Xiang Ren

Associate Professor, Computer Science Department, University of Southern California
1031 Downey Way GCS 402K, Los Angeles, CA 90089 USA
(213) 821-4067 | xiangren@usc.edu | https://seanre.com/

Employment History

University of Southern California — *Los Angeles, CA*
- Andrew and Erna Viterbi Early Career Chair — *10/ 2022 - present*
- Associate Professor (with tenure), Computer Science Department — *04/ 2023 - present*
- Assistant Professor, Computer Science Department — *02/ 2018 - 03/2023*
- Research Team Leader, Information Science Institutes (by courtesy) — *05/ 2020 - present*
- Director, USC Intelligence and Knowledge Discovery Research Lab — *02/ 2018 - present*

Sahara AI — *Los Angeles, CA*
- Co-founder — *03/ 2023 - present*

Allen Institute for Artificial Intelligence (AI2) — *Los Angeles, CA*
- Visiting Research Scientist (20%) — *04/ 2022 - 04/2023*

ACM SIGKDD — *Los Angeles, CA*
- Information Director, Appointed Chair — *01/ 2019 - present*

Snap Inc. — *Santa Monica, CA*
- Data Science Advisor — *03/ 2018 - 03/ 2019*

Stanford University — *Stanford, CA*
- Research Scholar, NLP Group & SNAP Group — *05/ 2017 - 02/ 2018*

Research Interests

Natural language processing
- LLM reasoning and generalization
- Model safety and trust
- Attribution, memorization, and provenance

Machine learning
- Efficient inference time method and scaling
- Continual learning & rapid generalization
- Interpretability; Post-hoc model explanation

Education History

Ph.D. in Computer Science

University of Illinois Urbana-Champaign, IL, 08/ 2012 - 02/ 2018

Advisor: Jiawei Han; Committee members: Chengxiang Zhai, Heng Ji, Tarek Abdelzaher
2018 ACM SIGKDD Doctoral Dissertation Award, C. W. Gear Outstanding Graduate Student Award

M.Sc. in Computer Science

University of Illinois Urbana-Champaign, IL, 04/ 2016

2016 David J. Kuck Outstanding M.S. Thesis Award (given to 1 student from the computer science department)

B.Eng. in Computer Science

Zhejiang University, Hangzhou, China, 05/ 2008 - 06/ 2012

Excellence Senior Thesis Award, Computer Science Department, Graduate with honors, Chu Kochen Honors College

Selected Honors and Awards

2024 **ACM WSDM Test of Time Award**
2023 **ACL Outstanding Paper Award**
2023 **Samsung AI Researcher of the Year**
2023 **MIT Tech Review Innovators Under 35 (Asia Pacific)**
2022 **NAACL Outstanding Paper Award**
2022 Andrew and Erna Viterbi Early Career Chair, USC
2022 Amazon Alexa Fairness in AI Award

2021 **NSF CAREER Award**
2021 Facebook Faculty Research Award
2021 Best Paper Award, NAACL TrustNLP Workshop
2020 **Best Paper Award Runner-up, The Web Conference (WWW)**
2020 Sony Faculty Research Award
2020 Snapchat Faculty Research Award
2019 **Forbes' Asia 30 Under 30, Healthcare & Science**
2019 JP Morgan AI Research Award
2019 Amazon Faculty Research Award
2019 Adobe Data Science Research Award
2018 Google Faculty Research Award
2018 ***Winner*, ACM SIGKDD Doctoral Dissertation Award**

Top Five Cited Papers (Total Citation on Google Scholar: 34,065)

[1] Hierarchical Graph Representation Learning with Differentiable Pooling (**Citations: 2,680**)
Rex Ying, Jiaxuan You, Chris Morris, **Xiang Ren**, William L. Hamilton, Jure Leskovec
Annual Conference on Neural Information Processing Systems (**NIPS**), 2018

[2] GraphRNN: A Deep Generative Model for Graphs (**Citations: 1,568**)
Jiaxuan You, Rex Ying, **Xiang Ren**, William L. Hamilton, Jure Leskovec
International Conference on Machine Learning (**ICML**), 2018

[3] Personalized Entity Recommendation: A Heterogeneous Information Network Approach (**Citations: 1,066**)
X. Yu, **Xiang Ren**, Y. Sun, Q. Gu, B. Sturt, U. Khandelwal, B. Norick, J. Han
7th ACM International Conference on Web Search and Data Mining (**WSDM**), 2014. (**Test of Time Award**)

[4] Faith and Fate: Limits of Transformers on Compositionality (**Citations: 1,051**)
Nouha Dziri, Ximing Lu, et al., **Xiang Ren**, Allyson Ettinger, Zaid Harchaoui, Yejin Choi
Advances in Neural Information Processing Systems (**NeurIPS**), 2023

[5] LLM-Blender: Ensembling Large Language Models with Pairwise Ranking and Generative Fusion (**Citations: 931**)
Dongfu Jiang, **Xiang Ren**, Bill Yuchen Lin
Annual Meeting of the Association for Computational Linguistics (**ACL**), 2023

Research Grants, Sponsorship, Gifts

Total funding (own share) ≈ $5.8 million

[11] **NSF Proto-OKN**: *Knowledge Graph Construction for Resilient, Trustworthy, and Secure Software Supply Chain*.
- **PI**, (10-01-2023 - 09-30-2026), $350,505 (my share)
- **Role:** Build a natural language interface to the software knowledge base system.

[10] **NSF CAREER**: *Teaching Machine through Human Explanations for Information Extraction*.
- **sole PI**, award# 2048211 (05-15-2021 - 04-30-2026), $513,058 (total)
- **Role:** Build and maintain domain-specific information extraction systems with human explanations.

[9] **NSF SciSIP**: *Modeling the Invention, Dissemination, and Translation of Scientific Concepts*.
- **PI**, award# 1829268 (08-01-2018 – 08-31-2020), $239,988 (own share); $503,974 (total). **Co-PI**: Daniel MacFarland
- **Role:** Develop neural models for measuring and modeling processes of scientific innovation.

[8] **DARPA Machine Commonsense (MCS)**: *Multi-modal Open World Grounded Learning and Inference (MOWGLI)*.
- **Co-PI** (06-01-2019 – 05-31-2023), $1,026,796 (own share); $8.6M (total).
- **PI**: Pedro Szekely. **Other Co-PIs**: Jay Pujara, Jure Leskovec, Sameer Singh, Deborah McGuinness
- **Role:** Address the challenge of machine common sense by reading the web and modeling structured knowledge.

[7] **DARPA Knowledge Management at Scale and Speed (KMASS)**: *LKNIC: Knowledge Needed in Context*.
- **Co-PI** (03-01-2022 – 02-31-2024), $526,613 (own share); $4.5M (total).
- **PI**: Jay Pujara. **Other Co-PIs**: Jonathan May, David Truam, Elizabeth Boschee
- **Role:** Revolutionize knowledge management through the use of a cognitively-inspired organization of multimodal background knowledge, a dialogue-driven agent for knowledge capture and dissemination.

[6] **DARPA Grounded Artificial Intelligence Language Acquisition (GAILA)**: *Learning Language from Image and Video Entities*.
- **Co-PI** (07-01-2019 – 5-31-2022), $372,747 (own share); $925k (total). **PI**: Ram Nevatia; **Other Co-PI**: Yan Liu.
- **Role:** Develop continual learning algorithms for vision-language models.

[5] **Schmidt Family Foundation Grant**: *Using Artificial Intelligence to Support Cognitive Task Analysis*.
- **Co-investigator** (06-15-2019 – 09-30-2020), $173k (own share); $340k (total). **PI**: Kenneth Yates; **Co-PI**: Rao Machiraju
- **Role:** Automatic construction of procedure knowledge graphs from interview transcripts.

[4] **IARPA Human Interpretable Attribution of Text Using Underlying Structure (HIATUS)**: *SADIRI: Stylometric Authorship Discernment and Interpretation for Realistic Inputs*.
- **Co-investigator** (09-10-2022 – 06-15-2025), $411,156 (own share)
- **PI**: Elizabeth Boschee. **Other Co-PIs**: Scott Miller
- **Role:** Develop enhanced methods for explaining language models for authorship detection.

[3] **IARPA Better Extraction from Text Towards Enhanced Retrieval (BETTER)**: *Cross-Lingual Event & Argument Retrieval*.
- **Co-investigator** (06-01-2019 – 11-30-2022), $479,682 (own share); $7.28M (total)
- **PI**: Elizabeth Boschee. **Other Co-PIs**: Scott Miller, James Allan
- **Role:** Develop enhanced methods for personalized, multilingual semantic extraction and retrieval from text.

[2] **DARPA Influence Campaign Awareness and Sensemaking (INCAS)**
- **Co-investigator**, (8-15-2021 – 12-31-2025), $464,803 (own share); $4.2M (total)
- **PI**: Emilio Ferrara. **Other Co-PIs**: Alessandro Flammini, Jacob Shapiro, Jonathan May
- **Role:** Develop explanation-based learning methods for campaign prediction models.

[1] **DARPA Systematizing Confidence in Open Research and Evidence (SCORE)**: *Macroscopic models for Reproducibility and Replicability of Research Claims in the Social and Behavioral Sciences*.
- **Co-PI**, (10-01-2019 – 12-31-2021), $274,104 (own share); $1.83M (total). **PI**: Jay Pujara. **Other Co-PI**: Kristina Lerman
- **Role:** Develop human-in-the-loop learning interfaces for information extraction from scientific papers.

[0] **Gift awards from Industry (sole PI, $830k+)**: Google Research Scholar Award ($60k), Facebook Sponsored Research Award ($150k), Amazon Trust ML grant ($70k), JP Morgan AI research award ($150k), Sony faculty innovation award ($100k), Google faculty research award ($80k), Amazon research award ($160k), Adobe data science award ($80k), Snapchat gifts ($20k)

Selected Media Coverage

2025 **NASDAQ** TradeTalk: "*Proactive AI Policy Balancing National Security, Privacy, and Economics*"
2025 **PYMNTS** "*Decentralization Emerges as a Test of Big Tech's AI Power*"
2024 **Reuters** "*Decentralized AI network Sahara raises fresh capital*"
2024 **Bloomberg** "*USC Professor Lands $43 Million in Funding for AI Startup*"
2023 **USC Viterbi** "*Dungeons & Dragons: Leading the Conversation through Theory-of-Mind*"
2021 **Communications of the ACM** "*The Best of NLP*"
2020 **The Register** While text-generating AI can write like humans, it lacks common sense
2020 **TechXplore** New test reveals AI still lacks common sense
2020 **Radio.com** AI still lacks common sense according to new research

Service

Community: K-12 outreach

Project pitch panel, Linked Learning Engineering Pathway at STEM Academy Hollywood, Los Angeles, 03/ 2021
Project advising, Linked Learning Engineering Pathway at Theodore Roosevelt High School, Los Angeles, 03/ 2021
Guest lecture, Theodore Roosevelt High School, Los Angeles, 02/ 2021

Professional: Organizer / Chair on External Activities

ACM SIGKDD Information Director, Sep 2018 - present
Commonsense Representation and Reasoning (**CRR**), Workshop Co-chair, co-located with ACL 2022
Federated Learning for NLP (**FedNLP**), Workshop Co-chair, co-located with ACL 2022
Deep Learning for Low-resource NLP (**DeepLo**), Workshop Co-chair, co-located with NAACL 2022
Trustworthy NLP (**TrustNLP**) Workshop Co-chair, co-located with NAACL 2021
The 2nd Learning from Limited Labeled Data Workshop (**LLD**), Workshop Co-chair, co-located with ICLR 2019
The 4th Workshop on Representation Learning for NLP (**Rep4NLP**), Workshop Co-chair, co-located with ACL 2019
First Conference on Automated Knowledge Base Construction (**AKBC**), Workshop Co-chair, 2018

Professional: Area Chair / Senior PC for International Conferences

Annual Meeting of the Association for Computational Linguistics (**ACL**), Area Chair, 2019-2021, 2023-2024
Conference on Empirical Methods in Natural Language Processing (**EMNLP**), Area Chair, 2018-2022
Annual Conference of the North American Chapter of the ACL (**NAACL**), Area Chair, 2021
The International Conference on Learning Representations (**ICLR**), Area Chair, 2020-2022
The Thirty-Fourth AAAI Conference on Artificial Intelligence (**AAAI**), Senior Area Chair, 2020
International Conference on Machine Learning (**ICML**), Area Chair, 2022

Professional: Program Committee for International Conferences

ACM SIGKDD Conference on Knowledge Discovery and Data Mining (**KDD**): 2015-21

Conference on Neural Information Processing Systems (**NeurIPS**): 2019-21
International Conference on Machine Learning (**ICML**): 2019-20
Annual Meeting of the Association for Computational Linguistics (**ACL**): 2017-19
Conference on Empirical Methods in Natural Language Processing (**EMNLP**): 2015-18
International World Wide Web Conference (**WWW**): 2017-20
The AAAI Conference on Artificial Intelligence (**AAAI**): 2018-19
International Joint Conference on Artificial Intelligence (**IJCAI**): 2018-19
ACM International Conference on Web Search and Data Mining (**WSDM**): 2017-19
Annual Conference of the North American Chapter of the Association for Computational Linguistics (**NAACL**): 2018-19
ACM SIGIR Conference on Research and Development in Information Retrieval (**SIGIR**): 2017-18
ACM Conference on Information and Knowledge Management (**CIKM**): 2017

Professional: Journal Reviewer
Transactions of the Association for Computational Linguistics (**TACL**)
IEEEs Transactions on Knowledge and Data Engineering (**TKDE**)
ACM Transactions on Knowledge Discovery from Data (**TKDD**)

Teaching

Classes taught at USC

Spring26	**Instructor,**	CSCI 662: Advanced Natural Language Processing	*USC*
Spring26	**Instructor,**	CSCI 444: Natural Language Processing	*USC*
Spring22	**Instructor,**	CSCI 566: Deep Learning and its Applications	*USC*
Fall21	**Instructor,**	CSCI 699: Representation Learning for NLP	*USC*
Fall120	**Instructor,**	CSCI 699: Representation Learning for NLP	*USC*
Fall2019	**Instructor,**	CSCI 699: Advanced Topics in Representation Learning for NLP	*USC*
Spring19	**Instructor,**	CSCI 699: Machine Learning for Knowledge Extraction and Reasoning	*USC*
Fall18	**Instructor,**	CSCI 699: Introduction to Information Extraction.	*USC*

Conference tutorials

[10] Learning from Task Instructions
The Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2023.

[9] Knowledge-Augmented Methods for Natural Language Processing
The Annual Conference of the Association for Computational Linguistics (**ACL**), 2022.

[8] Scalable Construction and Reasoning of Massive Knowledge Bases
The Annual Conference of the North American Chapter of the Association for Computational Linguistics (**NAACL**), 2018.

[7] Scalable Construction and Querying of Massive Knowledge Bases
27th International World Wide Web Conference (**WWW**), 2018.

[6] Construction and Querying of Large-scale Knowledge Bases
ACM Conference on Information and Knowledge Management (**CIKM**), 2017.

[5] Mining Entity-Relation-Attribute Structures from Massive Text Corpora
23th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (**KDD**), 2017.

[4] Building Structured Databases of Factual Knowledge from Massive Text Data
2017 International Conference on Management of Data (**SIGMOD**), 2017.

[3] Constructing Structured Information Networks from Massive Text Corpora
26th International World Wide Web Conference (**WWW**), 2017.

[2] Automatic Entity Recognition and Typing in Massive Text Data
2016 International Conference on Management of Data (**SIGMOD**), 2016.

[1] Automatic Entity Recognition and Typing in Massive Text Corpora
25th International World Wide Web Conference (**WWW**), 2016.

Invited Talks

[29] The API Privacy Paradox: How LLM APIs Unintentionally Reveal Model Details. Seoul National University; Berkeley RDI; Google; Stanford, 2025.

[28] AI in Web3, Pantera Summit, 2025.

[27] The Convergence of AI and Web3, UC Berkeley RDI Summit, 2025.

[26] Reflex or Reflect: When Do Language Tasks Need Slow Reasoning?, UCSD AI Seminar (SoCal), 2023.

[26] UMass NLP seminar, Apr 2023.

[25] UPenn NLP seminar, Dec 2022.

[24] HiTZ Basque Center for Language Technology, Sep 2022.

[23] Allen Institute for Artificial Intelligence, Mar 2022.

[22] Intel, Sep 2021.

[21] Amazon Alexa NU Summit, April 2021.

[20] UIUC NLP Seminar, Mar 2021.

[19] Google X, Sep 2020.

[18] CMU LTI Colloquium, Feb 2020.

[17] UT Austin, Jan 2019.

[16] Bloomberg, May 2019.

[15] JP Morgan, May 2019.

[14] IBM Research, Almaden, May 2019.

[13] Stanford NLP Seminar, May 2019.

[12] Stanford University, Aug 2017.

[11] Peking University, Computer Science Department, Jun 2017.

[10] Microsoft Research Asia, Jun 2017.

[9] Tsinghua University, Computer Science Department, Jun 2017.

[8] IBM Research Almaden, Jun 2017.

[7] Duke University, Computer Science Department, March 2017.

[6] University of California, Los Angeles, Computer Science Department, March 2017.

[5] UC Davis, Computer Science Department, March 2017.

[4] Georgia Tech, School of Computer Science, March 2017.

[3] University of Southern California, Computer Science, Feb. 2017.

[2] Google Research, August 2016.

[1] Microsoft Research, Redmond, August 2015.

Mentoring

Graduated PhDs advised at USC: Aaron Chan (CS, 2017-2022, sole advisor; **Graduation date:** December of 2022; **Placement**: Full-time research scientist at Meta AI), Yuchen Lin (CS, 2018-2022, sole advisor; **Graduation date:** December of 2022; **Placement**: Postdoc scholar at University of Washington Seattle and Allen Institute of Artificial Intelligence), Shushan Arakelyan (CS, 2017-present, sole advisor), Woojeong Jin (CS, 2018-2024, sole advisor; **Graduation date:** December of 2024; **Placement**: Applied Scientist at Hippocratic AI), Qinyuan Ye (CS, 2019-2025, sole advisor; **Graduation date:** December of 2025; **Placement**: Applied Scientist at Salesforce Research), Pei Zou (CS, 2019-2024, co-advised w/ Jay Pujara, primary advisor; **Graduation date:** December of 2024; **Placement**: Senior Applied Scientist at Microsoft), Xisen Jin (CS, 2019-2025, sole advisor; **Graduation date:** December of 2025; **Placement**: Research Scientist at Sahara AI), Jun Yan (CS, 2019-2024, sole advisor; **Graduation date:** December of 2024; **Placement**: Research Scientist at Google AI), Peifeng Wang (CS, 2019-2023, co-advised w/ Pedro Szekely, secondary co-advisor; **Graduation date:** December of 2023; **Placement**: Research Scientist at Salesforce Research), Soumya Sanyal (CS, 2020-2025, sole advisor; **Graduation date:** December of 2025; **Placement**: Applied Scientist at Amazon AWS)

Current PhD students advised at USC: Brihi Joshi (CS, Fall 2021 – present, co-advised w/ Swabha Swayamdipta; primary advisor Xiang Ren), Jacob Bremerman (CS, Fall 2021 – present, co-advised w/ Jonathan May; primary advisor Xiang Ren), Huihan Li (CS, Fall 2022 – present, sole advisor Xiang Ren), Sahana Ramnath (CS, Fall 2022 – present, sole advisor Xiang Ren), Sayan Ghosh (CS, Fall 2022 – present, sole advisor Xiang Ren), Hirona Arai (CS, Fall 2022 – present, sole advisor Xiang Ren), Matthew Finlayson (CS, Fall 2023 – present, co-advised w/ Swabha Swayamdipta; primary advisor Xiang Ren), Skyler Hallinan (CS, Fall 2024 – present, co-advised w/ Sai Praneeth Karimireddy; primary advisor Xiang Ren), Michael Duan (CS, Fall 2025 – present, sole advisor Xiang Ren), Sheryl Mathew (CS, Fall 2025 – present, co-advised w/ Julian J. Traum; primary advisor Xiang Ren)

PhD thesis committee member at USC: Sahil Garg (CS, 2014), Yaguang Li (CS, 2014), Fenxiao Chen (EE, 2014), Fenxiao Chen (EE, 2014), Akira Matsui (CS, 2017), Yilei Zeng (CS, 2017), Alana Shine (CS, 2017), Yu-Chuan Yen (CS, 2017); Tae Jin Park (ECE, 2015); Nazgol Tavabi (CS, 2015), Minh Pham (CS 2018), Rujun Han (CS 2018), Yury Zemlyanskiy (CS 2018), Akira Matsui (CS 2017), Aida Davani (CS, 2017), Brendan Kennedy (CS, 2016), Chaoyang He (CS, 2018), Thamme Gowda (CS, 2017), Dong-Ho Lee (CS, 2025),

Master students advised at USC (sole advisor): Hongtao Lin (CS, 2017-2018), Yozen Liu (CS, 2017-2018), Junyi Du (CS, 2017-2019), Ouyu Lan (CS, 2018-2019), Dongho Lee (CS, 2018-2020), Seyeon Lee (CS, 2019-present), Suji Kim (CS, 2019-present), Rahul Khanna (CS, 2019-present), Ravi Kiran Selvam, (CS, 2020-2021), Mahak Agarwal (CS, 2020-2021), Akshen Kadakia (CS, 2021-2022), Juan Zha (CS, 2021-present), Ziyi Liu (CS, 2021-present), Junfeng Liu (CS, 2021-present)

Undergraduate students advised at USC: Hao Chen (CS, 2018-2019), Yu Cao (CS progressive degree, 2018-2019, *USC Undergraudate Research Award*), Jamin Chen (CS, 2019-2020), Priya Irukulapati (CS, 2019-2020), Kelly Dickson (CS, 2019-2020), Ryan Moreno (CS, 2019-2020, *USC Undergraudate Research Award*, *Best Undergraduate Thesis Award*), Boyuan Long (CS, 2020-2021), Raymond Cheung (CS, 2020-2021), Wardah Jabeen (CS, 2020-2021), Adeline Liou (CS, 2020-2021), Jiashu Xu (CS, 2020-2021), Jordan McWilliams (CS, 2020-2021), Lindsey Feng (CS, 2021-2022), Kangmin Tan (CS, 2021-2022)

Publications

Book

Mining Structures of Factual Knowledge from Text: An Effort-Light Approach
Xiang Ren and Jiawei Han
Morgan & Claypool Publishers, 2018

Preprints

[P11] ARBOR: Locating Reasoning Failures in Long Chains of Thought through Reasoning State Tracking
Huihan Li, Siyuan Wang, Naga Vamsi Ramana Dinavahi, Xiaohui Liu, **Xiang Ren**
arXiv preprint, 2026

[P10] Language Model Activations Inhabit Privileged Error-Correcting Basins
Matthew Finlayson, Francisco Pernice, Eric Todd, Amir Zur, Daniel Wurgaft, Fenil R. Doshi, Vasudev Shyam, Matt Feiszli, Satchel Grant, Lucius Bushnaq, Tal Haklay, Usha Bhalla, Matthew Kowal, Thomas Fel, Jack Merullo, Atticus Geiger, **Xiang Ren**, Owen Lewis, Ekdeep Singh Lubana
arXiv preprint, 2026

[P9] ReLiveGym: Evaluating Long-Lived Agents over Weeks of Replayed Reality
Xisen Jin, Jingheng Li, Zhenglun Chen, Junyi Du, **Xiang Ren**
arXiv preprint, 2026

[P8] WhiteMatter: All-to-All Cross-Layer Connections via KV Mixing
Wenbo Zhang, **Xiang Ren**
arXiv preprint, 2026

[P7] Token Rankings are Unforgeable Language Model Signatures
Matthew Finlayson, Andreas Grivas, **Xiang Ren**, Swabha Swayamdipta
arXiv preprint, 2026

[P6] LATTICE: Evaluating Decision Support Utility of Crypto Agents
Aaron Chan, Tengfei Li, Tianyi Xiao, Angela Chen, Junyi Du, **Xiang Ren**
arXiv preprint, 2026

[P5] DIAL-SUMMER: A Structured Evaluation Framework of Hierarchical Errors in Dialogue Summaries
Sahana Ramnath, Nima Chitsazan, Mingyang Zhou, Chia-Hsuan Lee, Shi-Xiong Zhang, Stephen Rawls, Sambit Sahu, Sangwoo Cho, **Xiang Ren**, Genta Indra Winata, Akshaj Kumar Veldanda
arXiv preprint, 2026

[P4] OSGym: Scalable OS Infra for Computer Use Agents
Zengyi Qin, Jinyuan Chen, Yunze Man, Shengcao Cao, Ziqi Pang, Zhuoyuan Wang, Han Fang, Ling Zhu, Zixin Xie, Zibu Wei, Tianshu Ran, Haoran Geng, Ray Pan, Qizhen Sun, Zachary Bright, Yuyang Cai, Chongye Yang, Jiace Zhao, Tianrui Liu, Han Cao, Yeyang Zhou, Rui Wang, Song Wang, **Xiang Ren**, Bo Zhang, et al.
arXiv preprint, 2025

[P3] REALTALK: A 21-Day Real-World Dataset for Long-Term Conversation
Dong-Ho Lee, Adyasha Maharana, Jay Pujara, **Xiang Ren**, Francesco Barbieri
arXiv preprint, 2025

[P2] How FaR Are Large Language Models From Agents with Theory-of-Mind?
Pei Zhou, Aman Madaan, Srividya Pranavi Potharaju, Aditya Gupta, Kevin R. McKee, Ari Holtzman, Jay Pujara, **Xiang Ren**, Swaroop Mishra, Aida Nematzadeh, Shyam Upadhyay, Manaal Faruqui
arXiv preprint, 2023

[P1] DOMINO: A Dual-System for Multi-step Visual Language Reasoning
Peifang Wang, Olga Golovneva, Armen Aghajanyan, **Xiang Ren**, Muhao Chen, Asli Celikyilmaz, Maryam Fazel-Zarandi
arXiv preprint, 2023

Refereed conference articles

[195] Translating the Untranslatable: An Operationalizable Ontology for Untranslatability
Jacob Bremerman, Brihi Joshi, Hirona Arai, **Xiang Ren**, Jonathan May
Findings of the Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2026

[194] DRInQ: Evaluating Conversational Implicature with Controlled Context Variation
Hirona Jacqueline Arai, **Xiang Ren**
Annual Meeting of the Association for Computational Linguistics (**ACL**), 2026

[193] Believing without Seeing: Quality Scores for Contextualizing Vision-Language Model Explanations
Keyu He, Tejas Srinivasan, Brihi Joshi, **Xiang Ren**, Jesse Thomason, Swabha Swayamdipta
Annual Meeting of the Association for Computational Linguistics (**ACL**), 2026

[192] Proof-of-Guardrail in AI Agents and What (Not) to Trust from It
Xisen Jin, Michael Duan, Qin Lin, Aaron Chan, Zhenglun Chen, Junyi Du, **Xiang Ren**
ICML 2026 AI4GOOD Workshop, 2026

[191] OpaqueToolsBench: Learning Nuances of Tool Behavior Through Interaction
Skyler Hallinan, Thejas Venkatesh, **Xiang Ren**, Sai Praneeth Karimireddy, Ashwin Paranjape, Yuhao Zhang, Jack Hessel

ACM Conference on AI and Agentic Systems (**AIAS**), 2026

[190] Every Language Model Has a Forgery-Resistant Signature
Matthew Finlayson, **Xiang Ren**, Swabha Swayamdipta
International Conference on Representation Learning (**ICLR**), 2026

[189] Lookahead Tree-Based Rollouts for Enhanced Trajectory-Level Exploration in Reinforcement Learning with Verifiable Rewards
Shangyu Xing, Siyuan Wang, Chenyuan Yang, Xinyu Dai, **Xiang Ren**
International Conference on Representation Learning (**ICLR**), 2026

[188] Segment-Level Attribution for Selective Learning of Long Reasoning Traces
Siyuan Wang, Yanchen Liu, **Xiang Ren**
International Conference on Representation Learning (**ICLR**), 2026

[187] Function Induction and Task Generalization: An Interpretability Study with Off-by-One Addition
Qinyuan Ye, Robin Jia, **Xiang Ren**
International Conference on Representation Learning (**ICLR**), 2026

[186] Diverging Preferences: When do Annotators Disagree and do Models Know?
Michael J. Q. Zhang, Zhilin Wang, Jena D. Hwang, Yi Dong, Olivier Delalleau, Yejin Choi, Eunsol Choi, **Xiang Ren**, Valentina Pyatkin
International Conference on Machine Learning (**ICML**), 2025

[185] Attributing Culture-Conditioned Generations to Pretraining Corpora
Huihan Li, Arnav Goel, Keyu He, Xiang Ren
International Conference on Representation Learning (**ICLR**), 2025

[184] ELI-Why: Evaluating the Pedagogical Utility of Language Model Explanations
Brihi Joshi, Keyu He, Sahana Ramnath, Sadra Sabouri, Kaitlyn Zhou, Souti Chattopadhyay, Swabha Swayamdipta, Xiang Ren
Findings of the Annual Meeting of the Association for Computational Linguistics (**ACL**), 2025

[183] REL-A.I.: An Interaction-Centered Approach To Measuring Human-LM Reliance
Kaitlyn Zhou, Jena D. Hwang, **Xiang Ren**, Nouha Dziri, Dan Jurafsky, Maarten Sap
North American Chapter of the Association for Computational Linguistics (**NAACL**), 2025

[182] CAVE: Controllable Authorship Verification Explanations
Sahana Ramnath, Kartik Pandey, Elizabeth Boschee, Xiang Ren
North American Chapter of the Association for Computational Linguistics (**NAACL**), 2025

[181] Logits of API-Protected LLMs Leak Proprietary Information
Matthew Finlayson, **Xiang Ren**, Swabha Swayamdipta
Conference on Language Modeling (**COLM**), 2025

[180] The Surprising Effectiveness of Membership Inference with Simple N-Gram Coverage
Skyler Hallinan, Jaehun Jung, Melanie Sclar, Ximing Lu, Abhilasha Ravichander, Sahana Ramnath, Yejin Choi, Sai Praneeth Karimireddy, Niloofar Mireshghallah, Xiang Ren
Conference on Language Modeling (**COLM**), 2025

[179] Improving LLM Personas via Rationalization with Psychological Scaffolds
Brihi Joshi, **Xiang Ren**, Swabha Swayamdipta, Rik Koncel-Kedziorski, Tim Paek
Findings of the Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2025

[178] Amulet: Putting Complex Multi-Turn Conversations on the Stand with LLM Juries
Sahana Ramnath, Anurag Mudgil, Brihi Joshi, Skyler Hallinan, Xiang Ren
Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2025

[177] Mixing Inference-time Experts for Enhancing LLM Reasoning
Soumya Sanyal, Tianyi Xiao, Xiang Ren
Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2025

[176] Rethinking Backdoor Detection Evaluation for Language Models

Jun Yan, Wenjie Mo, **Xiang Ren**, Robin Jia
Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2025

[175] Diagnosing Memorization in Chain-of-Thought Reasoning, One Token at a Time
Huihan Li, You Chen, Siyuan Wang, Yixin He, Ninareh Mehrabi, Rahul Gupta, Xiang Ren
Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2025

[174] Stepwise Informativeness Search for Improving LLM Reasoning
Siyuan Wang, Enda Zhao, Zhongyu Wei, Xiang Ren
Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2025

[173] Demystifying Language Model Forgetting with Low-rank Example Associations
Xisen Jin, Xiang Ren
Advances in Neural Information Processing Systems (**NeurIPS**), 2025

[172] Better Language Model Inversion by Compactly Representing Next-Token Distributions
Murtaza Nazir, Matthew Finlayson, John X. Morris, **Xiang Ren**, Swabha Swayamdipta
Advances in Neural Information Processing Systems (**NeurIPS**), 2025

[171] Tailoring Self-Rationalizers with Multi-Reward Distillation
Sahana Ramnath, Brihi Joshi, Skyler Hallinan, Ximing Lu, Liunian Harold Li, Aaron Chan, Jack Hessel, Yejin Choi, Xiang Ren
International Conference on Representation Learning (**ICLR**), 2024

[170] (InThe)WildChat: 570K ChatGPT Interaction Logs In The Wild
Wenting Zhao, **Xiang Ren**, Jack Hessel, Claire Cardie, Yejin Choi
International Conference on Representation Learning (**ICLR**) **Spotlight**, 2024

[169] Phenomenal Yet Puzzling: Testing Inductive Reasoning Capabilities of Language Models with Hypothesis Refinement
Linlu Qiu, Liwei Jiang, Ximing Lu, Melanie Sclar, Valentina Pyatkin, Chandra Bhagavatula, Bailin Wang, Yoon Kim, Yejin Choi, Nouha Dziri, Xiang Ren
International Conference on Representation Learning (**ICLR**) **Oral**, 2024

[168] Instruction-following evaluation through verbalizer manipulation
Shiyang Li, Jun Yan, Hai Wang, Zheng Tang, **Xiang Ren**, Vijay Srinivasan, Hongxia Jin
Findings of the North American Chapter of the Association for Computational Linguistics (**NAACL**), 2024

[167] Virtual Prompt Injection for Instruction-Tuned Large Language Models
Jun Yan, Vikas Yadav, Shiyang Li, Lichang Chen, Zheng Tang, Hai Wang, Vijay Srinivasan, **Xiang Ren**, Hongxia Jin
North American Chapter of the Association for Computational Linguistics (**NAACL**), 2024

[166] Are Machines Better at Complex Reasoning? Unveiling Human-Machine Inference Gaps in Entailment Verification
Soumya Sanyal, Tianyi Xiao, Jiacheng Liu, Wenya Wang, Xiang Ren
Findings of the Annual Meeting of the Association for Computational Linguistics (**ACL**), 2024

[165] Relying on the Unreliable: The Impact of Language Models' Reluctance to Express Uncertainty
Kaitlyn Zhou, Jena D. Hwang, **Xiang Ren**, Maarten Sap
Annual Meeting of the Association for Computational Linguistics (**ACL**), 2024

[164] Can LLMs Reason with Rules? Logic Scaffolding for Stress-Testing and Improving LLMs
Siyuan Wang, Zhongyu Wei, Yejin Choi, Xiang Ren
Annual Meeting of the Association for Computational Linguistics (**ACL**), 2024

[163] What Will My Model Forget? Forecasting Forgotten Examples in Language Model Refinement
Xisen Jin, Xiang Ren
International Conference on Machine Learning (**ICML**) **Spotlight**, 2024

[162] WildChat: 1M ChatGPT Interaction Logs in the Wild
Wenting Zhao, **Xiang Ren**, Jack Hessel, Claire Cardie, Yejin Choi
International Conference on Representation Learning (**ICLR**), 2024

[161] PlaSma: Procedural Knowledge Models for Language-based Planning and Re-Planning
Faeze Brahman, Chandra Bhagavatula, Valentina Pyatkin, Jena D. Hwang, Xiang Lorraine Li, Hirona Jacqueline Arai, Soumya Sanyal, Keisuke Sakaguchi, **Xiang Ren**, Yejin Choi
International Conference on Representation Learning (**ICLR**), 2024

[160] Symbolic Working Memory Enhances Language Models for Complex Rule Application
Siyuan Wang, Zhongyu Wei, Yejin Choi, Xiang Ren
Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2024

[159] In Search of the Long-Tail: Systematic Generation of Long-Tail Knowledge via Logical Rule Guided Search
Huihan Li, Yuting Ning, Zeyi Liao, Siyuan Wang, Xiang Lorraine Li, Ximing Lu, Faeze Brahman, Wenting Zhao, Yejin Choi, Xiang Ren
Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2024

[158] WinoViz: Probing Visual Properties of Objects Under Different States
Woojeong Jin, Tejas Srinivasan, Jesse Thomason, Xiang Ren
ICLR Workshop on Secure and Trustworthy Large Language Models, 2024

[157] WildVis: Open Source Visualizer for Million-Scale Chat Logs in the Wild
Yuntian Deng, Wenting Zhao, Jack Hessel, **Xiang Ren**, Claire Cardie, Yejin Choi
Conference on Empirical Methods in Natural Language Processing (**EMNLP**) **Demo**, 2024

[156] CULTURE-GEN: Revealing Global Cultural Perception in Language Models through Natural Language Prompting
Huihan Li, Liwei Jiang, Jena D. Huang, Hyunwoo Kim, Sebastin Santy, Taylor Sorensen, Bill Yuchen Lin, Nouha Dziri, **Xiang Ren**, Yejin Choi
ACL 2024 C3NLP Workshop, 2024

[155] Stress-Testing Long-Context Language Models with Lifelong ICL and Task Haystack
Xiaoyue Xu, Qinyuan Ye, Xiang Ren
Advances in Neural Information Processing Systems (**NeurIPS**) Dataset and Benchmark Track, 2024

[154] Self-Discover: Large Language Models Self-Compose Reasoning Structures
Pei Zhou, Jay Pujara, **Xiang Ren**, Xinyun Chen, Heng-Tze Cheng, Quoc V. Le, Ed H. Chi, Denny Zhou, Swaroop Mishra, Huaixiu Steven Zheng
Advances in Neural Information Processing Systems (**NeurIPS**), 2024

[153] Phenomenal Yet Puzzling: Testing Inductive Reasoning Capabilities of Language Models with Hypothesis Refinement
Linlu Qiu, Liwei Jiang, Ximing Lu, Melanie Sclar, Valentina Pyatkin, Chandra Bhagavatula, Bailin Wang, Yoon Kim, Yejin Choi, Nouha Dziri, Xiang Ren
International Conference on Representation Learning (**ICLR**) **Oral**, 2023

[152] (InThe)WildChat: 570K ChatGPT Interaction Logs In The Wild
Wenting Zhao, **Xiang Ren**, Jack Hessel, Claire Cardie, Yejin Choi, Yuntian Deng
International Conference on Representation Learning (**ICLR**) **Spotlight**, 2023

[152] Tailoring Self-Rationalizers with Multi-Reward Distillation
Sahana Ramnath, Brihi Joshi, Skyler Hallinan, Ximing Lu, Liunian Harold Li, Aaron Chan, Jack Hessel, Yejin Choi, Xiang Ren
International Conference on Representation Learning (**ICLR**), 2023

[151] PlaSma: Making Small Language Models Better Procedural Knowledge Models for (Counterfactual) Planning
Faeze Brahman, Chandra Bhagavatula, Valentina Pyatkin, Jena D Hwang, Xiang Lorraine Li, Hirona J Arai, Soumya Sanyal, Keisuke Sakaguchi, **Xiang Ren**, Yejin Choi
International Conference on Representation Learning (**ICLR**), 2023

[150] SwiftSage: A Generative Agent with Fast and Slow Thinking for Complex Interactive Tasks
Bill Yuchen Lin, Yicheng Fu, Karina Yang, Prithviraj Ammanabrolu, Faeze Brahman, Shiyu Huang, Chandra Bhagavatula, Yejin Choi, Xiang Ren
Conference on Neural Information Processing Systems (**NeurIPS**) **Spotlight**, 2023

[149] Inference-Time Policy Adapters (IPA): Tailoring Extreme-Scale LMs without Fine-tuning
Ximing Lu, Faeze Brahman, Peter West, Jaehun Jang, Khyathi Chandu, Abhilasha Ravichander, Lianhui Qin, Prithviraj Ammanabrolu, Liwei Jiang, Sahana Ramnath, Nouha Dziri, Jillian Fisher, Bill Yuchen Lin, Skyler Hallinan, **Xiang Ren**, Sean Welleck, Yejin Choi

Conference on Neural Information Processing Systems (**NeurIPS**) **Spotlight**, 2023

[148] Inference-Time Policy Adapters (IPA): Tailoring Extreme-Scale LMs without Fine-tuning
Ximing Lu, Faeze Brahman, Peter West, Jaehun Jang, Khyathi Chandu, Abhilasha Ravichander, Lianhui Qin, Prithviraj Ammanabrolu, Liwei Jiang, Sahana Ramnath, Nouha Dziri, Jillian Fisher, Bill Yuchen Lin, Skyler Hallinan, **Xiang Ren**, Sean Welleck, Yejin Choi
The Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2023

[147] How Predictable Are Large Language Model Capabilities? A Case Study on BIG-bench
Qinyuan Ye, Harvey Yiyun Fu, **Xiang Ren**, Robin Jia
The Conference on Empirical Methods in Natural Language Processing (**Findings of EMNLP**), 2023

[146] Estimating Large Language Model Capabilities without Labeled Test Data
Harvey Yiyun Fu, Qinyuan Ye, Albert Xu, **Xiang Ren**, Robin Jia
The Conference on Empirical Methods in Natural Language Processing (**Findings of EMNLP**), 2023

[145] Exploring Distributional Shifts in Large Language Models for Code Analysis
Shushan Arakelyan, Rocktim Jyoti Das, Yi Mao, Xiang Ren
The Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2023

[144] Bootstrap Your Own Skills: Learning to Solve New Tasks with Large Language Model Guidance
Jesse Zhang, Jiahui Zhang, Karl Pertsch, Ziyi Liu, **Xiang Ren**, Minsuk Chang, Shao-Hua Sun, Joseph J Lim
Conference on Robot Learning (**CORL**), 2023

[143] SCOTT: Self-Consistent Chain-of-Thought Distillation
Peifeng Wang, Zhengyang Wang, Zheng Li, Yifan Gao, Bing Yin, **Xiang Ren**
Annual Meeting of the Association for Computational Linguistics (**ACL**), 2023
ACL Outstanding Paper Award

[142] Are Machine Rationales (Not) Useful to Humans? Measuring and Improving Human Utility of Free-text Rationales
Brihi Joshi, Ziyi Liu, Sahana Ramnath, Aaron Chan, Zhewei Tong, Shaoliang Nie, Qifan Wang, Yejin Choi, Xiang Ren
Annual Meeting of the Association for Computational Linguistics (**ACL**), 2023

[141] FiD-ICL: A Fusion-in-Decoder Approach for Efficient In-Context Learning
Qinyuan Ye, Iz Beltagy, Matthew E. Peters, **Xiang Ren**, Hannaneh Hajishirzi
Annual Meeting of the Association for Computational Linguistics (**ACL**), 2023

[140] Symbolic Chain-of-Thought Distillation: Small Models Can Also "Think" Step-by-Step
Liunian Harold Li, Jack Hessel, Youngjae Yu, **Xiang Ren**, Kai-Wei Chang and Yejin Choi
Annual Meeting of the Association for Computational Linguistics (**ACL**), 2023

[139] CoNAL: Anticipating Outliers with Large Language Models
Albert Xu, **Xiang Ren**, Robin Jia
Annual Meeting of the Association for Computational Linguistics (**ACL**), 2023

[138] PairReranker: Pairwise Learning to Rank for Decoding Language Models
Dongfu Jiang, **Xiang Ren**, Bill Yuchen Lin
Annual Meeting of the Association for Computational Linguistics (**ACL**), 2023

[137] APOLLO: A Simple Approach for Adaptive Pretraining of Language Models for Logical Reasoning
Soumya Sanyal, Yichong Xu, Shuohang Wang, Ziyi Yang, Reid Pryzant, Wenhao Yu, Chenguang Zhu, Xiang Ren
Annual Meeting of the Association for Computational Linguistics (**ACL**), 2023

[136] REV: Information-Theoretic Evaluation of Free-Text Rationales
Hanjie Chen, Faeze Brahman, **Xiang Ren**, Yangfeng Ji, Yejin Choi, Swabha Swayamdipta
Annual Meeting of the Association for Computational Linguistics (**ACL**), 2023

[135] Textual Backdoor Attacks with Iterative Trigger Injection
Jun Yan, Vansh Gupta, Xiang Ren
Annual Meeting of the Association for Computational Linguistics (**ACL**), 2023

[134] Cross-lingual Lifelong Learning
Meryem M'hamdi, **Xiang Ren**, Jonathan May
Annual Meeting of the Association for Computational Linguistics (**ACL**), 2023

[133] An AI Dungeon Master's Guide: Learning to Converse and Guide with Intents and Theory-of-Mind in Dungeons and Dragons
Pei Zhou, Andrew Zhu, Jennifer Hu, Jay Pujara, **Xiang Ren**, Chris Callison-Burch, Yejin Choi, Prithviraj Ammanabrolu
Annual Meeting of the Association for Computational Linguistics (**ACL**), 2023

[132] Dataless Knowledge Fusion by Merging Weights of Language Models

Xisen Jin, **Xiang Ren**, Daniel Preotiuc-Pietro, Pengxiang Cheng
International Conference on Representation Learning (**ICLR**), 2023

[131] PINTO: Faithful Language Reasoning Using Prompt-Generated Rationales
Peifeng Wang, Aaron Chan, Filip Ilievski, Muhao Chen, Xiang Ren
International Conference on Representation Learning (**ICLR**), 2023

[130] On Grounded Planning for Embodied Tasks with Language Models
Bill Yuchen Lin, Chengsong Huang, Qian Liu, Wenda Gu, Sam Sommerer, **Xiang Ren**
The AAAI Conference on Artificial Intelligence (**AAAI**), 2023

[129] Retweet-BERT: Characterizing Echo Chambers by Leveraging Language Features and Information Diffusion in Social Media
Julie Jiang, **Xiang Ren**, and Emilio Ferrara
The International AAAI Conference on Web and Social Media (**ICWSM**), 2022

[128] Reflect, Not Reflex: Inference-Based Common Ground Improves Dialogue Response Quality
Pei Zhou, Hyundong J. Cho, Pegah Jandaghi, Dong-Ho Lee, Bill Yuchen Lin, Jay Pujara, **Xiang Ren**
The Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2022

[127] Machine Translation Robustness to Natural Asemantic Variation
Jacob Bremerman, **Xiang Ren**, Jonathan May
The Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2022

[126] ER-Test: Evaluating Explanation Regularization Methods for NLP Models
Brihi Joshi, Aaron Chan, Ziyi Liu, Shaoliang Nie, Maziar Sanjabi, Hamed Firooz, **Xiang Ren**
Findings of The Conference on Empirical Methods in Natural Language Processing (**EMNLP Findings**), 2022

[125] RobustLR: Evaluating Robustness to Logical Perturbation in Deductive Reasoning
Soumya Sanyal, Zeyi Liao, **Xiang Ren**
The Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2022

[124] Eliciting Transferability in Multi-task Learning with Task-level Mixture-of-Experts
Qinyuan Ye, Juan Zha, **Xiang Ren**
Findings of The Conference on Empirical Methods in Natural Language Processing (**EMNLP Findings**), 2022

[123] Unsupervised Cross-Task Generalization via Retrieval Augmentation
Bill Yuchen Lin, Kangmin Tan, Chris Miller, Beiwen Tian, **Xiang Ren**
Conference on Neural Information Processing Systems (**NeurIPS**), 2022

[122] NS3: Neuro-Symbolic Semantic Code Search
Shushan Arakelyan, Anna Hakhverdyan, Miltiadis Allamanis, Christophe Hauser, Luis Antonio Garcia, **Xiang Ren**
Conference on Neural Information Processing Systems (**NeurIPS**), 2022

[121] UniREx: A Unified Learning Framework for Language Model Rationale Extraction
Aaron Chan, Maziar Sanjabi, Lambert Mathias, Liang Tan, Shaoliang Nie, Xiaochang Peng, **Xiang Ren**, Hamed Firooz
International Conference on Machine Learning (**ICML**), 2022

[120] Lifelong Pretraining: Continually Adapting Language Models to Emerging Corpora
Xisen Jin, Dejiao Zhang, Henghui Zhu, Wei Xiao, Shang-Wen Li, Xiaokai Wei, Andrew Arnold, **Xiang Ren**
Conference of the North American Chapter of the Association for Computational Linguistics (**NAACL**), 2022

[119] Sparse Distillation: Speeding Up Text Classification by Using Bigger Models
Qinyuan Ye, Madian Khabsa, Mike Lewis, Sinong Wang, **Xiang Ren**, Aaron Jaech
Conference of the North American Chapter of the Association for Computational Linguistics (**NAACL**), 2022

[118] FedNLP: A Research Platform for Federated Learning in Natural Language Processing
Bill Yuchen Lin, Chaoyang He, Zihang Zeng, Hulin Wang, Yufen Huang, Mahdi Soltanolkotabi, **Xiang Ren**, Salman Avestimehr
Conference of the North American Chapter of the Association for Computational Linguistics (**Findings of NAACL**), 2022

[117] On the Robustness of Reading Comprehension Models to Entity Renaming
Jun Yan, Yang Xiao, Sagnik Mukherjee, Bill Yuchen Lin, Robin Jia, **Xiang Ren**
Conference of the North American Chapter of the Association for Computational Linguistics (**NAACL**), 2022

[116] NewsEdits: A Dataset of News Article Revision Histories and a Novel Document-Level Reasoning Challenge
Alexander Spangher, **Xiang Ren**, Jonathan May, Nanyun Peng
Conference of the North American Chapter of the Association for Computational Linguistics (**NAACL**), 2022.
NAACL Outstanding Paper Award

[115] FaiRR: Faithful and Robust Deductive Reasoning over Natural Language
Soumya Sanyal, Harman Singh, **Xiang Ren**
Annual Meeting of the Association for Computational Linguistics (**ACL**), 2022

[114] A Good Prompt Is Worth Millions of Parameters?: Low-resource Prompt-based Learning for Vision-Language Models
Woojeong Jin, Yu Cheng, Yelong Shen, Weizhu Chen, **Xiang Ren**
Annual Meeting of the Association for Computational Linguistics (**ACL**), 2022

[113] Fixing Errors in the Wild: On Continual Model Refinement in Out-of-Distribution Data Streams
Bill Yuchen Lin, Sida Wang, Xi Victoria Lin, Robin Jia, Lin Xiao, **Xiang Ren**, Scott Yih
Annual Meeting of the Association for Computational Linguistics (**ACL**), 2022

[112] Good Examples Make A Faster Learner: Simple Demonstration-based Learning for Low-resource NER
Dong-Ho Lee, Mahak Agarwal, Akshen Kadakia, Jay Pujara, **Xiang Ren**
Annual Meeting of the Association for Computational Linguistics (**ACL**), 2022

[111] Think Before You Speak: Using Self-talk to Generate Implicit Commonsense Knowledge for Response Generation
Pei Zhou, Karthik Gopalakrishnan, Behnam Hedayatnia, Seokhwan Kim, Jay Pujara, **Xiang Ren**, Yang Liu, Dilek Hakkani-Tur
Annual Meeting of the Association for Computational Linguistics (**ACL**), 2022

[110] Leveraging Visual Knowledge in Language Tasks: An Empirical Study on Intermediate Pre-training for Cross-Modal Knowledge Transfer
Woojeong Jin, Dong-Ho Lee, Chenguang Zhu, Jay Pujara, **Xiang Ren**
Annual Meeting of the Association for Computational Linguistics (**ACL**), 2022

[109] KG-FiD: Infusing Knowledge Graph in Fusion-in-Decoder for Open-Domain Question Answering
Donghan Yu, Chenguang Zhu, Yuwei Fang, Wenhao Yu, Shuohang Wang, Yichong Xu, **Xiang Ren**, Yiming Yang, Michael Zeng
Annual Meeting of the Association for Computational Linguistics (**ACL**), 2022

[108] Contextualized Scene Imagination for Generative Commonsense Reasoning
Peifeng Wang, Jonathan Zamora, Junfeng Liu, Filip Ilievski, Muhao Chen, **Xiang Ren**
International Conference on Representation Learning (**ICLR**), 2022

[107] Refining Language Models with Compositional Explanations
Huihan Yao, Ying Chen, Qinyuan Ye, Xisen Jin, **Xiang Ren**
Annual Conference on Neural Information Processing Systems (**NeurIPS**), 2021 (**Spotlight**, ac rate: 3%)

[106] SalKG: Learning From Knowledge Graph Explanations for Commonsense Reasoning
Aaron Chan, Boyuan Long, Jiashu Xu, Soumya Sanyal, Tanishq Gupta, **Xiang Ren**
Annual Conference on Neural Information Processing Systems (**NeurIPS**), 2021 (ac rate: 26%)

[105] Gradient Based Memory Editing for Task-Free Continual Learning
Xisen Jin, Arka Sadhu, Junyi Du, **Xiang Ren**
Annual Conference on Neural Information Processing Systems (**NeurIPS**), 2021 (ac rate: 26%)

[104] **CrossFit**: A Few-shot Learning Challenge for Cross-task Generalization in NLP
Qinyuan Ye, Yuchen Lin, **Xiang Ren**
The Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2021 (ac rate: 24.6%)

[103] **RICA**: Evaluating Robust Inference Capabilities Based on Commonsense Axioms
Pei Zhou, Rahul Khanna, Bill Yuchen Lin, Daniel Ho, Jay Pujara, **Xiang Ren**
The Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2021 (ac rate: 24.6%)

[102] *Learn Continually, Generalize Rapidly*: Lifelong Knowledge Accumulation for Few-shot Learning
Xisen Jin, Bill Yuchen Lin, Mohammad Rostami, **Xiang Ren**
The Conference on Empirical Methods in Natural Language Processing (**Findings of EMNLP**), 2021

[101] Discretized Integrated Gradients for Explaining Language Models
Soumya Sanyal, **Xiang Ren**
The Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2021 (ac rate: 24.6%)

[100] On the Influence of Masking Policies in Intermediate Pre-training
Qinyuan Ye, Belinda Z. Li, Sinong Wang, Benjamin Bolte, Hao Ma, Wen-tau Yih, **Xiang Ren**, Madian Khabsa
The Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2021 (ac rate: 24.6%)

[99] *Lawyers are Dishonest?* : Quantifying Representational Harms in Commonsense Knowledge Resources
Ninareh Mehrabi, Pei Zhou, Fred Morstatter, Jay Pujara, **Xiang Ren**, Aram Galstyan
The Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2021 (ac rate: 24.6%)

[98] Probing Commonsense Explanation in Dialogue Response Generation
Pei Zhou, Pegah Jandaghi, Hyundong Cho, Bill Yuchen Lin, Jay Pujara, **Xiang Ren**
The Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2021 (ac rate: 24.6%)

[97] ECONET: Effective Continual Pretraining of Language Models for Event Temporal Reasoning
Rujun Han, **Xiang Ren**, Nanyun Peng
The Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2021 (ac rate: 24.6%)

[96] Extract, Denoise and Enforce: Evaluating and Improving Concept Preservation for Text-to-Text Generation
Yuning Mao, Wenchang Ma, Deren Lei, Jiawei Han, **Xiang Ren**
The Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2021 (ac rate: 24.6%)

[95] Cross-Attention is All You Need: Adapting Pretrained Transformers for Machine Translation
Mozhdeh Gheini, **Xiang Ren**, Jonathan May
The Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2021 (ac rate: 24.6%)

[94] MSD: Saliency-aware Knowledge Distillation for Multimodal Understanding
Woojeong Jin, Maziar Sanjabi, Shaoliang Nie, Liang Tan, **Xiang Ren**, Hamed Firooz
The Conference on Empirical Methods in Natural Language Processing (**Findings of EMNLP**), 2021

[93] **RockNER**: Evaluating the Robustness of Named Entity Recognition with Natural Adversarial Attacks
Bill Yuchen Lin, Wenyang Gao, Jun Yan, Ryan Moreno, **Xiang Ren**
The Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2021 (ac rate: 24.6%)

[92] **ForecastQA**: A Question Answering Challenge for Event Forecasting
Woojeong Jin, Rahul Khanna, Suji Kim, Dong-Ho Lee, Fred Morstatter, Aram Galstyan, **Xiang Ren**
Annual Meeting of the Association for Computational Linguistics (**ACL**), 2021 (ac rate: 21.3%)

[91] Learning to Generate Task-specific Adapters from Task Description
Qinyuan Ye, **Xiang Ren**
Annual Meeting of the Association for Computational Linguistics (**ACL**), 2021 (ac rate: 21.3%)

[90] *Common Sense Beyond English*: Evaluating and Improving Multilingual Language Models for Common Sense Reasoning
Bill Yuchen Lin, Seyeon Lee, Xiaoyang Qiao, **Xiang Ren**
Annual Meeting of the Association for Computational Linguistics (**ACL**), 2021 (ac rate: 21.3%)

[89] AdaTag: Multi-Attribute Value Extraction from Product Profiles with Adaptive Decoding
Jun Yan, Nasser Zalmout, Yan Liang, Christan Grant, **Xiang Ren**, Xin Luna Dong
Annual Meeting of the Association for Computational Linguistics (**ACL**), 2021 (ac rate: 21.3%)

[88] Learning Contextualized Knowledge Structures for Commonsense Reasoning
Jun Yan, Mrigank Raman, Aaron Chan, Tianyu Zhang, Ryan Rossi, Handong Zhao, Sungchul Kim, Nedim Lipka, **Xiang Ren**
Annual Meeting of the Association for Computational Linguistics (**Findings of ACL**), 2021

[88] **RiddleSense**: A Challenge Dataset for Answering Riddle-Style Commonsense Questions
Bill Yuchen Lin, Ziyi Wu, Yichi Yang, Dong-Ho Lee, **Xiang Ren**
Annual Meeting of the Association for Computational Linguistics (**Findings of ACL**), 2021

[87] Do Language Models Perform Generalizable Commonsense Inference?
Peifeng Wang, Filip Ilievski, Muhao Chen, **Xiang Ren**
Annual Meeting of the Association for Computational Linguistics (**Findings of ACL**), 2021

[86] Semi-Automated Protocol Disambiguation and Code Generation
Jane Yen, Tamás Lévai, Qinyuan Ye, **Xiang Ren**, Ramesh Govindan, Barath Raghavan
Annual conference of ACM SIGCOMM (**SIGCOMM**), 2021

[85] Pre-training Text-to-Text Transformers for Concept-centric Common Sense
Wangchunshu Zhou, Dong-Ho Lee, Ravi Kiran Selvam, Seyeon Lee, Bill Yuchen Lin, **Xiang Ren**
International Conference on Representation Learning (**ICLR**), 2021

[84] Learning to Deceive Knowledge Graph Augmented Models via Targeted Perturbation
Mrigank Raman, Siddhant Agarwal, Peifeng Wang, Aaron Chan, Hansen Wang, Sungchul Kim,
Ryan Rossi, Handong Zhao, Nedim Lipka, **Xiang Ren**
International Conference on Representation Learning (**ICLR**), 2021
Best Paper Award Runner-up at NeurIPS 2020 KR2ML Workshop

[83] Differentiable Open-Ended Commonsense Reasoning
Bill Yuchen Lin, Haitian Sun, Bhuwan Dhingra, Manzil Zaheer, **Xiang Ren**, William W. Cohen
Conference of the North American Chapter of the Association for Computational Linguistics (**NAACL**), 2021 (ac rate: 26%)

[82] On Transferability of Bias Mitigation Effects in Language Model Fine-Tuning
Xisen Jin, Francesco Barbieri, Brendan Kennedy, Aida Mostafazadeh Davani, Leonardo Neves, **Xiang Ren**
Conference of the North American Chapter of the Association for Computational Linguistics (**NAACL**), 2021 (ac rate: 26%)

[81] X-METRA-ADA: Cross-lingual Meta-Transfer learning Adaptation to Natural Language Understanding
Meryem M'hamdi, Doo Soon Kim, Franck Dernoncourt, Trung Bui, **Xiang Ren**, Jonathan May
Conference of the North American Chapter of the Association for Computational Linguistics (**NAACL**), 2021 (ac rate: 26%)

[80] TaxoClass: Hierarchical Multi-Label Text Classification Using Only Class Names
Jiaming Shen, Wenda Qiu, Yu Meng, Jingbo Shang, **Xiang Ren**, Jiawei Han
Conference of the North American Chapter of the Association for Computational Linguistics (**NAACL**), 2021 (ac rate: 26%)

[79] IsoBN: Fine-Tuning BERT with Isotropic Batch Normalization
Wenxuan Zhou, Bill Yuchen Lin, **Xiang Ren**
The AAAI Conference on Artificial Intelligence (**AAAI**), 2021

[78] Time-Series Event Prediction with Evolutionary State Graph
Wenjie Hu, Yang Yang, Ziqiang Cheng, Carl Yang, **Xiang Ren**
ACM International Web Search and Data Mining conference (**WSDM**), 2021

[77] Visually Grounded Continual Learning of Compositional Phrases
Xisen Jin, Junyi Du, Arka Sadhu, Ram Nevatia, **Xiang Ren**

The 2020 Conference on Empirical Methods in Natural Language Processing **EMNLP**), 2020 (ac rate: 22.4%)

[76] Scalable Multi-Hop Relational Reasoning for Knowledge-Aware Question Answering
Yanlin Feng, Xinyue Chen, Bill Yuchen Lin, Peifeng Wang, Jun Yan, **Xiang Ren**
The 2020 Conference on Empirical Methods in Natural Language Processing **EMNLP**), 2020 (ac rate: 22.4%)

[75] Recurrent Event Network: Autoregressive Structure Inference over Temporal Knowledge Graphs
Woojeong Jin, Meng Qu, Xisen Jin, **Xiang Ren**
The 2020 Conference on Empirical Methods in Natural Language Processing **EMNLP**), 2020 (ac rate: 22.4%)

[74] Birds have four legs?! NumerSense: Probing Numerical Commonsense Knowledge of Pre-trained Language Models
Bill Yuchen Lin, Seyeon Lee, Rahul Khanna, **Xiang Ren**
The 2020 Conference on Empirical Methods in Natural Language Processing **EMNLP**), 2020 (ac rate: 22.4%)

[73] Learning Collaborative Agents with Rule Guidance for Knowledge Graph Reasoning
Deren Lei, Gangrong Jiang, Xiaotao Gu, Kexuan Sun, Yuning Mao, **Xiang Ren**
The 2020 Conference on Empirical Methods in Natural Language Processing **EMNLP**), 2020 (ac rate: 22.4%)

[72] Multi-document Summarization with Maximal Marginal Relevance-guided Reinforcement Learning
Yuning Mao, Yanru Qu, Yiqing Xie, **Xiang Ren**, Jiawei Han
The 2020 Conference on Empirical Methods in Natural Language Processing **EMNLP**), 2020 (ac rate: 22.4%)

[71] SynSetExpan: An Iterative Framework for Joint Entity Set Expansion and Synonym Discovery
Jiaming Shen, Wenda Qiu, Jingbo Shang, Michelle Vanni, **Xiang Ren**, Jiawei Han
The 2020 Conference on Empirical Methods in Natural Language Processing **EMNLP**), 2020 (ac rate: 22.4%)

[70] Teaching Machine Comprehension with Compositional Explanations
Qinyuan Ye, Xiao Huang, **Xiang Ren**
The 2020 Conference on Empirical Methods in Natural Language Processing (**EMNLP** Findings), 2020

[69] CommonGen: A Constrained Text Generation Challenge for Generative Commonsense Reasoning
Bill Yuchen Lin, Ming Shen, Pei Zhou, Chandra Bhagavatula, Yejin Choi, **Xiang Ren**
The 2020 Conference on Empirical Methods in Natural Language Processing (**EMNLP** Findings), 2020

[68] RICA: Evaluating Robust Inference Capabilities Based on Commonsense Axioms
Pei Zhou, Rahul Khanna, Bill Yuchen Lin, Daniel Ho, Jay Pujara, **Xiang Ren**
The 2020 Conference on Empirical Methods in Natural Language Processing (**EMNLP** Findings), 2020

[67] Connecting the Dots: A Knowledgeable Path Generator for Commonsense Question Answering
Peifeng Wang, Nanyun Peng, Pedro Szekely, **Xiang Ren**
The 2020 Conference on Empirical Methods in Natural Language Processing (**EMNLP** Findings), 2020

[66] Will this Idea Spread Beyond Academia? Understanding Knowledge Transfer of Scientific Concepts across Text Corpora
Hancheng Cao, Mengjie Cheng, Zhepeng Cen, Daniel McFarland, **Xiang Ren**
The 2020 Conference on Empirical Methods in Natural Language Processing (**EMNLP** Findings), 2020

[65] Learning to Contextually Aggregate Multi-Source Supervision for Sequence Labeling
Ouyu Lan*, Xiao Huang*, Bill Yuchen Lin, He Jiang, Liyuan Liu, X. Ren
Annual Meeting of the Association for Computational Linguistics **ACL**), 2020 (ac rate: 22.7%)

[64] Facet-Aware Evaluation for Extractive Summarization
Yuning Mao, Liyuan Liu, Qi Zhu, X. Ren, Jiawei Han
Annual Meeting of the Association for Computational Linguistics **ACL**), 2020 (ac rate: 22.7%)

[63] TriggerNER: Learning with Entity Triggers as Explanation for Named Entity Recognition
Bill Yuchen Lin, Dongho Lee, Ming Shen, Xiao Huang, Ryan Moreno, Prashant Shiralkar, X. Ren.
Annual Meeting of the Association for Computational Linguistics **ACL**), 2020 (ac rate: 22.7%)

[62] Contextualizing Hate Speech Classifiers with Post-hoc Explanation
Brendan Kennedy*, Xisen Jin*, Aida Mostafazadeh Davani, Morteza Dehghani, X. Ren.
Annual Meeting of the Association for Computational Linguistics **ACL**), 2020 (ac rate: 22.7%)

[61] Temporal Attribute Prediction via Joint Modeling of Multi-Relational Structure Evolution
Sankalp Garg*, Navodita Sharma*, Woojeong Jin, **Xiang Ren**.
International Joint Conference on Artificial Intelligence (**IJCAI**), 2020

[60] Alleviate Dataset Shift Problem in Fine-grained Entity Typing with Virtual Adversarial Training
Haochen Shi, Siliang Tang, Xiaotao Gu, Bo Chen, **Xiang Ren**.
International Joint Conference on Artificial Intelligence (**IJCAI**), 2020

[59] Learning from Explanations with Neural Execution Tree
Ziqi Wang, Yujia Qin, Wenxuan Zhou, Jun Yan, Qinyuan Ye, Leonardo Neves, Zhiyuan Liu, **Xiang Ren**.
International Conference on Representation Learning (**ICLR**), 2020 (ac rate: 26.5%)

[58] Towards Hierarchical Importance Attribution: Explaining Compositional Semantics for Neural Sequence Models
Xisen Jin, Junyi Du, Zhongyu Wei, Xiangyang Xue, **Xiang Ren**.

International Conference on Representation Learning (**ICLR**), 2020 (**Spotlight**)

[57] NERO: A Neural Rule Grounding Framework for Label-Efficient Relation Extraction
Wenxuan Zhou, Hongtao Lin, Bill Yuchen Lin, Ziqi Wang, Junyi Du, Leonardo Neves, **Xiang Ren**.
The Web Conference (**WWW**), 2020. **Runner-up of Best Paper Award**

[56] KagNet: Learning to Answer Commonsense Questions with Knowledge-Aware Graph Networks
Bill Yuchen Lin, Xinyue Chen, Jamin Chen, **Xiang Ren**.
Conference on Empirical Methods in Natural Language Processing **EMNLP**), 2019 (ac rate: 24.6%)

[55] Collaborative Policy Learning for Open Knowledge Graph Reasoning
Cong Fu, Tong Chen, Meng Qu, Woojeong Jin, **Xiang Ren**.
Conference on Empirical Methods in Natural Language Processing **EMNLP**), 2019 (ac rate: 24.6%)

[54] Looking Beyond Label Noise: Shifted Label Distribution Matters in Distantly Supervised Relation Extraction
Qinyuan Ye, Liyuan Liu, Maosen Zhang, **Xiang Ren**.
Conference on Empirical Methods in Natural Language Processing **EMNLP**), 2019 (ac rate: 24.6%)

[53] Learning Dynamic Context Augmentation for Global Entity Linking
Xiyuan Yang, Xiaotao Gu, Sheng Lin, Siliang Tang, Yueting Zhuang, Fei Wu, Zhigang Chen, Guoping Hu, **Xiang Ren**.
Conference on Empirical Methods in Natural Language Processing **EMNLP**), 2019 (ac rate: 24.6%)

[52] Hierarchical Text Classification with Reinforced Label Assignment
Yuning Mao, Jingjing Tian, Jiawei Han, **Xiang Ren**.
Conference on Empirical Methods in Natural Language Processing **EMNLP**), 2019 (ac rate: 24.6%)

[51] HMEAE: Hierarchical Modular Event Argument Extraction
Xiaozhi Wang, Ziqi Wang, Xu Han, Zhiyuan Liu, Juanzi Li, Peng Li, Maosong Sun, Jie Zhou, **Xiang Ren**.
Conference on Empirical Methods in Natural Language Processing **EMNLP**), 2019 (ac rate: 24.6%)

[50] Reporting the Unreported: Event Extraction for Analyzing the Local Representation of Hate Crimes
A. M. Davani, L. Yeh, M. Atari, B. Kennedy, G. P. Wightman, E. Gonzalez, N. Delong, R. Bhatia, A. Mirinjian, X. Ren and M. Dehghani.
Conference on Empirical Methods in Natural Language Processing **EMNLP**), 2019 (ac rate: 24.6%)

[49] Eliciting Knowledge from Experts: Automatic Transcript Parsing for Cognitive Task Analysis
J. Du, H. Jiang, J. Shen, X. Ren.
Annual Meeting of the Association for Computational Linguistics (**ACL**), 2019

[48] Improving Distantly-supervised Entity Typing with Compact Latent Space Clustering
B. Chen, X. Gu, Y. Hu, S. Tang, Y. Zhuang, X. Ren.
2019 Conference of the North American Chapter of the Association for Computational Linguistics (**NAACL**), 2019

[47] Posterior-regularized REINFORCE for Instance Selection in Distant Supervision
Q. Zhang, S. Tang, F. Wu, S. Pu, Y. Zhuang, X. Ren.
2019 Conference of the North American Chapter of the Association for Computational Linguistics (**NAACL**), 2019

[46] Characterizing and Forecasting User Engagement with In-App Action Graphs: A Case Study of Snapchat
Y. Liu, X. Shi, L. Pierce, X. Ren.
ACM SIGKDD International Conference on Knowledge (**KDD**), 2019

[45] AtSNE: Efficient and Robust Visualization on GPU through Hierarchical Optimization
C. Fu, Y. Zhang, D. Cai, X. Ren.
ACM SIGKDD International Conference on Knowledge (**KDD**), 2019

[44] Mining Entity Synonyms with Efficient Neural Set Generation
J. Shen, R. Lyu, X. Ren, M. Vanni, B. Sadler, J. Han.
The AAAI Conference on Artificial Intelligence (**AAAI**), 2019

[43] Cross-relation Cross-bag Attention for Distantly-supervised Relation Extraction
Y. Yuan, L. Liu, S. Tang, Z. Zhang, Y. Zhuang, F. Wu, X. Ren.
The AAAI Conference on Artificial Intelligence (**AAAI**), 2019

[42] Learning Dual Retrieval Module for Semi-supervised Relation Extraction
Hongtao Lin, Jun Yan, Qu Meng, **Xiang Ren**.
The Web Conference (**WWW**), 2019

[41] Jointly Learning Explainable Rules for Recommendation with Knowledge Graph
Weizhi Ma, Min Zhang, Yue Cao, Woojeong Jin, Chenyang Wang, Yiqun Liu, Shaoping Ma, **Xiang Ren**.
The Web Conference (**WWW**), 2019

[40] Integrating Local Context and Global Cohesiveness for Open Information Extraction
Qi Zhu, **Xiang Ren**, Jingbo Shang, Yu Zhang, Ahmed El-Kishky, Jiawei Han.
ACM International Conference on Web Search and Data Mining (**WSDM**), 2019

[39] Open-Schema Event Profiling for Massive News Corpora
Quan Yuan, **Xiang Ren**, Wenqi He, Chao Zhang, Xinhe Geng, Lifu Huang, Heng Ji, Chin-Yew Lin, Jiawei Han.

ACM Conference on Information and Knowledge Management (**CIKM**), 2018

[38] Hierarchical Graph Representation Learning with Differentiable Pooling
Rex Ying, Jiaxuan You, Chris Morris, **Xiang Ren**, William L. Hamilton, Jure Leskovec
Annual Conference on Neural Information Processing Systems (**NIPS**), 2018

[37] Efficient Contextualized Representation: Language Model Pruning for Sequence Labeling
Lucas Liu, **Xiang Ren**, Jingbo Shang, Jian Peng, Jiawei Han
Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2018

[36] Learning Named Entity Tagger using Domain-Specific Dictionary
Jingbo Shang, Lucas liu, Xiaotao Gu, **Xiang Ren**, Teng Ren, Jiawei Han
Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2018

[35] GraphRNN: A Deep Generative Model for Graphs
Jiaxuan You, Rex Ying, **Xiang Ren**, William L. Hamilton, Jure Leskovec
International Conference on Machine Learning (**ICML**), 2018 (**Spotlight**)

[34] End-to-End Reinforcement Learning for Automatic Taxonomy Induction
Yuning Mao, **Xiang Ren**, Jiaming Shen, Xiaotao Gu, Jiawei Han
Annual Meeting of the Association for Computational Linguistics (**ACL**), 2018

[33] HiExpan: Task-Guided Taxonomy Construction by Hierarchical Tree Expansion
J. Shen, Z. Wu, D. Lei, C. Zhang, X. Ren, M. T. Vanni, B. M. Sadler, J. Han.
ACM SIGKDD International Conference on Knowledge (**KDD**), 2018

[32] Weakly-supervised Relation Extraction by Pattern-enhanced Embedding Learning
Meng Qu, **Xiang Ren**, Yu Zhang, Jiawei Han
The Web Conference (**WWW**), 2018

[31] Open Information Extraction with Global Structure Constraints
Qi Zhu, **Xiang Ren**, Jingbo Shang, Yu Zhang, Frank Xu, Jiawei Han
The Web Conference (**WWW**), poster paper, 2018
Runner-up of Best Poster Award

[30] Indirect Supervision for Relation Extraction using Question-Answer Pairs
Ellen Wu, **Xiang Ren**, Frank Xu, Ji Li, Jiawei Han
ACM International Conference on Web Search and Data Mining (**WSDM**), 2018

[29] Empowering Sequence Labeling with Task-Aware Neural Language Model
L. Liu, J. Shang, X. Ren, F. Xu, H. Gui, J. Peng, J. Han.
The AAAI Conference on Artificial Intelligence (**AAAI**), 2018
Open-sourced software received over 350 stars on Github as of May 2018

[28] Dynamic Network Embedding by Modeling Triadic Closure Process
Lekui Zhou, Yang Yang, **Xiang Ren**, Fei Wu, Yueting Zhuang
The AAAI Conference on Artificial Intelligence (**AAAI**), 2018

[27] Heterogeneous Supervision for Relation Extraction: A Representation Learning Approach
L. Liu, X. Ren, Q. Zhu, S. Zhi, H. Gui, H. Ji, J. Han.
2017 Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2017

[26] An Attention-based Collaboration Framework for Multi-View Network Representation Learning
M. Qu, J. Tang, J. Shang, X. Ren, M. Zhang, J. Han
26th ACM International on Conference on Information and Knowledge Management (**CIKM**), 2017

[25] Automatic Synonym Discovery with Knowledge Bases
M. Qu, X. Ren, J. Han
23nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (**KDD**), 2017

[24] MetaPAD: Meta Pattern-driven Attribute Discovery in Massive Text Corpora
M. Jiang, J. Shang, T. Cassidy, X. Ren, L. M. Kaplan, T. P. Hanratty, J. Han
23nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (**KDD**), 2017

[23] CoType: Joint Extraction of Typed Entities and Relations with Knowledge Bases
X. Ren, Z. Qiu, W. He, M. Qu, C. R. Voss, H. Ji, T .F.Abdelzaher, J. Han
26th International Conference on World Wide Web (**WWW**), 2017.

[22] Comparative Document Analysis for Large Text Corpora
X. Ren, Y. Lv, K. Wang, J. Han.
10th ACM International Conference on Web Search and Data Mining (**WSDM**), 2017

[21] AFET: Automatic Fine-Grained Entity Typing by Hierarchical Partial-Label Embedding
X. Ren, W. He, M. Qu, H. Ji, C. R. Voss, J. Han.

2016 Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2016, Pages 1369-1378

[20] Label Noise Reduction in Entity Typing by Heterogeneous Partial-Label Embedding
X. Ren, W. He, M. Qu, H. Ji, C. R. Voss, J. Han.
22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (**KDD**), 2016, Pages 1825-1834

[19] FacetGist: Collective Extraction of Document Facets in Large Technical Corpora
T. A. Siddiqui*, X. Ren*, A. Parameswaran, J. Han
25th ACM International on Conference on Information and Knowledge Management (**CIKM**), 2016, Pages 871-880

[18] Representing Documents via Latent Keyphrase Inference
J. Liu, X. Ren, J. Shang, T. Cassidy, C. R. Voss, J. Han
25th International Conference on World Wide Web (**WWW**), 2016, Pages 1057-1067

[17] Precision Matrix Estimation in High Dimensional Gaussian Graphical Models with Faster Rates
L. Wang, X. Ren, Q. Gu
19th International Conference on Artificial Intelligence and Statistics (**AISTATS**), 2016, Pages 177–185

[16] ClusType: Effective Entity Recognition and Typing by Relation Phrase-Based Clustering
X. Ren, A. El-Kishky, C. Wang, F. Tao, C. R. Voss, H. Ji, J. Han
21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (**KDD**), 2015, Pages 995-1004

[15] Synonym Discovery for Structured Entities on Heterogeneous Graphs
X. Ren, and T. Cheng
24th International Conference on World Wide Web (**WWW**), 2015, Pages 443-453

[14] Mining Quality Phrases from Massive Text Corpora
J. Liu, J. Shang, C. Wang, X. Ren, J. Han
2015 ACM SIGMOD International Conference on Management of Data (**SIGMOD**), 2015, Pages 1729-1744
Won grand prize of Yelp Data Challenges in 2015.

[13] ClusCite: Effective Citation Recommendation by Information Network-Based Clustering
X. Ren, J. Liu, X. Yu, U. Khandelwal, Q. Gu, L. Wang, J. Han
20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (**KDD**), 2014, Pages 821-830

[12] Heterogeneous Graph-Based Intent Learning from Queries, Web Pages and Wikipedia Concepts
X. Ren, Y. Wang, X. Yu, J. Yan, Z. Chen, J. Han
7th ACM International Conference on Web Search and Data Mining (**WSDM**), 2014, Pages 23-32

[11] Personalized Entity Recommendation: A Heterogeneous Information Network Approach
X. Yu, X. Ren, Y. Sun, Q. Gu, B. Sturt, U. Khandelwal, B. Norick, J. Han
7th ACM International Conference on Web Search and Data Mining (**WSDM**), 2014, Pages 283-292
Most cited paper in WSDM 2014

[10] Automatic Construction and Ranking of Topical Keyphrases on Collections of Short Documents
M. Danilevsky, C. Wang, N. Desai, X. Ren, J. Guo, J. Han
2014 SIAM International Conference on Data Mining (**SDM**), 2014, Pages 398-406

[9] Semantic Frame-Based Document Representation for Comparable Corpora
H. Kim, X. Ren, Y. Sun, C. Wang, J. Han
IEEE 13th International Conference on Data Mining (**ICDM**), 2013, Pages 350-359
(Acceptance rate: 94/809 = 11.6%)

[8] HeteRec: Entity Recommendation in Heterogeneous Information Networks with Implicit User Feedback
X. Yu, X. Ren, Y. Sun, B. Sturt, U. Khandelwal, Q. Gu, B. Norick, J. Han
7th ACM Recommender Systems Conference (**RecSys**), 2013, Pages 347-350

[7] Repairing Sparse Low Rank Texture
X. Liang, X. Ren, Z. Zhang, Y. Ma
14th European Conference on Computer Vision (**ECCV**), 2012, Pages 482-495

[6] A Non-convex Relaxation Approach to Sparse Dictionary Learning
J. Shi*, X. Ren*, G. Dai, J. Wang, Z. Zhang
2011 IEEE Conference on Computer Vision and Pattern Recognition (**CVPR**), 2011, Pages 1809-1816

Refereed Journal Articles

[5] Automated Phrase Mining from Massive Text Corpora
J. Shang, J. Liu, M. Jiang, X. Ren, J. Han
IEEE Transactions on Knowledge and Data Engineering (**TKDE**), 2018 (**IF: 8.78**)

[4] Linearized Alternating Direction Method with Adaptive Penalty and Warm Starts for Fast Solving Transform Invariant Low-Rank Textures
X. Ren and Z. Lin
International Journal of Computer Vision (**IJCV**), Volume 104 Issue 1, August 2013, Pages 1-14 (**IF: 13.21**)

Other scholarly publications

[3] LEAN-LIFE: A Label-Efficient Annotation Framework Towards Learning from Annotator Explanation
D. Lee, R. Khanna, B. Lin, S. Lee, Q. Ye, E. Boschee, L. Neves, X. Ren
2020 Annual Meeting of the Association for Computational Linguistics (**ACL**) Demo track, 2020

[2] LifeNet: A Structured Network-Based Knowledge Exploration and Analytics System for Life Sciences
X. Ren, J. Shen, M. Qu, X. Wang, Z. Wu, Q. Zhu, M. Jiang, F. Tao, S. Sinha, D. Liem, P. Ping, R. Weinshilboum, J. Han
2017 Annual Meeting of the Association for Computational Linguistics (**ACL**) Demo track, 2017

[1] Cross-media Event Extraction and Recommendation
D. Lu, C. R. Voss, F. Tao, X. Ren, R. Guan, R. Korolov, T. Zhang, D. Wang, H. Li, T. Cassidy, H. Ji, S, Chang, J. Han, W. Wallace, J. Hendler, M. Si, L. Kaplan
2016 Conference of the North American Chapter of the Association for Computational Linguistics (**NAACL**) Demo track, 2016