Few-Shot Event Detection as Prompt-Guided Identification and Localization.” [Sha Li](#), Liyuan Liu, Yiqing Xie, Heng Ji, Jiawei Han. <https://arxiv.org/pdf/2202.07615.pdf>
- “Dynamic Global Memory for Document-level Argument Extraction.” Xinya Du, [Sha Li](#), Heng Ji. ACL 2022.
- “Eider: Empowering Document-level Relation Extraction with Efficient Evidence Extraction and Inference-stage Fusion.” Yiqing Xie, Jiaming Shen, [Sha Li](#), Yuning Mao, Jiawei Han. ACL Findings 2022.
- “Document-level Argument Extraction by Conditional Generation.” [Sha Li](#), Heng Ji, Jiawei Han. NAACL 2021.

Event Structure Learning

- “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.

- “Non-Sequential Graph Script Induction via Multimedia Grounding.” Yu Zhou, [Sha Li](#), Manling Li, Xudong Lin, Shih-Fu Chang, Mohit Bansal and Heng Ji. ACL 2023.
- “The Future is not One-dimensional: Graph Modeling based Complex Event Schema Induction for Event Prediction” Manling Li, [Sha Li](#), Zhenhailong Wang, Lifu Huang, Kyunghyun Cho, Heng Ji and Jiawei Han, Clare Voss. EMNLP 2021.

Procedure Understanding

- “OpenPI-C: A Better Benchmark and Stronger Baseline for Open-Vocabulary State Tracking. Xueqing Wu, [Sha Li](#), Heng Ji. ACL 2023 Findings.
- “A Language First Approach for Procedural Planning. Jiateng Liu, [Sha Li](#), Zhenhailong Wang, Manling Li and Heng Ji. ACL 2023 Findings.

Information Extraction Applications

- “Enhancing Knowledge Selection for Grounded Dialogues via Document Semantic Graphs.” [Sha Li](#), Mahdi Namazifar, Di Jin, Mohit Bansal, Heng Ji, Yang Liu, Dilek Hakkani-Tur. NAACL 2022.
- “RESIN-11: Schema-guided Event Prediction for 11 Newsworthy Scenarios.” Xinya Du, Zixuan Zhang, [Sha Li](#), Ziqi Wang, Pengfei Yu, Hongwei Wang, Tuan Lai, Xudong Lin, Iris Liu, Ben Zhou, Haoyang Wen, Manling Li, Darryl Hannan, Jie Lei, Hyounghun Kim, Rotem Dror, Haoyu Wang, Michael Regan, Qi Zeng, Qing Lyu, Charles Yu, Carl N. Edwards, Xiaomeng Jin, Yizhu Jiao, Ghazaleh Kazeminejad, Zhenhailong Wang, Chris Callison-Burch, Carl Vondrick, Mohit Bansal, Dan Roth, Jiawei Han, Shih-Fu Chang, Martha Palmer and Heng Ji. NAACL 2022 Demo.
- “RESIN: A Dockerized Schema-Guided Cross-document Cross-lingual Cross-media Information Extraction and Event Tracking System.” Haoyang Wen, Ying Lin, Tuan Lai, Xiaoman Pan, [Sha Li](#), Xudong Lin, Ben Zhou, Manling Li, Haoyu Wang, Hongming Zhang, Xiaodong Yu, Alexander Dong, Zhenhailong Wang, Yi Fung, Piyush Mishra, Qing Lyu, Didac Surs, Brian Chen, Susan Windisch Brown, Martha Palmer, Chris Callison-Burch, Carl Vondrick, Jiawei Han, Dan Roth, Shih-Fu Chang and Heng Ji. NAACL 2021 Demo.

Graph Mining

- “When Do GNNs Work: Understanding and Improving Neighborhood Aggregation.” Yiqing Xie*, [Sha Li](#)*, Carl Yang, Raymond Chi-Wing Wong, Jiawei Han. IJCAI 2020.
- “GMCM: Graph-based Micro-behavior Conversion Model for Post-click Conversion Rate Estimation.” Wentian Bao, Hong Wen, [Sha Li](#), Xiao-Yang Liu, Quan Lin and Keping Yang. SIGIR 2020.
- “NetTaxo: Automated Topic Taxonomy Construction from Text-Rich Network.” Jingbo Shang, Xinyang Zhang, Liyuan Liu, [Sha Li](#), Jiawei Han. WWW 2020.
- “Relation Learning on Social Networks with Multi-Modal Graph Edge Variational Autoencoders.” Carl Yang, Jieyu Zhang, Haonan Wang, [Sha Li](#), Yu Shi, Myunghwan Kim, Matt Walker and Jiawei Han. WSDM 2020.

COMMUNITY SERVICE

Reviewer for ACL, EMNLP, ARR, EACL, AAI

TEACHING EXPERIENCE

Teaching Assistant Introduction to Data Mining for MCS-DS, Head TA

Spring 2018-2020

- Outstanding TA Award, Spring 2019.