

# Muhao Chen

Assistant Professor (Research) of Computer Science, USC.

Email: [muhaoche@usc.edu](mailto:muhaoche@usc.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; Knowledge Acquisition from Unstructured Data; Knowledge-driven AI for Interdisciplinary Tasks (Computational Biology, Medicine, and Geoinformatics).

## EDUCATION AND ACADEMIC EXPERIENCE

University of Southern California, Los Angeles, CA

2020.9-date

- **Assistant Professor (Research)**, Department of Computer Science
- **Research Lead**, Information Sciences Institute
- **Director**, USC Language Understanding and Knowledge Acquisition Group

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. ✉Indicating corresponding authors of journal publications.

*Published in top-tier NLP and AI venues such as TACL, EMNLP, ACL, NAACL, CoNLL, AAAI, ICLR, SIGIR, etc.*

### Tutorials

- T1 Wenpeng Yin, **Muhao Chen**, Ben Zhou, Qiang Ning, Kai-Wei Chang, Dan Roth. Indirectly Supervised Natural Language Processing. In **ACL**, 2023.
- T2 **Muhao Chen**, Lifu Huang, Manling Li, Ben Zhou, Heng Ji, Dan Roth. New Frontiers of Information Extraction. In **NAACL**, 2022.
- T3 **Muhao Chen**, Hongming Zhang, Qiang Ning, Manling Li, Heng Ji, Kathleen McKeown, Dan Roth. Event-centric Natural Language Processing. In **ACL**, 2021.
- T4 **Muhao Chen**, Hongming Zhang, Qiang Ning, Manling Li, Heng Ji, Dan Roth. Event-centric Natural Language Understanding. In **AAAI**, 2021.
- T5 Jay Pujara, Pedro Szekely, Huan Sun, **Muhao Chen**. From Tables to Knowledge: Recent Advances in Table Understanding. In **KDD**, 2021.
- T6 **Muhao Chen**, Kai-Wei Chang, Dan Roth. Recent Advances in Transferable Representation Learning. In **AAAI**, 2020.

### Refereed Journal Publication and Book Chapters

- J1 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.
- J2 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.
- J3 Mohammad Rostami, Hangfeng He, **Muhao Chen**, Dan Roth. Transfer Learning via Representation Learning. *Federated and Transfer Learning*. Springer, 2022 (Book Chapter)

- J4 Jyun-Yu Jiang, Chelsea J.-T. Ju, Junheng Hao, **Muhao Chen**, Wei Wang<sup>✉</sup>. 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.
- J5 Guangyu Zhou\*, **Muhao Chen**<sup>✉\*</sup>, Chelsea J. T. Ju\*, Zheng Wang, Jyun-Yu Jiang, Wei Wang<sup>✉</sup>. Mutation effect estimation on protein-protein interactions using deep contextualized representation learning. **NAR Genom. Bioinform**, vol. 2 (2). Oxford University Press. 2020.
- J6 Zequn Sun, Qingheng Zhang, Wei Hu<sup>✉</sup>, 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
- J7 **Muhao Chen**<sup>✉\*</sup>, 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.
- J8 Carlo Zaniolo<sup>✉</sup>, 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 Publication in Conference Proceedings

- P1 Wenxuan Zhou, Sheng Zhang, Tristan Naumann, **Muhao 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.
- P2 Jiashu Xu, Mingyu Derek Ma, **Muhao 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.
- P3 Tanay Dixit, Fei Wang, **Muhao 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.
- P4 Shudi Hou, Yu Xia, **Muhao Chen**, Sujian Li. Contrastive Bootstrapping for Label Refinement. In *the 61<sup>st</sup> Annual Meeting of the Association for Computational Linguistics (ACL)*, 2023.
- P5 Fei Wang\*, James Y. Huang\*, Tianyi Yan, Wenxuan Zhou, **Muhao 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.
- P6 Keming Lu, I-Hung Hsu, Wenxuan Zhou, Mingyu Derek Ma, **Muhao 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.
- P7 Xinze Li, Yixin Cao, **Muhao 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.
- P8 Haoyu Wang, Hongming Zhang, Yuqian Deng, Jacob Gardner, Dan Roth, **Muhao 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.
- P9 Xiaocong Yang, James Y. Huang, Wenxuan Zhou, **Muhao 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.
- P10 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.
- P11 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.
- P12 Peifeng Wang, Aaron Chan, Filip Ilievski, **Muhao Chen**, Xiang Ren. PINTO: Faithful Language Reasoning Using Prompted-Generated Rationales. In *the 11<sup>th</sup> International Conference on Learning Representations (ICLR)*, 2023.
- P13 Wenxuan 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.
- P14 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.
- P15 Fei Wang, Kaiqiang Song, Hongming Zhang, Lifeng Jin, Sangwoo Cho, Wenlin Yao, Xiaoyang Wang, **Muhao Chen**, Dong Yu. Saliency Allocation as Guidance for Abstractive Summarization. In *the 37th Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2022.
- P16 Keming Lu, I-Hung Hsu, Wenxuan 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.

- P17 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.
- P18 Ehsan Qasemi, Filip Ilevski, **Muhao Chen**, Pedro Szekely. PaCo: Preconditions Attributed to Commonsense Knowledge. In *the 37th Conference on Empirical Methods in Natural Language Processing (EMNLP)* - Findings, 2022.
- P19 Yiwei Wang, **Muhao Chen**, Wenxuan 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.
- P20 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.
- P21 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.
- P22 Wenxuan Zhou, Qiang Ning, Heba Elfardy, 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.
- P23 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.
- P24 Yiwei Wang, **Muhao Chen**, Wenxuan 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.
- P25 Wenxuan 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.
- P26 Peifeng Wang, Jonathan Zamora, Junfeng Liu, Filip Ilevski, **Muhao Chen**, Xiang Ren. Contextualized Scene Imagination for Generative Commonsense Reasoning. In *the 10<sup>th</sup> International Conference on Learning Representations (ICLR)*, 2022.
- P27 Wenxuan 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 (AACL)*, 2022.
- P28 Ehsan Qasemi, Piyush Khanna, Qiang Ning, **Muhao Chen**. PlnKS: Preconditioned Commonsense Inference with Minimal Supervision. In *the 2<sup>nd</sup> Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics (AACL)*, 2022.
- P29 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.
- P30 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)
- P31 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.
- P32 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.
- P33 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.
- P34 Xiyang Zhang, **Muhao Chen**, Jonathan May. Saliency-Aware Event Chain Modeling for Narrative Understanding. In *Proceedings of the 36th Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2021.
- P35 Fei Wang, Kexuan Sun, Jay Pujara, Pedro Szekely, **Muhao Chen**. Table-based Fact Verification With Saliency-aware Learning. In *the 36th Conference on Empirical Methods in Natural Language Processing (EMNLP)* - Findings, 2021.
- P36 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.
- P37 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.
- P38 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.
- P39 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.

- P40 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.
- P41 **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.
- P42 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.
- P43 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.
- P44 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.
- P45 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.
- P46 **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**
- P47 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.
- P48 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.
- P49 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.
- P50 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.
- P51 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.
- P52 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.
- P53 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**, ~0.8%)
- P54 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
- P55 **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
- P56 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.
- P57 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.
- P58 **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.
- P59 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.
- P60 **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.

- P61 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.
- P62 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.
- P63 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.
- P64 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.
- P65 **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.
- P66 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.
- P67 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.
- P68 **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)
- P69 **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.
- P70 **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
- P71 Haochen Chen, Xiaofei Sun, Yingtao Tian, Bryan Perozzi, **Muhao Chen** and Steven Skiena. Enhanced Network Embeddings via Exploiting Edge Labels. In *the 27<sup>th</sup> ACM Conference on Information and Knowledge Management (CIKM)*. ACM 2018.
- P72 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 19<sup>th</sup> ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, ACM 2018.
- P73 **Muhao Chen**, Qi Zhao, Pengyuan Du, Carlo Zaniolo, Mario Gerla. Demand-driven Cache Allocation Based on Context-aware Collaborative Filtering. In *Proceedings of the 19<sup>th</sup> ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, ACM 2018.
- P74 **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.
- P75 **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 Workshop and System Demonstration Papers

- W1 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.
- W2 Zubo 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
- W3 Changjun Fan, Yizhou Sun, Li Zeng, Yang-Yu Liu, **Muhao Chen**, Zhong Liu. Dismantle Large Networks through Deep Reinforcement Learning. In *ICLR Workshops*, 2019.
- W4 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.
- W5 **Muhao Chen**, Tao Zhou, Pei Zhou, Carlo Zaniolo. Multi-graph Affinity Embeddings for Multilingual Knowledge Graphs. **Contributed talk** in the *6<sup>th</sup> Workshop on Automated Knowledge Base Construction at NIPS (AKBC)*. 2017.
- W6 **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)

W7 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.

W8 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

- *Amazon Research Award*. 2023
- *Keston Exploratory Research Award*. 2023
- *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

Total funds acquired as PI or Co-PI: **\$1.38M** (my share)

<b>Amazon Research Award: <i>On Faithfulness of Information Extraction</i></b>	2023-2024
• \$73.8K unrestricted fund + \$20K AWS credits. (Sole PI)	
<b>Keston Exploratory Research Award: <i>Multi-document Newsworthy Event Monitoring and Forecasting</i></b>	2023-2024
• \$100K unrestricted fund from Keston Family Foundation. (PI; w/ Jonathan May)	
<b>Cisco Faculty Research Award: <i>Robust Knowledge Extraction from Text</i></b>	2022-2023
• \$70.4K unrestricted fund. (Sole PI)	
<b>DARPA KMASS: <i>Knowledge Needed in Context</i></b>	2022-2025
• \$866K my share out of \$6.64M in total. (Co-PI, Lead of TA-B)	
<b>NSF Cloud Access Grant</b>	2022-2023
• \$80K cloud computing credits. (Sole PI)	
<b><i>Robust Table-Text Understanding with Structure-Aware Equivariance Pre-training</i></b>	2022-2023
• Gift award from Amazon Alexa. PI of \$50K unrestricted fund.	
<b>Google GCP Grant</b>	2022-2023
• \$5K Google cloud credits.	
<b><i>Improved Performance, Analytics and Summarization of Synergistic Anticipation of Geopolitical Events</i></b>	2022-2023
• ARLIS INFER research grant. \$50K my share. (Co-PI)	
<b>NSF CRII: III: <i>Knowledge Graph Completion with Transferable Representation Learning</i></b>	2021-2023
• NSF IIS Grant, \$175K. (Sole PI)	
<b>DARPA MCS: <i>Multi-modal Open World Grounded Learning and Inference</i></b>	2021-2023
• \$400K my share. (Senior Personnel)	

## TEACHING AND MENTORING

### Teaching

Instructor at USC

- CSCI 544: Applied Natural Language Processing (Graduate-level course; Enrollment: 307) Spring 2022

Teaching assistant (2015.9-2016.6), associate (2016.9-2017.6), fellow (2017.9-2018.6) at UCLA

- Data Structures and Algorithms; Introduction to Computer Science.

Lab Members/Students at USC



Fei Wang, Ph.D. Student in Computer Science.	Fall 2022-date
• Viterbi Honors Program; <b>USC CS Departmental Best Research Award 2022; Annenberg Fellow.</b>	
• Joined the lab as a M.S. student in Fall 2020	
Bangzheng Li, Ph.D. Student in Computer Science.	Fall 2022-date
• <b>Provost PhD Fellow.</b>	
• Joined the lab as an undergraduate researcher in Summer 2021	
Qin (Jacqueline) Liu, Ph.D. Student in Computer Science. (USC PhD Fellowship)	Fall 2022-date
Tenghao Huang, Ph.D. Student in Computer Science (Viterbi-ISI Fellow; Co-advised with Xuezhe Ma)	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
Eric (Ehsan) Qasemi, PhD Student in Computer Science. (Co-advised with Pedro Szekely)	Fall 2020-date
Wenxuan Zhou, PhD Student in Computer Science.	Fall 2020-date
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-date
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

#### Visiting Scholars/Students at USC

Tianqing Fang. PhD Student in Computer Science, HKUST (Hong Kong PhD Fellow)	Summer 2022-Spring 2023
Zhaotian Weng. Undergraduate, Tsinghua University. (USC-THU Summer Research Program)	Summer 2022
Xiaocong Yang. Undergraduate, Tsinghua University.	Summer-Fall 2022
Amani Rune Maina-Kilaas. Undergraduate, Harvey Mudd College (NSF REU Student)	Summer 2022
Tanay Dixit. Undergraduate, IIT Madras, Computer Science (NSF REU Student)	Summer 2022
Devadutta Dash. Undergraduate, IIT Varanasi, Computer Science (IUSSTF-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; Now PhD student at UCSB)	Summer 2021
Piyush Khanna, Undergraduate, Delhi Tech Univ., Computer Science	Summer 2021
Yueguan Wang. Undergraduate, Tsinghua Univ., EE (USC-THU Summer Research Program)	Summer 2021

#### Dissertation/Qualification Committee

Jianzhi Yang, USC PhD Student in Quantitative and Computational Biology. (Chair: Mark Chaisson)
Tejas Srinivasan, USC PhD Student in Computer Science. (Chair: Jesse Thomason)
Binh Vu, USC PhD Student in Computer Science. (Chair: Craig Knoblock)
Yizhou Zhang, USC PhD Student in Computer Science. (Chair: Yan Liu)
Jihoon Sohn, USC PhD in Mathematics, 2022. (Chair: Francis Bonahon)
Tianran Zhang, UCLA PhD in Bioengineering, 2022. (Chair: Alex Bui)
Peifeng Wang, USC PhD Student in Computer Science. (Chair: Pedro Szekely)
Michiel de Jong, USC PhD Student in Computer Science. (Chair: Leana Golubchik)
Minh Pham, USC PhD in Computer Science, 2022. (Chair: Craig Knoblock)
Kexuan Sun, USC PhD Student in Computer Science. (Chair: Jay Pujara)
Xuelu (Shirley) Chen, UCLA PhD in Computer Science, 2021. (Chair: Carlo Zaniolo)
Jeong Hyun An, USC M.S. in Computer Science, 2022. (Committee Chair)

#### Mentoring (Before USC)

Haoyu Wang, UPenn MS Student in CIS (Now PhD student at UPenn). <b>Research mentor</b>	Fall 2019-Spring 2022
Tianran Zhang, UCLA PhD student in Bioengineering. <b>Research mentor</b>	Fall 2019-Winter 2021
Junheng Hao, UCLA PhD student in Computer Science. <b>Research mentor</b>	Fall 2017-Fall 2019
Xuelu (Shirley) Chen, UCLA PhD student in Computer Science. <b>Research mentor</b>	Fall 2017-Winter 2020
Pei Zhou, UCLA undergraduate student (Now PhD student at USC CS). <b>Research mentor</b>	Winter 2017-Spring 2019
Weijia Shi, UCLA undergraduate student (Now PhD student at UW CSE). <b>Research mentor</b>	Summer 2018-Spring 2019

Ankith Uppunda, UCLA undergraduate student. **Research mentor.**

Spring 2019

Zhubo Deng, UCLA undergraduate student. **Research mentor.**

Spring 2019

Xiaoshuang Wei, UCLA MS student in Computer Science (Now software engineer at Google). **Thesis mentor**

Spring 2017

## PROFESSIONAL SERVICE

### Panelist:

2021, 2022: NSF CISE Core Panel

### Conference Organization/Senior Committee Member:

2023: AAAI (**Area Chair**), AACL (**Senior Area Chair** – Information Retrieval and Text Mining), ACL (**Area Chair** – Information Extraction), NLPCC (Area Chair)

2022: AAAI (Senior PC), NAACL (**Senior Area Chair** – Information Extraction), NAACL SRW (Faculty Mentor), EMNLP (**Area Chair** – Commonsense Reasoning), AACL (**Area Chair** – Information Retrieval and Text Mining)

2021: AAAI (Senior PC), IJCAI (Senior PC)

2019: IEEE AIKE (Doctoral Consortium Chair)

### Workshop Organization Member:

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, AACL-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.

### Editorial Board: Frontiers in Big Data

**Journal Reviewer:** TPAMI, TACL, 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.

### University Service

NSF REU Site, **Co-organizer**

2022

USC ISI Institutional AI Seminar, **Organizer**

2021.7-date

Department of Computer Science, **Faculty Hiring Committee**

2021, 2022

Viterbi School of Engineering, **Fellowship Committee**

2021

Department of Computer Science, **PhD Admission Fellowship Committee**

2021

## PRESENTATIONS

### *Invited talks, colloquia and tutorials*

1. Indirectly Supervised Natural Language Processing. *Half-day Tutorial at ACL*. July, 2023.
2. Robust and Indirectly Supervised Knowledge Acquisition. *CS Research Colloquium, ASU*. March 2023
3. Robust and Indirectly Supervised Knowledge Acquisition. *CS Research Colloquium, UMN*. March 2023
4. Robust and Indirectly Supervised Knowledge Acquisition. *CS Research Colloquium, UC Davis*. Feb 2023
5. Robust and Indirectly Supervised Knowledge Acquisition. *CS Research Colloquium, UCSB*. Feb 2023
6. Robust and Indirectly Supervised Information Extraction. *Invited talk, AI & ML Seminar, UCI*. Nov 2022
7. Robust and Indirectly Supervised Information Extraction. *Invited talk, NL Seminar, UMN*. Nov 2022
8. Robust and Indirectly Supervised Information Extraction. *Invited talk, CS Department Seminar (CS201), UCLA*. Oct 2022.
9. Robust and Indirectly Supervised Information Extraction. *Invited talk, CS Department Seminar (remote), Nanjing University*. Sept 2022.



10. Robust and Indirectly Supervised Information Extraction. *Invited talk, Apple, San Jose*. Aug 2022.
11. Robust and Indirectly Supervised Information Extraction. *Invited talk, Microsoft CSR Distinguished Talk Series*. July 2022.
12. New Frontiers of Information Extraction. *Half-day Tutorial at NAACL*. July, 2022.
13. Understanding Event Processes in Natural Language. *Invited talk, NLP Seminar, UC Santa Cruz*. April 2022.
14. Understanding Event Processes in Natural Language. *Invited talk, CS Department Seminar, Rutgers–New Brunswick*. March 2022.
15. Understanding Event Processes in Natural Language. *Invited talk, ML Seminar, Purdue University*. Sept 2021.
16. Understanding Event Processes in Natural Language. *Invited talk, Tencent AI Lab, Seattle*. Sept 2021.
17. Understanding Event Processes in Natural Language. *Invited talk, University of Central Florida*. Aug 2021.
18. Understanding Event Processes in Natural Language. *Invited talk, IBM Research Almaden*. July 2021.
19. Understanding Event Processes in Natural Language. *Invited talk, CS Department Seminar, Fudan University*. June 2021.
20. Understanding Event Processes in Natural Language. *Invited talk, CS Department Seminar, Nanjing University*. June 2021.
21. Understanding Event Processes in Natural Language. *Invited talk, NLP Seminar, National University of Singapore*. May 2021.
22. Understanding Event Processes in Natural Language. *Invited talk, NLP Seminar, OSU*. May 2021
23. Understanding Event Processes in Natural Language. *Invited talk, CS Research Colloquium, UCSB*. April 2021
24. Understanding Event Processes in Natural Language. *Invited talk, Frontier Topics in Vision and Language, ASU*. Mar 2021.
25. Understanding Event Processes in Natural Language. *Invited talk at Machine Learning and Big Data Seminar, UCLA*. Nov 2020
26. Event-Centric Natural Language Processing. *Half-day Tutorial at ACL*. Aug, 2021.
27. From Tables to Knowledge: Recent Advances in Table Understanding. *Half-day Tutorial at KDD*. Aug, 2021.
28. Understanding Event Processes in Natural Language. Understanding Event Processes in Natural Language. *Invited talk at Language Technology Seminar, University of Cambridge*. Nov 2020
29. Event-Centric Natural Language Understanding. *Half-day Tutorial at AAAI*. Feb, 2021.
30. Knowledge Acquisition with Transferable Representation Learning. *CS Research Colloquium, USC*. Nov 2020
31. Knowledge Acquisition with Transferable Representation Learning. *AI Seminar, USC ISI*. Los Angeles, Jan 2020.
32. Knowledge Acquisition with Transferable Representation Learning. *IBM Research Almaden*. San Jose, Jan 2020.
33. Knowledge Acquisition with Transferable Representation Learning. *Invited talk at FDSiF, Fudan Univ., Shanghai, China*. Dec 2019.
34. Knowledge Acquisition with Transferable Representation Learning. *Invited talk at SNAP Seminar, Stanford University*. Jan 2019.
35. Knowledge Acquisition with Transferable Representation Learning. *Machine Learning and Big Data Seminar, UCLA*. Jan 2019.
36. Recent Advances in Transferable Representation Learning. *Half-day tutorial at AAAI*. NYC, NY, Feb 2020.
37. Neural Article Pair Modeling for Wikipedia Sub-article Matching. *Google Search Intelligence Seminar Talk*. Sept 2017
38. 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 |
| <ul style="list-style-type: none"><li>• <b>Research Intern</b>, <i>Embedding Edge-attributed Relational Hierarchies</i>. [C35]</li></ul>                                     |               |
| Google, Mountain View, CA (Google Knowledge Graph)   | 2017.6~2017.9 |
| <ul style="list-style-type: none"><li>• <b>Research Intern</b>, <i>Neural Article Pair Modeling for Large-scale Sub-article Relation Extraction</i>. [C43,C40,C27]</li></ul> |               |
| Google, Mountain View, CA (Procella Real-time Data Infrastructure)   | 2016.6~2016.9 |
| <ul style="list-style-type: none"><li>• <b>System SDE Intern</b>, <i>Dynamic Shard-partitioning for Large-scale Windowed Event Streams</i>.</li></ul>                        |               |
| Teradata Labs, Los Angeles, CA (Optimizer Group)   | 2015.6~2015.9 |
| <ul style="list-style-type: none"><li>• <b>R&amp;D Intern</b>, <i>Multi-task Learning for Cost-based Database Optimizers</i>.</li></ul>                                      |               |

