

# Muhaao Chen

Assistant Professor of Computer Science, UC Davis.

Email: [muhchen@ucdavis.edu](mailto:muhchen@ucdavis.edu) Homepage: <https://muhaochen.github.io/> Group Website: <https://luka-group.github.io/>

## RESEARCH INTERESTS

**Natural Language Processing:** Robustness, Generalizability, and Indirect Supervision in Machine Learning; AI Safety; Multi-modal and Agentic AI.

## EDUCATION AND ACADEMIC EXPERIENCE

University of California, Davis, CA

2023.11-present

- **Assistant Professor (step 6),** Department of Computer Science
- **Faculty Member,** Graduate Group in Applied Mathematics

University of Southern California, Los Angeles, CA

- **Adjunct Assistant Research Professor,** Department of Computer Science
- **Assistant Research Professor,** Department of Computer Science

2023.11-present

2020.9-2023.10

University of Pennsylvania, Philadelphia, PA

2019.7-2020.8

*Postdoctoral Fellow, Hosted by Dan Roth, Eduardo D. Glandt Distinguished Professor of Computer and Information Science*

University of California, Los Angeles, CA

*Ph.D. in Computer Science*

2014.9-2019.6

Dissertation: Multi-relational Representation Learning and Knowledge Acquisition

Advisors:

- Carlo Zaniolo, *Distinguished Professor of Computer Science, N.E. Friedmann Chair in Knowledge Science* (Committee Chair)
- Kai-Wei Chang, *Associate Professor of Computer Science*
- Wei Wang, *Leonard Kleinrock Chair Professor of Computer Science*

Fudan University, Shanghai, China

*B.S. in Computer Science. Advisor: X. Sean Wang, Dorothean Chair Professor of Computer Science*

2010.9-2014.6

## PUBLICATION

\*Indicating equal contributions.

### Tutorials

- T1 **Muhaao Chen**, Chaowei Xiao, Huan Sun, Lei Li, Leon Derczynski, Anima Anandkumar, Fei Wang. Combating Security and Privacy Issues in the Era of Large Language Models. In **NAACL**, 2024.
- T2 Wenpeng Yin, **Muhaao Chen**, Rui Zhang, Ben Zhou, Fei Wang, Dan Roth. Enhancing LLM Capabilities Beyond Scaling Up. In **EMNLP**, 2024.
- T3 Wenpeng Yin, **Muhaao Chen**, Ben Zhou, Qiang Ning, Kai-Wei Chang, Dan Roth. Indirectly Supervised Natural Language Processing. In **ACL**, 2023.
- T4 **Muhaao Chen**, Lifu Huang, Manling Li, Ben Zhou, Heng Ji, Dan Roth. New Frontiers of Information Extraction. In **NAACL**, 2022.
- T5 **Muhaao Chen**, Hongming Zhang, Qiang Ning, Manling Li, Heng Ji, Kathleen McKeown, Dan Roth. Event-centric Natural Language Processing. In **ACL**, 2021.
- T6 **Muhaao Chen**, Hongming Zhang, Qiang Ning, Manling Li, Heng Ji, Dan Roth. Event-centric Natural Language Understanding. In **AAAI**, 2021.
- T7 Jay Pujara, Pedro Szekely, Huan Sun, **Muhaao Chen**. From Tables to Knowledge: Recent Advances in Table Understanding. In **KDD**, 2021.
- T8 **Muhaao Chen**, Kai-Wei Chang, Dan Roth. Recent Advances in Transferable Representation Learning. In **AAAI**, 2020.

### Refereed Publication in Conference Proceedings

- P1 Qin Liu, Fei Wang, Chaowei Xiao, Muhan Chen. SudoLM: Learning Access Control of Parametric Knowledge with Authorization Alignment. In *the 63<sup>rd</sup> Annual Meeting of the Association for Computational Linguistics (ACL)*, 2025.
- P2 Tenghao Huang, Kinjal Basu, Ibrahim Abdelaziz, Pavan Kapanipathi, Jonathan May, Muhan Chen. R2D2: Remembering, Replay-ing and Dynamic Decision Making with a Reflective Agentic Memory. In *the 63<sup>rd</sup> Annual Meeting of the Association for Computational Linguistics (ACL)*, 2025.
- P3 Weidi Luo, Shenghong Dai, Xiaogeng Liu, Suman Banerjee, Huan Sun, Muhan Chen, Chaowei Xiao. AGrail: A Lifelong Agent Guardrail with Effective and Adaptive Safety Detection. In *the 63<sup>rd</sup> Annual Meeting of the Association for Computational Linguistics (ACL)*, 2025.
- P4 Xiaofei Wen, Wenzuan Zhou, Wenjie Jacky Mo, Muhan Chen. ThinkGuard: Deliberative Slow Thinking Leads to Cautious Guardrails. In *the 63<sup>rd</sup> Annual Meeting of the Association for Computational Linguistics (ACL) - Findings*, 2025.
- P5 Fei Wang, Xingyu Fu, James Y. Huang, Zekun Li, Qin Liu, Xiaogeng Liu, Mingyu Derek Ma, Nan Xu, Wenzuan Zhou, Kai Zhang, Tianyi Lorena Yan, Wenjie Jacky Mo, Hsiang-Hui Liu, Pan Lu, Chunyuan Li, Chaowei Xiao, Kai-Wei Chang, Dan Roth, Sheng Zhang, Hoifung Poon, **Muhan Chen**. MuirBench: A Comprehensive Benchmark for Robust Multi-image Understanding. *International Conference on Learning Representations (ICLR)*, 2025.
- P6 Terry Tong, Fei Wang, Zhe Zhao, **Muhan Chen**. BadJudge: Backdoor Vulnerabilities of LLM-As-A-Judge. *International Conference on Learning Representations (ICLR)*, 2025.
- P7 Yingzi Ma, Jiongxiao Wang, Fei Wang, Siyuan Ma, Jiazhao Li, Jinsheng Pan, Xiuju Li, Furong Huang, Lichao Sun, Bo Li, Yejin Choi, **Muhan Chen**, Chaowei Xiao. Benchmarking Vision Language Model Unlearning via Fictitious Facial Identity Dataset. *International Conference on Learning Representations (ICLR)*, 2025.
- P8 Nan Xu, Fei Wang, Sheng Zhang, Hoifung Poon, Muhan Chen. From Introspection to Best Practices: Principled Analysis of Demonstrations in Multimodal In-Context Learning. In *the 22<sup>nd</sup> Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*, 2025.
- P9 Wenjie Jacky Mo, Jiashu Xu, Qin Liu, Jiongxiao Wang, Jun Yan, Hadi Askari, Chaowei Xiao, Muhan Chen. Test-time Backdoor Mitigation for Black-Box Large Language Models with Defensive Demonstrations. In *the 22<sup>nd</sup> Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL) - Findings*, 2025.
- P10 Yiwei Wang, Muhan Chen, Nanyun Peng, Kai-Wei Chang. Vulnerability of Large Language Models to Output Prefix Jailbreaks: Impact of Positions on Safety. In *the 22<sup>nd</sup> Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL) - Findings*, 2025.
- P11 Hadi Askari, Anshuman Chhabra, Muhan Chen, Prasant Mohapatra. Assessing LLMs for Zero-shot Abstractive Summarization Through the Lens of Relevance Paraphrasing. In *the 22<sup>nd</sup> Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL) - Findings*, 2025.
- P12 Zilan Wang, Junfeng Guo, Jiacheng Zhu, Yiming Li, Heng Huang, **Muhan Chen**, Zhengzhong Tu. SleeperMark: Towards Robust Watermark against Fine-Tuning Text-to-image Diffusion Models. In *the IEEE/CVF Computer Vision and Pattern Recognition Conference (CVPR)*, 2025.
- P13 Tenghao Huang, Donghee Lee, John Sweeney, Jiatong Shi, Emily Steliotes, Matthew Lange, Jonathan May, Muhan Chen. Food-Puzzle: Developing Large Language Model Agents as Flavor Scientists. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD) - Datasets and Benchmarks Track*, 2025.
- P14 Jianpeng Chen, Wangzhi Zhan, Haohui Wang, Zian Jia, Jingru Gan, Junkai Zhang, Jingyuan Qi, Tingwei Chen, Lifu Huang, Muhan Chen, Ling Li, Wei Wang, Dawei Zhou. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD) - Datasets and Benchmarks Track*, 2025.
- P15 Wenzuan Zhou, Sheng Zhang, Yu Gu, **Muhan Chen**, Hoifung Poon. UniversalNER: Targeted Distillation from Large Language Models for Open Named Entity Recognition. *International Conference on Learning Representations (ICLR)*, 2024.
- P16 Xiaogeng Liu, Nan Xu, **Muhan Chen**, Chaowei Xiao. AutoDAN: Generating Stealthy Jailbreak Prompts on Aligned Large Lan-guage Models. *International Conference on Learning Representations (ICLR)*, 2024.
- P17 Qin Liu, Fei Wang, Chaowei Xiao, **Muhan Chen**. From Shortcuts to Triggers: Backdoor Defense with Denoised PoE. In *the 21<sup>st</sup> Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*, 2024.
- P18 Bangzheng Li, Ben Zhou, Fei Wang, Xingyu Fu, Dan Roth, **Muhan Chen**. Deceptive Semantic Shortcuts on Reasoning Chains: How Far Can Models Go without Hallucination? In *the 21<sup>st</sup> Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*, 2024.

- P19 Jiashu Xu, Mingyu Derek Ma, Fei Wang, Chaowei Xiao, **Muhao Chen**. Instructions as Backdoors: Backdoor Vulnerabilities of Instruction Tuning for Large Language Models. In *the 21st Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*, 2024.
- P20 Jiashu Xu, Fei Wang, Mingyu Derek Ma, Pang Wei Koh, Chaowei Xiao, **Muhao Chen**. Instructional Fingerprinting of Large Language Models. In *the 21st Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*, 2024.
- P21 Tianyang Liu, Fei Wang, **Muhao Chen**. Rethinking Tabular Data Understanding with Large Language Models. In *the 21st Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*, 2024.
- P22 Victoria Graf, Qin Liu, **Muhao Chen**. Two Heads are Better than One: Nested PoE for Robust Defense Against Multi-Backdoors. In *the 21st Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*, 2024.
- P23 Lingbo Mo, Boshi Wang, **Muhao Chen**, Huan Sun. How Trustworthy are Open-Source LLMs? An Assessment under Malicious Demonstrations Shows their Vulnerabilities. In *the 21st Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*, 2024.
- P24 Tenghao Huang, Dongwon Jung, Vaibhav Kumar, Mohammad Kachuee, Xiang Li, Puyang Xu, **Muhao Chen**. Planning and Editing What You Retrieve for Enhanced Tool Learning. In *the 21st Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL) - Findings*, 2024.
- P25 Nan Xu, Fei Wang, Ben Zhou, Bangzheng Li, Chaowei Xiao, **Muhao Chen**. Cognitive Overload: Jailbreaking Large Language Models with Overloaded Logical Thinking. In *the 21st Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL) - Findings*, 2024.
- P26 Tianqing Fang, Zhaowei Wang, Wenxuan Zhou, Hongming Zhang, Yangqiu Song, **Muhao Chen**. Getting Sick After Seeing a Doctor? Diagnosing and Mitigating Knowledge Conflicts in Event Temporal Reasoning. In *the 21st Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL) - Findings*, 2024.
- P27 Jiongxiao Wang, Junlin Wu, **Muhao Chen**, Yevgeniy Vorobeychik, Chaowei Xiao. On the Exploitability of Reinforcement Learning with Human Feedback for Large Language Models. In *the 62nd Annual Meeting of the Association for Computational Linguistics (ACL)*, 2024.
- P28 Tianyi Yan, Fei Wang, James Y. Huang, Wenxuan Zhou, Fan Yin, Aram Galstyan, Wenpeng Yin, **Muhao Chen**. Contrastive Instruction Tuning. In *the 62nd Annual Meeting of the Association for Computational Linguistics (ACL) - Findings*, 2024.
- P29 Hanzi Xu, **Muhao Chen**, Lifu Huang, Slobodan Vucetic, Wenpeng Yin. X-Shot: A Unified System to Handle Frequent, Few-shot and Zero-shot Learning Simultaneously in Classification. In *the 62nd Annual Meeting of the Association for Computational Linguistics (ACL) - Findings*, 2024.
- P30 Fei Wang, Wenxuan Zhou, James Y. Huang, Nan Xu, Sheng Zhang, Hoifung Poon, **Muhao Chen**. mDPO: Conditional Preference Optimization for Multimodal Large Language Models. In *the 39th Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2024.
- P31 Xiaofei Wen, Bangzheng Li, Tenghao Huang, **Muhao Chen**. Red Teaming Language Models for Processing Contradictory Dialogues. In *the 39th Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2024.
- P32 Yufei Tian, Tenghao Huang, Miri Liu, Derek Jiang, Alexander Spanger, **Muhao Chen**, Jonathan May, Nanyun Peng. Are Large Language Models Capable of Generating Human-Level Narratives? In *the 39th Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2024. **Outstanding Paper Award**
- P33 Qin Liu, Fei Wang, Nan Xu, Tianyi Lorena Yan, Tao Meng, **Muhao Chen**. Monotonic Paraphrasing Improves Generalization of Language Model Prompting. In *the 39th Conference on Empirical Methods in Natural Language Processing (EMNLP) - Findings*, 2024.
- P34 Terry Tong, Qin Liu, Jiashu Xu, **Muhao Chen**. Securing Multi-turn Conversational Language Models from Distributed Backdoor Attacks. In *the 39th Conference on Empirical Methods in Natural Language Processing (EMNLP) - Findings*, 2024.
- P35 Jiongxiao Wang, Jiazhao Li, Yiquan Li, Xiangyu Qi, Junjie Hu, Yixuan Li, Patrick McDaniel, **Muhao Chen**, Bo Li, Chaowei Xiao. Mitigating Fine-tuning based Jailbreak Attack with Backdoor Enhanced Safety Alignment. In *the 38th Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2024.
- P36 Yu Wang, Xiaogeng Liu, Yu Li, **Muhao Chen**, Chaowei Xiao. AdaShield: Safeguarding Multimodal Large Language Models from Structure-based Attack via Adaptive Shield Prompting. In *The 18th European Conference on Computer Vision (ECCV)*, 2024.
- P37 Tianqing Fang, Wenxuan Zhou, Fangyu Liu, Hongming Zhang, Yangqiu Song, **Muhao Chen**. On-the-fly Denoising for Data Augmentation in Natural Language Understanding. In *the 18th Conference of the European Chapter of the Association for Computational Linguistics (EACL) - Findings*, 2024.

- P38 Qin Liu, Wenjie Mo, Terry Tong, Jiashu Xu, Fei Wang, Chaowei Xiao, **Muhaao Chen**. Mitigating Backdoor Threats to Large Language Models: Advancement and Challenges. In *Proceedings of the 60<sup>th</sup> Annual Allerton Conference (Allerton)*, 2024. **Invited Paper**
- P39 James Y. Huang, Wenlin Yao, Kaiqiang Song, Hongming Zhang, **Muhaao Chen**, Dong Yu. Bridging Continuous and Discrete Spaces: Interpretable Sentence Representation Learning via Compositional Operations. In *the 38th Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2023. **Outstanding Paper Award**
- P40 Haoyu Wang, Hongming Zhang, Yueguan Wang, Yuqian Deng, **Muhaao Chen**, Dan Roth. Are All Steps Equally Important? Benchmarking Essentiality Detection of Events. In *the 38th Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2023.
- P41 Zekun Li, Wenzuan Zhou, Yao-Yi Chiang, **Muhaao Chen**. GeoLM: Empowering Language Models for Geospatially Grounded Language Understanding. In *the 38th Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2023.
- P42 Yiwei Wang, Yujun Cai, **Muhaao Chen**, Yuxuan Liang, Bryan Hooi. Primacy Effect of ChatGPT. In *the 38th Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2023.
- P43 Fei Wang, Wenjie Mo, Yiwei Wang, Wenzuan Zhou, **Muhaao Chen**. A Causal View of Entity Bias in (Large) Language ModelsA Causal View of Entity Bias in (Large) Language Models. In *the 38th Conference on Empirical Methods in Natural Language Processing (EMNLP)* - Findings, 2023.
- P44 Wenzuan Zhou, Sheng Zhang, Hoifung Poon, **Muhaao Chen**. Context-faithful Prompting for Large Language Models. In *In the 38th Conference on Empirical Methods in Natural Language Processing (EMNLP)* - Findings, 2023.
- P45 Nan Xu, Fei Wang, Mingtao Dong, **Muhaao Chen**. Dense Retrieval as Indirect Supervision for Large-space Decision Making. In *the 38th Conference on Empirical Methods in Natural Language Processing (EMNLP)* - Findings, 2023.
- P46 Tenghao Huang, Ehsan Qasemi, Bangzheng Li, He Wang, Faeze Brahman, **Muhaao Chen**, Snigdha Chaturvedi. Affective and Dynamic Beam Search for Story Generation. In *the 38th Conference on Empirical Methods in Natural Language Processing (EMNLP)* - Findings, 2023.
- P47 Shikhar Singh, Ehsan Qasemi, **Muhaao Chen**. VIPHY: Probing “Visible” Physical Commonsense Knowledge. In *the 38th Conference on Empirical Methods in Natural Language Processing (EMNLP)* - Findings, 2023.
- P48 Yiwei Wang, Bryan Hooi, Fei Wang, Yujun Cai, Yuxuan Liang, Wenzuan Zhou, Jing Tang, Manjuan Duan, **Muhaao Chen**. How Fragile is Relation Extraction under Entity Replacements? In *the 27<sup>th</sup> SIGNLL Conference on Computational Natural Language Learning (CoNLL)*, 2023.
- P49 Wenzuan Zhou, Sheng Zhang, Tristan Naumann, **Muhaao Chen**, Hoifung Poon. Continual Contrastive Finetuning Improves Low-Resource Relation Extraction. In *the 61<sup>st</sup> Annual Meeting of the Association for Computational Linguistics (ACL)*, 2023.
- P50 Jiashu Xu, Mingyu Derek Ma, **Muhaao Chen**. Can NLI Provide Proper Indirect Supervision for Low-resource Biomedical Relation Extraction? In *the 61<sup>st</sup> Annual Meeting of the Association for Computational Linguistics (ACL)*, 2023.
- P51 Tanay Dixit, Fei Wang, **Muhaao Chen**. Improving Factuality of Abstractive Summarization without Sacrificing Summary Quality. In *the 61<sup>st</sup> Annual Meeting of the Association for Computational Linguistics (ACL)*, 2023.
- P52 Shudi Hou, Yu Xia, **Muhaao Chen**, Sujian Li. Contrastive Bootstrapping for Label Refinement. In *the 61<sup>st</sup> Annual Meeting of the Association for Computational Linguistics (ACL)*, 2023.
- P53 Fei Wang\*, James Y. Huang\*, Tianyi Yan, Wenzuan Zhou, **Muhaao Chen**. Robust Natural Language Understanding with Residual Attention Debiasing. In *the 61<sup>st</sup> Annual Meeting of the Association for Computational Linguistics (ACL)* - Findings, 2023.
- P54 Keming Lu, I-Hung Hsu, Wenzuan Zhou, Mingyu Derek Ma, **Muhaao Chen**. Multi-hop Evidence Retrieval for Cross-document Relation Extraction. In *the 61<sup>st</sup> Annual Meeting of the Association for Computational Linguistics (ACL)* - Findings, 2023.
- P55 Xinze Li, Yixin Cao, **Muhaao Chen**, Aixin Sun. Take a Break in the Middle: Investigating Subgoals towards Hierarchical Script Generation. In *the 61<sup>st</sup> Annual Meeting of the Association for Computational Linguistics (ACL)* - Findings, 2023.
- P56 Haoyu Wang, Hongming Zhang, Yuqian Deng, Jacob Gardner, Dan Roth, **Muhaao Chen**. Extracting or Guessing? Improving Faithfulness of Event Temporal Relation Extraction. In *the 17<sup>th</sup> Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, 2023.
- P57 Xiaocong Yang, James Y. Huang, Wenzuan Zhou, **Muhaao Chen**. Parameter-Efficient Tuning with Special Token Adaptation. In *the 17<sup>th</sup> Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, 2023.
- P58 Peifeng Wang, Aaron Chan, Filip Ilievski, **Muhaao Chen**, Xiang Ren. PINTO: Faithful Language Reasoning Using Prompted-Generated Rationales. In *the 11<sup>th</sup> International Conference on Learning Representations (ICLR)*, 2023.

- P59 Bonan Kou, **Muhao Chen**, Tianyi Zhang. Automated Summarization of Stack Overflow Posts. In the *Proceedings of the 45th IEEE/ACM International Conference on Software Engineering (ICSE)*, 2023.
- P60 Zhongkai Zhao, Bonan Kou, Mohamed Yilmaz Ibrahim, **Muhao Chen**, Tianyi Zhang. Knowledge-based Version Incompatibility Detection for Deep Learning. In the *31<sup>st</sup> ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE)*, 2023.
- P61 Tai Nguyen, Yifeng Di, Joohan Lee, **Muhao Chen**, Tianyi Zhang. Software Entity Recognition with Noise-Robust Learning. In the *Proceedings of the 38th IEEE/ACM International Conference on Automated Software Engineering (ASE)*, 2023.
- P62 Minh Pham, Craig Knoblock, **Muhao Chen**. Detecting Semantic Errors in Tables using Textual Evidence. In the *Proceedings of IEEE International Conference on Big Data (BigData)*, 2023.
- P63 Wenzuan Zhou, Fangyu Liu, Huan Zhang, **Muhao Chen**. Sharpness-Aware Minimization with Dynamic Reweighting. In the *37th Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2022.
- P64 Nan Xu, Fei Wang, Bangzheng Li, Mingtao Dong, **Muhao Chen**. Does Your Model Classify Entities Reasonably? Diagnosing and Mitigating Spurious Correlations in Entity Typing. In the *37th Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2022.
- P65 Fei Wang, Kaiqiang Song, Hongming Zhang, Lifeng Jin, Sangwoo Cho, Wenlin Yao, Xiaoyang Wang, **Muhao Chen**, Dong Yu. Salience Allocation as Guidance for Abstractive Summarization. In the *37th Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2022.
- P66 Keming Lu, I-Hung Hsu, Wenzuan Zhou, Mingyu Derek Ma, **Muhao Chen**. Summarization as Indirect Supervision for Relation Extraction. In the *37th Conference on Empirical Methods in Natural Language Processing (EMNLP) - Findings*, 2022.
- P67 Zekun Li, Jina Kim, Yao-Yi Chiang, **Muhao Chen**. SpaBERT: Pretrained Language Models on Geographic Data for Geo-Entity Representation. In the *37th Conference on Empirical Methods in Natural Language Processing (EMNLP) - Findings*, 2022.
- P68 Ehsan Qasemi, Filip Ilievski, **Muhao Chen**, Pedro Szekely. PaCo: Preconditions Attributed to Commonsense Knowledge. In the *37th Conference on Empirical Methods in Natural Language Processing (EMNLP) - Findings*, 2022.
- P69 Yiwei Wang, **Muhao Chen**, Wenzuan Zhou, Yujun Cai, Yuxuan Liang, Dayiheng Liu, Baosong Yang, Juncheng Liu, Bryan Hooi. Should We Rely on Entity Mentions for Relation Extraction? Debiasing Relation Extraction with Counterfactual Analysis. In the *20th Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*, 2022.
- P70 Fei Wang, Zhewei Xu, Pedro Szekely, **Muhao Chen**. Robust (Controlled) Table-to-Text Generation with Structure-Aware Equivariance Learning. In the *20th Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*, 2022.
- P71 James Y. Huang, Bangzheng Li, Jiashu Xu, **Muhao Chen**. Unified Semantic Typing with Meaningful Label Inference. In the *20th Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*, 2022.
- P72 Wenzuan Zhou, Qiang Ning, Heba Elfandy, Kevin Small, **Muhao Chen**. Answer Consolidation: Formulation and Benchmarking. In the *20th Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*, 2022.
- P73 Juncheng Liu, Zequn Sun, Bryan Hooi, Yiwei Wang, Dayiheng Liu, Baosong Yang, Xiaokui Xiao, **Muhao Chen**. Dangling-Aware Entity Alignment with Mixed High-Order Proximities. In the *20th Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL) - Findings*, 2022.
- P74 Yiwei Wang, **Muhao Chen**, Wenzuan Zhou, Yujun Cai, Yuxuan Liang, Bryan Hooi. GraphCache: Message Passing as Caching for Sentence-Level Relation Extraction. In the *20th Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL) - Findings*, 2022.
- P75 Wenzuan Zhou\*, Fangyu Liu\*, Ivan Vulić, Nigel Collier, **Muhao Chen**. Prix-LM: Pretraining for Multilingual Knowledge Base Construction. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (ACL)*, 2022.
- P76 Peifeng Wang, Jonathan Zamora, Junfeng Liu, Fangyu Liu, **Muhao Chen**, Xiang Ren. Contextualized Scene Imagination for Generative Commonsense Reasoning. In the *10<sup>th</sup> International Conference on Learning Representations (ICLR)*, 2022.
- P77 Wenzuan Zhou, **Muhao Chen**. An Improved Baseline for Sentence-level Relation Extraction. In the *2<sup>nd</sup> Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics (AAACL)*, 2022.
- P78 Ehsan Qasemi, Piyush Khanna, Qiang Ning, **Muhao Chen**. PInKS: Preconditioned Commonsense Inference with Minimal Supervision. In the *2<sup>nd</sup> Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics (AAACL)*, 2022.
- P79 Jihoon Sohn, Mingyu Derek Ma, **Muhao Chen**. Bending the Future: Autoregressive Modeling of Temporal Knowledge Graphs in Curvature-Variable Hyperbolic Spaces. In the *4<sup>th</sup> Conference on Automated Knowledge Base Construction (AKBC)*, 2022.

- P80 Bonan Kou, Yifeng Di, **Muhao Chen**, Tianyi Zhang. SOSum: A Dataset of Stack Overflow Post Summaries. In *Proceedings of the 19<sup>th</sup> International Conference on Mining Software Repositories (MSR)*, 2022. (Data/Tool Showcase Track)
- P81 Wenxuan Zhou, Fangyu Liu, **Muhao Chen**. Contrastive Out-of-Distribution Detection for Pretrained Transformers. In *Proceedings of the 36th Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2021.
- P82 Wenxuan Zhou, **Muhao Chen**. Learning from Noisy Labels for Entity-Centric Information Extraction. In *Proceedings of the 36th Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2021.
- P83 Haoyu Wang, Hongming Zhang, **Muhao Chen**, Dan Roth. Learning Constraints and Descriptive Segmentation for Subevent Detection. In *Proceedings of the 36th Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2021.
- P84 Xiyang Zhang, **Muhao Chen**, Jonathan May. Salience-Aware Event Chain Modeling for Narrative Understanding. In *Proceedings of the 36th Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2021.
- P85 Fei Wang, Kexuan Sun, Jay Pujara, Pedro Szekely, **Muhao Chen**. Table-based Fact Verification With Salience-aware Learning. In *the 36th Conference on Empirical Methods in Natural Language Processing (EMNLP) - Findings*, 2021.
- P86 Mingyu Derek Ma, **Muhao Chen**, Te-lin Wu, Nanyun Peng. HyperExpan: Taxonomy Expansion with Hyperbolic Representation Learning. In *the 36th Conference on Empirical Methods in Natural Language Processing (EMNLP) - Findings*, 2021.
- P87 Fangyu Liu, **Muhao Chen**, Dan Roth, Nigel Collier. Visual Pivoting for (Unsupervised) Entity Alignment. In *the 35<sup>th</sup> AAAI Conference on Artificial Intelligence (AAAI)*, 2021.
- P88 Cunchao Zhu, **Muhao Chen**, Changjun Fan, Guangquan Cheng, Yan Zhang. Learning from History: Modeling Temporal Knowledge Graphs with Sequential Copy-Generator Networks. In *the 35<sup>th</sup> AAAI Conference on Artificial Intelligence (AAAI)*, 2021.
- P89 Zequn Sun, **Muhao Chen**, Wei Hu. Knowing the No-match: Entity Alignment with Dangling Cases. In *Proceedings of the 59<sup>th</sup> Annual Meeting of the Association for Computational Linguistics (ACL)*, 2021.
- P90 Peifeng Wang, Filip Ilievski, Muhao Chen, Xiang Ren. Do Language Models Perform Generalizable Commonsense Inference? In *Proceedings of the 59<sup>th</sup> Annual Meeting of the Association for Computational Linguistics (ACL) - Findings*, 2021.
- P91 **Muhao Chen**, Weijia Shi, Ben Zhou, Dan Roth. Cross-lingual Entity Alignment with Incidental Supervision. In *the 16<sup>th</sup> Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, 2021.
- P92 Xuelu Chen\*, Michael Boratko\*, **Muhao Chen**, Shib Sankar Dasgupta, Xiang Li, Andrew McCallum. Probabilistic Box Embeddings for Uncertain Knowledge Graph Reasoning. In *the 19<sup>th</sup> Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*, 2021.
- P93 Fei Wang, Kexuan Sun, **Muhao Chen**, Jay Pujara, Pedro Szekely. Retrieving Complex Tables with Multi-Granular Graph Representation Learning. In *Proceedings of the 44<sup>th</sup> ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, 2021.
- P94 Minh Pham, Craig Knoblock, **Muhao Chen**, Binh Vu, Jay Pujara. SPADE: A Semi-supervised Probabilistic Approach for Detecting Errors in Tables. In *Proceedings of the 30<sup>th</sup> International Joint Conference on Artificial Intelligence (IJCAI)*, 2021.
- P95 Kexuan Sun, Fei Wang, **Muhao Chen**, Jay Pujara. Tabular Functional Block Detection with Embedding-based Agglomerative Cell Clustering. In *Proceedings of the 30<sup>th</sup> ACM International Conference on Information and Knowledge Management (CIKM)*, 2021.
- P96 **Muhao Chen**, Hongming Zhang, Haoyu Wang, Dan Roth. “What Are You Trying to Do?” Semantic Typing of Event Processes. In *Proceedings of the 24<sup>th</sup> SIGNLL Conference on Computational Natural Language Learning (CoNLL)*, 2020. **Best Paper Nomination**
- P97 Zequn Sun, **Muhao Chen**, Wei Hu, Chengming Wang. Knowledge Association with Hyperbolic Representation Learning of Knowledge Graphs. In *Proceedings of the 25<sup>th</sup> Conference on Empirical Methods in Natural Language Processing (EMNLP)* 2020.
- P98 Haoyu Wang, **Muhao Chen**, Hongming Zhang, Dan Roth. Joint Constrained Learning for Event-Event Relation Extraction. In *Proceedings of the 25<sup>th</sup> Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2020.
- P99 Hongming Zhang, **Muhao Chen**, Haoyu Wang, Y. Song, Dan Roth. Analogous Process Structure Induction for Sub-event Sequence Prediction. In *Proceedings of the 25<sup>th</sup> Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2020.
- P100 Xuelu Chen, **Muhao Chen**, Changjun Fan, Ankith Uppunda, Yizhou Sun, Carlo Zaniolo. Multilingual Knowledge Graph Completion via Ensemble Knowledge Transfer. In *Proceedings of the 25<sup>th</sup> Conference on Empirical Methods in Natural Language Processing (EMNLP) - Findings*, 2020.
- P101 Zequn Sun, Chengming Wang, Wei Hu, **Muhao Chen**, Jian Dai, Wei Zhang, Yuzhong Qu. Knowledge Graph Alignment Network with Gated Multi-hop Neighborhood Aggregation. In *the 34<sup>th</sup> AAAI Conference on Artificial Intelligence (AAAI)*, 2020.
- P102 Changping Meng, **Muhao Chen**, Jie Mao, Jennifer Neville. ReadNet: A Hierarchical Transformer Framework for Readability Analysis. In *the 42<sup>nd</sup> European Conference on Information Retrieval (ECIR)*, 2020.

- P103 Junheng Hao, Chelsea J. T. Ju, **Muhao Chen**, Yizhou Sun, Carlo Zaniolo, Wei Wang. Bio-JOIE: Joint Representation Learning of Biological Knowledge Bases. In *the 11<sup>th</sup> ACM SIGBio Conference on Bioinform., Comput. Bio. and Health Inform. (ACM-BCB)*, 2020. **SIGBio Best Student Paper Award**
- P104 Tianran Zhang, **Muhao Chen**, Alex Bui. Diagnostic Prediction with Sequence-of-sets Representation Learning for Clinical Events. In *Proceedings of the 18<sup>th</sup> International Conference on Artificial Intelligence in Medicine (AIME)*, 2020
- P105 **Muhao Chen**, Yingtao Tian, Haochen Chen, Kai-Wei Chang, Steve Skiena, Carlo Zaniolo. Learning to Represent Bilingual Dictionaries. In *Proceedings of the 23rd SIGNLL Conference on Computational Natural Language Learning (CoNLL)*, 2019
- P106 Junheng Hao, **Muhao Chen**, Wenchao Yu, Yizhou Sun, Wei Wang. Universal Representation Learning of Knowledge Bases by Jointly Embedding Ontological Concepts and Instances. In *Proceedings of the 25<sup>th</sup> ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, 2019.
- P107 Xuelu Chen, **Muhao Chen**, Weijia Shi, Yizhou Sun, Carlo Zaniolo. Uncertain Knowledge Graphs Embeddings. In *the 33<sup>rd</sup> International Conference on Artificial Intelligence (AAAI)*, 2019.
- P108 **Muhao Chen\***, Weijia Shi\*, Pei Zhou , Kai-Wei Chang. Retrofitting Contextualized Word Embeddings with Paraphrases. In *Proceedings of the 24th Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2019.
- P109 Pei Zhou, Weijia Shi , Jieyu Zhao, Kuan-Hao Huang, **Muhao Chen**, Ryan Cotterell, Kai-Wei Chang. Examining Gender Bias in Languages with Grammatical Gender. In *Proceedings of the 24th Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2019.
- P110 **Muhao Chen**, Chris Quirk. Embedding Edge-attributed Relational Hierarchies. In *Proceedings of the 42<sup>nd</sup> ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, 2019.
- P111 Qingheng Zhang, Zequn Sun, Wei Hu, **Muhao Chen**, Lingbing Guo, Yuzhong Qu. Multi-view Knowledge Graph Embedding for Entity Alignment. In *Proceedings of the 28<sup>th</sup> International Joint Conference on Artificial Intelligence (IJCAI)*, 2019. Also in **ISWC 2020 Invited Sister Conference Track**.
- P112 Zequn Sun, Jiacheng Huang, Wei Hu, **Muhao Chen**, Yuzhong Qu. TransEdge: Translating Relation-contextualized Embeddings for Knowledge Graphs. In *the 18<sup>th</sup> International Semantic Web Conference (ISWC)*, 2019.
- P113 Haochen Chen, Syed Fahad Sultan, Yingtao Tian, **Muhao Chen**, Steven Skiena. Fast and Accurate Network Embeddings via Very Sparse Random Projection. In *Proceedings of the 28<sup>th</sup> ACM International Conference on Information and Knowledge Management (CIKM)*, 2019.
- P114 Changjun Fan, Yuhui Ding, Li Zeng, **Muhao Chen**, Yizhou Sun and Zhong Liu. Learning to Identify High Betweenness Centrality Nodes from Scratch: A Novel Graph Neural Network Approach. In *Proceedings of the 28<sup>th</sup> ACM International Conference on Information and Knowledge Management (CIKM)*, 2019.
- P115 **Muhao Chen**, Changping Meng, Gang Huang, Carlo Zaniolo. Learning to Differentiate Between Main-articles and Sub-articles in Wikipedia. In *IEEE International Conference on Big Data (BigData)*, 2019.
- P116 Yingtao Tian, Haochen Chen, Bryan Perozzi, **Muhao Chen**, Xiaofei Sun, Steven Skiena. Social Relation Inference via Label Propagation. In *the 41<sup>st</sup> European Conference on Information Retrieval (ECIR)*, 2019.
- P117 Qi Zhao, **Muhao Chen**, Pengyuan Du, Tuan Le, Mario Gerla. Towards Efficient Cellular Traffic Offloading via Dynamic MPTCP Path Configuration with SDN. In *IEEE International Conference on Computing, Networking and Communications (ICNC)*, 2019.
- P118 **Muhao Chen**, Gang Huang, Changping Meng, Carlo Zaniolo. Neural Article Pair Modeling for Wikipedia Sub-article Matching. In *the 29<sup>th</sup> European Conference on Machine Learning (ECML)*, 2018 (**Plenary Presentation**, ~1.7% acceptance rate)
- P119 **Muhao Chen**, Yingtao Tian, Kai-Wei Chang, Steven Skiena, Carlo Zaniolo. Co-training Embeddings of Knowledge Graphs and Entity Descriptions for Cross-lingual Entity Alignment. In *the 27<sup>th</sup> International Joint Conference on Artificial Intelligence (IJCAI)*, 2018.
- P120 **Muhao Chen**, Yingtao Tian, Xuelu Chen, Zijun Xue, Carlo Zaniolo. On2Vec: Embedding-based Relation Prediction for Ontology Population. In *Proceedings of the 17<sup>th</sup> SIAM International Conference on Data Mining (SDM)*. SIAM, 2018
- P121 Haochen Chen, Xiaofei Sun, Yingtao Tian, Bryan Perozzi, **Muhao Chen** and Steven Skiena. Enhanced Network Embeddings via Exploiting Edge Labels. In *the 27th ACM Conference on Information and Knowledge Management (CIKM)*. ACM 2018.
- P122 Pengyuan Du, Seunghyun Yoo, Qi Zhao, **Muhao Chen**, Mario Gerla. Towards Opportunistic Resource Sharing in Mobile Social Networks - an Evolutionary Game Theoretic Approach. In *Proceedings of the 19th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, ACM 2018.

P123 **Muhao Chen**, Qi Zhao, Pengyuan Du, Carlo Zaniolo, Mario Gerla. Demand-driven Cache Allocation Based on Context-aware Collaborative Filtering. In *Proceedings of the 19th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, ACM 2018.

P124 **Muhao Chen**, Yingtao Tian, Mohan Yang, Carlo Zaniolo. Multilingual Knowledge Graph Embeddings for Cross-lingual Knowledge Alignment. In *Proceedings of the 26<sup>th</sup> International Joint Conference on Artificial Intelligence (IJCAI)*. 2017.

P125 **Muhao Chen**, Shi Gao, X. Sean Wang. Converting Spatiotemporal Data Among Heterogeneous Granularity Systems. In *Proceedings of the 25<sup>th</sup> IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. IEEE, 2016.

## Refereed Journal Publication and Book Chapters

- J1 James Y. Huang, Wenzuan Zhou, Fei Wang, Fred Morstatter, Sheng Zhang, Hoifung Poon, Muhao Chen. Offset Unlearning for Large Language Models. *Transactions on Machine Learning Research (TMLR)*. 2025
- J2 Juan Zambrano Chaves, Mr Shih-Cheng Huang, Yanbo Xu, Mr Hanwen Xu, Dr Naoto Usuyama, Dr Sheng Zhang, Fei Wang, Yujia Xie, Mahmoud Khademi, Ziyi Yang, Hany Awadalla, Julia Gong, Houdong Hu, Jianwei Yang, Chunyuan Li, Jianfeng Gao, Yu Gu, Cliff Wong, Mu Wei, Tristan Naumann, Muhao Chen, Matthew Lungren, Akshay Chaudhari, Serena Yeung-Levy, Curtis Langlotz, Sheng Wang, Hoifung Poon. A Clinically Accessible Small Multimodal Radiology Model and Evaluation Metric for Chest X-Ray Findings. *Nature Communications*. 16 (1). Nature Portfolio. 2025.
- J3 Yao-Yi Chiang, **Muhao Chen**, Weiwei Duan, Jina Kim, Craig Knoblock, Stefan Leyk, Zekun Li, Yijun Lin, Min Namgung, Basel Shbita, Johannes H. Uhl. GeoAI for the Digitization of Historical Maps. *Handbook of GeoAI*. Taylor & Francis, 2023 (Book Chapter)
- J4 Bangzheng Li, Wenpeng Yin, **Muhao Chen**. Ultra-fine Entity Typing with Indirect Supervision from Natural Language Inference. *Transactions of the Association for Computational Linguistics (TACL)*. MIT Press, 2022.
- J5 Tianran Zhang, **Muhao Chen**, Alex Bui. AdaDiag: Adversarial Domain Adaptation of Diagnostic Prediction with Clinical Event Sequences. *Journal of Biomedical Informatics (JBI)*, vol. 134. Elsevier, 2022.
- J6 Mohammad Rostami, Hangfeng He, **Muhao Chen**, Dan Roth. Transfer Learning via Representation Learning. *Federated and Transfer Learning*. Springer, 2022 (Book Chapter)
- J7 Jyun-Yu Jiang, Chelsea J.-T. Ju, Junheng Hao, **Muhao Chen**, Wei Wang. Circular RNA Prediction based on Junction Encoders and Deep Interaction among Splice Sites. *Bioinformatics*, vol. 37. Oxford University Press, 2021. Full Paper of **ISMB/ECCB**, 2021.
- J8 Guangyu Zhou\*, **Muhao Chen**\*, Chelsea J. T. Ju\*, Zheng Wang, Jyun-Yu Jiang, Wei Wang. Mutation effect estimation on protein-protein interactions using deep contextualized representation learning. *NAR Genom. Bioinform*, vol. 2 (2). Oxford University Press. 2020.
- J9 Zequn Sun, Qingheng Zhang, Wei Hu, Chengming Wang, **Muhao Chen**, Chengkai Li, Yuzhong Qu. A Benchmarking Study of Embedding-based Entity Alignment for Knowledge Graphs. *Proceedings of the VLDB Endowment (PVLDB)*, vol. 13. ACM. 2020
- J10 **Muhao Chen**\*, Chelsea J. T. Ju\*, Guangyu Zhou, Tianran Zhang, Kai-Wei Chang, Carlo Zaniolo, Wei Wang. Multifaceted Protein-Protein Interaction Prediction Based on Siamese Residual RCNN. *Bioinformatics*, vol. 35 (14) Oxford University Press. Full Paper of **ISMB/ECCB**, 2019.
- J11 Carlo Zaniolo, Shi Gao, Maurizio Atzori, **Muhao Chen**, Jiaqi Gu. User-Friendly Temporal Queries on Historical Knowledge Bases. *Information and Computation*, Vol. 259 (3). Elsevier, 2018.

## Refereed Workshop and System Demonstration Papers

- w1 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. MetaScientist: A Human-AI Synergistic Framework for Automated Mechanical Metamaterial Design. In *the 22<sup>nd</sup> Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*, 2025. (demo)
- w2 Weijia Shi, **Muhao Chen**, Yingtao Tian, Kai-Wei Chang. Learning Bilingual Word Embeddings Using Lexical Definitions. In *Proceedings of ACL Workshop on Representation Learning for NLP (RepL4NLP)*, 2019.
- w3 Zhubo Deng, Pei Zhou, Weijia Shi, **Muhao Chen**, Kai-Wei Chang. Computational Analysis of French-origin Reborrowing Process for English Loanwords. In *ICDM Workshop on Multilingual Cognitive Services (ICDMW)*, 2019
- w4 Changjun Fan, Yizhou Sun, Li Zeng, Yang-Yu Liu, **Muhao Chen**, Zhong Liu. Dismantle Large Networks through Deep Reinforcement Learning. In *ICLR Workshops*, 2019.
- w5 Pei Zhou, **Muhao Chen**, Kai-Wei Chang, Carlo Zaniolo. Quantification and Analysis of Scientific Language Variation by Research Fields. In *Proceedings of the ICDM Workshops (ICDMW)*, 2018.

- w6 **Muhao Chen**, Tao Zhou, Pei Zhou, Carlo Zaniolo. Multi-graph Affinity Embeddings for Multilingual Knowledge Graphs. **Contributed talk** in the 6th Workshop on Automated Knowledge Base Construction at NIPS (AKBC). 2017.
- w7 **Muhao Chen**, Carlo Zaniolo. Learning Multi-faceted Knowledge Graph Embeddings for Natural Language Processing. In Proceedings of the 26<sup>th</sup> International Joint Conference on Artificial Intelligence (IJCAI). 2017 (Extended abstract)
- w8 Tao Zhou, **Muhao Chen**, Demetri Terzopoulos, Jie Yu. Attention-based Natural Language Person Retrieval. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). IEEE, 2017.
- w9 Shi Gao, **Muhao Chen**, Maurizio Atzori, Carlo Zaniolo. SPARQL<sup>T</sup> and its User-Friendly Interface for Managing and Querying the History of Knowledge Bases. In the 14<sup>th</sup> International Semantic Web Conference (ISWC), 2015 (demo).

## AWARDS

- *Outstanding Paper Award*. EMNLP 2024
- *Outstanding Paper Award*. EMNLP 2023
- *Amazon Research Award*. 2023
- *Keston Exploratory Research Award*. 2023
- *Amazon Research Award*. 2022
- *Cisco Faculty Research Award*. 2022
- *NSF CISE Research Initiation Initiative (CRII) Award*. 2021
- *ACM SIGBio Best Student Paper Award*. 2020
- *UCLA Dissertation Fellowship*. 2018-2019
- *Tung OOCL Scholarships*. Tung's Foundation of Hong Kong & The Oriental Overseas Container Line. 2012, 2013
- *Wang-Dao Fellowship*. President of Fudan University. 2013
- *Chun-Tsung Fellowship*. Hui-Chun Chin and Tsung-Dao Lee Chinese Undergraduate Research Endowment. 2012

## FUNDING

- NSF CICI: UCSS: Enabling a Safe and Directive Multi-modal Foundation Model Ecosystem for Food Science Research** 2023-2026  
• \$200K my share. (Co-PI) 2025-2028
- Amazon Trusted AI Prize** 2024-2025  
• \$250K research fund. (PI)
- DARPA FoundSci: MetaScientist: Bioinspired Meta-Material Design with an Autonomous Scientist** 2024-2025  
• \$180K my share. (Co-PI)
- Amazon Research Award: Robust (Controlled) NLG with Structure -Aware Equivariance Learning** 2023-2024  
• \$70K unrestricted fund + \$40K AWS credits. Spring 2023 call. (Sole PI)
- NSF Proto-OKN: Knowledge Graph Construction for Resilient, Trustworthy, and Secure Software Supply Chains** 2023-2026  
• NSF ITR Grant, \$370K my share. (Co-PI)
- Amazon Research Award: On Faithfulness of Information Extraction** 2023-2024  
• \$73.8K unrestricted fund + \$20K AWS credits. Spring 2022 call. (Sole PI)
- Keston Exploratory Research Award: Multi-document Newsworthy Event Monitoring and Forecasting** 2023-2024  
• \$100K unrestricted fund from Keston Family Foundation. (PI)
- Cisco Research Award: Robust Knowledge Extraction from Text.** 2022-2023  
• \$70.4K unrestricted fund. (Sole PI)
- DARPA KMASS Phase 1: Knowledge Needed in Context.** 2022-2023  
• \$308K my share. (Co-PI)
- Improved Performance, Analytics and Summarization of Synergistic Anticipation of Geopolitical Events.** 2022-2023  
• ARLIS research grant. \$50K my share. (Co-PI)
- NSF CRII: III: Knowledge Graph Completion with Transferable Representation Learning.** 2021-2023  
• NSF IIS Grant, \$175K. (Sole PI)
- DARPA MCS: Multi-modal Open World Grounded Learning and Inference.** 2020-2023

- \$600K my share. (Senior Personnel)

## TEACHING AND MENTORING

### Teaching

ECS 271: Machine Learning and Discovery

- Spring 2024 (UC Davis), Spring 2025 (UC Davis)

ECS 289G / ECS 282: Advanced Natural Language Processing

- Winter 2024 (UC Davis), Winter 2025 (UC Davis)

CSCI 544: Applied Natural Language Processing

- Spring 2022 (USC)

### Lab Members/Students

*Ph.D. Students:*

Dongwon Jung, Ph.D. Student in Computer Science.

Fall 2024-date

Xiaofei Wen, Ph.D. Student in Computer Science

Fall 2024-date

Jacky Mo, Ph.D. Student in Computer Science (Co-advised with Zhe Zhao)

Fall 2024-date

Rui Cai, Ph.D. Student in Computer Science (Co-advised with Zhe Zhao)

Fall 2024-date

Hadi Askari, Ph.D. Student in Computer Science

Winter 2024-date

Bangzheng Li, Ph.D. Student in Computer Science.

Fall 2022-date

- **Provost PhD Fellow.**

- Joined the lab as an undergraduate researcher in Summer 2021

Qin (Jacqueline) Liu, Ph.D. Student in Computer Science.

Fall 2022-date

Tenghao Huang, Ph.D. Student in Computer Science (Viterbi-ISI Fellow)

Fall 2022-date

Nan (Nancy) Xu, Ph.D. Student in Computer Science

Spring 2022-date

James Y. Huang, PhD Student in Computer Science.

Fall 2021-date

Fei Wang, Ph.D. Student in Computer Science

Fall 2020-date

- Viterbi Honors Program; USC CS Departmental Best Research Award; Annenberg Fellow. Amazon ML Fellowship.

*M.S. and Undergraduate Students:*

Keming (Luke) Lu, MS Student in Industrial and System Engineering.

Fall 2021-date

Shikhar Singh, MS Student in Computer Science.

Fall 2021-Summer 2022

Tianyi (Lorena) Yan, Undergraduate Student, Computer Science (CURVE Fellowship)

Fall 2022-date

Jacky Mo, Undergraduate Student, Computer Science (CURVE Fellowship)

Fall 2022-Spring 2024

Mingtao Dong, Undergraduate Student, Computer Science (Provost's Research Fellowship)

Spring 2022-date

Jiashu Xu, Undergraduate Student, Computer Science (CURVE Fellowship)

Fall 2021- Spring 2022

Zhewei Xu, Undergraduate Student, Computer Science (CURVE Fellowship)

Fall 2021-Spring 2022

Logan Norman, Undergraduate Student, Computer Science (Viterbi Fellowship)

Fall 2021-Spring 2022

### Alumni

Wenxuan Zhou, PhD in Computer Science, USC. 2023

- Dissertation Title: *Robust and Generalizable Knowledge Acquisition from Text*

- Employment: Research Scientist at Google Deepmind

Eric (Ehsan) Qasemi, PhD in Computer Science, USC. 2023 (Co-advised with Pedro Szekely)

- Dissertation Title: *Multi-modal Preconditioned Inference of Commonsense Knowledge*

- Employment: ML Scientist at Oracle Cloud

Peifeng Wang, PhD in Computer Science, USC. 2023 (Co-advised with Xiang Ren)

- Dissertation: Externalized Reasoning in Language Models for Scalable, Trustworthy AI

- Employment: Research Scientist at Salesforce AI

### Visiting Scholars/Students

Tianqing Fang. PhD Student in Computer Science, HKUST (Hong Kong PhD Fellow)

Summer 2022-Spring 2023

Xiaocong Yang. Undergraduate, Tsinghua University. (Joined UIUC as a Master Student)

Summer-Fall 2022

Amani Rune Maina-Kilaas. Undergraduate, Harvey Mudd College (NSF REU Student; Joined MIT as a PhD Student) Summer 2022

Tanay Dixit. Undergraduate, IIT Madras, Computer Science (NSF REU Student; Joined UIUC as a Master Student)	Summer 2022
Devadutta Dash. Undergraduate, IIT Varanasi, Computer Science (IUSSFT-Viterbi Research Intern)	Summer 2022
Jihoon Sohn. USC PhD Student in Mathematics	Spring 2021-Spring 2022
Bangzheng Li. UIUC, B.S. in Computer Science (Now PhD student in my lab)	Summer 2021-Fall 2021
Shanxiu He. Undergraduate, UCLA Computer Science (USC ISI NLP Intern; Joined UCSB as a PhD Student)	Summer 2021
Piyush Khanna, Undergraduate, Delhi Tech Univ., Computer Science (Joined UCSD as a Master Student)	Summer 2021
Yueguan Wang. Undergraduate, Tsinghua Univ., EE (USC-THU Summer Research Program)	Summer 2021

### Dissertation/Qualification Committee

#### *Dissertation Committee:*

Jihoon Sohn, USC PhD in Mathematics, 2022. (Dissertation Chair: Francis Bonahon)  
Minh Pham, USC PhD in Computer Science, 2022. (Dissertation Chair: Craig Knoblock)  
Xuelu (Shirley) Chen, UCLA PhD in Computer Science, 2021. (Dissertation Chair: Carlo Zaniolo)  
Tianran Zhang, UCLA PhD in Bioengineering, 2022. (Dissertation Chair: Alex Bui)  
Jeong Hyun An, USC M.S. in Computer Science, 2022. (Dissertation Committee Chair)

#### *Qualification Committee:*

Hsiao-Ying Lu, UC Davis PhD student in Computer Science (Qualification Committee)  
Soroush Koohpayegani, UC Davis PhD Student in Computer Science (Dissertation Committee)  
Sairamvinay Vijayaraghavan, UC Davis PhD student in Computer Science (Qualification Committee)  
I-Hung Hsu, USC PhD Student in Computer Science. (Qualification Committee)  
Haowen Lin, USC PhD Student in Computer Science. (Qualification Committee)  
Jianzhi Yang, USC PhD Student in Quantitative and Computational Biology. (Qualification Committee)  
Tejas Srinivasan, USC PhD Student in Computer Science. (Qualification Committee)  
Binh Vu, USC PhD Student in Computer Science. (Qualification Committee)  
Yizhou Zhang, USC PhD Student in Computer Science. (Qualification Committee)  
Michiel de Jong, USC PhD Student in Computer Science. (Qualification Committee)  
Kexuan Sun, USC PhD Student in Computer Science. (Qualification Committee)

## PROFESSIONAL SERVICE

### Conference Organization/Senior Committee Member:

2024: NAACL (Area Chair), EMNLP (Area Chair), ACL (Area Chair), LREC-COLING (Area Chair – Information Extraction)  
2023: AAAI (Area Chair), ACL (Senior Area Chair – Information Retrieval and Text Mining), ACL (Area Chair), EMNLP (Area Chair – Information Extraction), NLPCC (Area Chair)  
2022: AAAI (Senior PC), NAACL (Senior Area Chair), NAACL SRW (Faculty Mentor), EMNLP (Area Chair), AACL (Area Chair)  
2021: AAAI (Senior PC), IJCAI (Senior PC)  
2019: IEEE AIKE (Doctoral Consortium Chair)

### Workshop Organization Member:

AAAI spring Symposium on AI for Engineering and Scientific Discoveries, 2025  
LLMSEC@ACL, 2025  
AI4Research@IJCAI, 2024  
Indirect, Weak and Self Supervision for Knowledge Extraction (Wise-Supervision@AKBC), 2022  
Deep Learning on Graphs for Natural Language Processing (DLG4NLP@NAACL), 2022

### PC Member:

2021: ACL, EACL, NAACL, WSDM, WWW.  
2020: AAAI, ACL-IJCNLP, AKBC, COLING, EMNLP, IJCAI, ISWC, KDD, SIGIR, WSDM, \*SEM.  
2019: AAAI, AKBC, NAACL, ACL, EMNLP-IJCNLP, BigData, NLPCC, WISE, ICSC.  
2018: AAAI, EMNLP, BigData, NLPCC, SoCal NLP.

### Society Organization

ACL Special Interest Group on NLP Security (SIGSEC) -- Founding Officer and Secretary, 2023.9-date

**Editorial Board:** Frontiers in Big Data

**Journal Reviewer:** TPAMI, TACL, AIJ, TNNLS, TASLP, Pattern Recognition, Bioinformatics, PLOS Computational Biology, Briefings in Bioinformatics, Cell Systems, Clinical and Translational Medicine, Comput. & Struct. Biotechnol., TKDD, TOIS, W3J, TKDE, TII, BMC Genomics, BMC Medical Genomics, BMC Human Genetics, Human Genomics, Quantitative Biology, GeoInformatica, Information Sciences, Nature Machine Intelligence.

**Panelist:**

2025: DoE Panel

2024: DoE Panel

2021, 2022: NSF CISE Core Panels

**University Service**

Department of Computer Science, UC Davis <b>CS Colloquia Organizer</b>	2025.1-date
Departments of Linguistics and Statistics, UC Davis <b>Faculty Hiring Committee</b>	2025
UC Davis, Fellowship Review Committee	2024
UC LFRP, Planning Committee	2024
Department of Computer Science, UC Davis <b>Faculty Hiring Committee</b>	2024
NSF REU Site, <b>Co-organizer</b>	2022, 2023
USC ISI Institutional AI Seminar, <b>Organizer</b>	2021.7-2022.12
Department of Computer Science, USC <b>Faculty Hiring Committee</b>	2021, 2022
Viterbi School of Engineering, USC <b>Fellowship Committee</b>	2021
Department of Computer Science, USC <b>PhD Admission Fellowship Committee</b>	2021

## PRESENTATIONS

*Invited talks, colloquia and tutorials*

1. Mitigating Backdoor Threats to Large Language Models. *EECS Colloquium, UC Merced*. Nov 2024
2. Mitigating Backdoor Threats to Large Language Models. *Invited talk, Allerton Conference*. Sept 2024
3. Mitigating Backdoor Threats to Large Language Models. *Invited talk, Adobe*. Aug 2024
4. Combating Security and Privacy Issues in the Era of Large Language Models. *Half-day Tutorial at NAACL*. June 2024
5. Indirectly Supervised Natural Language Processing. *Half-day Tutorial at ACL*. July, 2023.
6. Robust and Indirectly Supervised Knowledge Acquisition. *CS Research Colloquium, ASU*. March 2023
7. Robust and Indirectly Supervised Knowledge Acquisition. *CS Research Colloquium, UMN*. March 2023
8. Robust and Indirectly Supervised Knowledge Acquisition. *CS Research Colloquium, UC Davis*. Feb 2023
9. Robust and Indirectly Supervised Knowledge Acquisition. *CS Research Colloquium, UCSB*. Feb 2023
10. Robust and Indirectly Supervised Information Extraction. *Invited talk, AI & ML Seminar, UCI*. Nov 2022
11. Robust and Indirectly Supervised Information Extraction. *Invited talk, NL Seminar, UMN*. Nov 2022
12. Robust and Indirectly Supervised Information Extraction. *Invited talk, CS Department Seminar (CS201), UCLA*. Oct 2022.
13. Robust and Indirectly Supervised Information Extraction. *Invited talk, CS Department Seminar (remote), Nanjing University*. Sept 2022.
14. Robust and Indirectly Supervised Information Extraction. *Invited talk, Apple, San Jose*. Aug 2022.
15. Robust and Indirectly Supervised Information Extraction. *Invited talk, Microsoft CSR Distinguished Talk Series*. July 2022.
16. New Frontiers of Information Extraction. *Half-day Tutorial at NAACL*. July, 2022.
17. Understanding Event Processes in Natural Language. *Invited talk, NLP Seminar, UC Santa Cruz*. April 2022.
18. Understanding Event Processes in Natural Language. *Invited talk, CS Department Seminar, Rutgers–New Brunswick*. March 2022.
19. Understanding Event Processes in Natural Language. *Invited talk, ML Seminar, Purdue University*. Sept 2021.
20. Understanding Event Processes in Natural Language. *Invited talk, Tencent AI Lab, Seattle*. Sept 2021.

21. Understanding Event Processes in Natural Language. *Invited talk, University of Central Florida*. Aug 2021.
22. Understanding Event Processes in Natural Language. *Invited talk, IBM Research Almaden*. July 2021.
23. Understanding Event Processes in Natural Language. *Invited talk, CS Department Seminar, Fudan University*. June 2021.
24. Understanding Event Processes in Natural Language. *Invited talk, CS Department Seminar, Nanjing University*. June 2021.
25. Understanding Event Processes in Natural Language. *Invited talk, NLP Seminar, National University of Singapore*. May 2021.
26. Understanding Event Processes in Natural Language. *Invited talk, NLP Seminar, OSU*. May 2021
27. Understanding Event Processes in Natural Language. *Invited talk, CS Research Colloquium, UCSB*. April 2021
28. Understanding Event Processes in Natural Language. *Invited talk, Frontier Topics in Vision and Language, ASU*. Mar 2021.
29. Understanding Event Processes in Natural Language. *Invited talk at Machine Learning and Big Data Seminar, UCLA*. Nov 2020
30. Event-Centric Natural Language Processing. *Half-day Tutorial at ACL*. Aug, 2021.
31. From Tables to Knowledge: Recent Advances in Table Understanding. *Half-day Tutorial at KDD*. Aug, 2021.
32. Understanding Event Processes in Natural Language. Understanding Event Processes in Natural Language. *Invited talk at Language Technology Seminar, University of Cambridge*. Nov 2020
33. Event-Centric Natural Language Understanding. *Half-day Tutorial at AAAI*. Feb, 2021.
34. Knowledge Acquisition with Transferable Representation Learning. *CS Research Colloquium, USC*. Nov 2020
35. Knowledge Acquisition with Transferable Representation Learning. *AI Seminar, USC ISI*. Los Angeles, Jan 2020.
36. Knowledge Acquisition with Transferable Representation Learning. *IBM Research Almaden*. San Jose, Jan 2020.
37. Knowledge Acquisition with Transferable Representation Learning. *Invited talk at FDSiF. Fudan Univ., Shanghai, China*. Dec 2019.
38. Knowledge Acquisition with Transferable Representation Learning. *Invited talk at SNAP Seminar, Stanford University*. Jan 2019.
39. Knowledge Acquisition with Transferable Representation Learning. *Machine Learning and Big Data Seminar, UCLA*. Jan 2019.
40. Recent Advances in Transferable Representation Learning. *Half-day tutorial at AAAI*. NYC, NY, Feb 2020.
41. Neural Article Pair Modeling for Wikipedia Sub-article Matching. *Google Search Intelligence Seminar Talk*. Sept 2017
42. Reasoning Across Multiple Spatiotemporal Granularity Systems. *Invited talk at Teradata Labs. El Segundo, CA, USA*. Mar 2015

## INDUSTRIAL EXPERIENCE

Microsoft Research, Redmond, WA (NLP Group)	2018.6~2018.9
• <b>Research Intern</b> , <i>Embedding Edge-attributed Relational Hierarchies</i> . [P110]	
Google, Mountain View, CA (Google Knowledge Graph)	2017.6~2017.9
• <b>Research Intern</b> , <i>Neural Article Pair Modeling for Large-scale Sub-article Relation Extraction</i> . [P118,P115,P102]	
Google, Mountain View, CA (Procella Real-time Data Infrastructure)	2016.6~2016.9
• <b>System SDE Intern</b> , <i>Dynamic Shard-partitioning for Large-scale Windowed Event Streams</i> .	
Teradata Labs, Los Angeles, CA (Optimizer Group)	2015.6~2015.9
• <b>R&amp;D Intern</b> , <i>Multi-task Learning for Cost-based Database Optimizers</i> .	