

National Scholarship (top 2%), USTC 2007, 2008

Nominated by CSE (only 1 or 2 nominees) for:

Early Career Distinguished Scholar (University-Wide), OSU 2021
Sloan Fellowship 2021
BEWEL Leadership in Innovation Award 2021

Grants

Summary: ~2.84M as PI and ~23.59M in total.

- [1] **Huan Sun (Sole PI)**. [Amazon Alexa Prize Taskbot Challenge 2](#), 01/2023-09/2023. **\$250,000**. Only 10 teams across the world were selected.
- [2] **Huan Sun (Sole PI)**. Google Research Scholar Award, “Empowering Pre-trained Language Models for Multi-step Reasoning,” 04/2022. **\$60,000**.
- [3] **Huan Sun (PI)**. Cisco, “Building Transparent Interactive Natural Language Interfaces to Databases,” 04/2022. **\$200,000**. Co-PIs: Yu Su and Michael White.
- [4] **Huan Sun (Sole PI)**. [Amazon Alexa Prize Taskbot Challenge 1](#), 06/2021-05/2022. **\$250,000**.
- [5] **Huan Sun (Senior Personnel)**. NSF AI Institute, “[ICICLE: Intelligent CyberInfrastructure with Computational Learning in the Environment](#),” 2021-2026. Around **\$20 Million** in total (PI: Panda Dhabaleswar, OSU). Contribution: NLP, knowledge graphs, conversational AI (~1 RA support).
- [6] **Huan Sun (Senior Personnel)**. NSF Convergence Accelerator Track F, “Actionable Sensemaking Tools for Curating and Authenticating Information in the Presence of Misinformation during Crises,” 2021-2022. **\$749,999** in total (PI: Srinivasan Parthasarathy, OSU). Contribution: NLP, knowledge graphs.
- [7] **Huan Sun (Co-PI)**. President’s Research Excellence Grant, 2021-2022. **\$50,000**. In collaboration with Dr. Xiaoxue Wang (PI) and Dr. Xia Ning (Co-PI).
- [8] **Huan Sun (Sole PI)**. National Science Foundation (NSF), “**CAREER**: Towards Interactive and Transparent Question Answering with Applications in the Clinical Domain,” 06/2020-05/2025. **\$499,965**.
- [9] **Huan Sun (Sole PI)**. Google Faculty Award, “The Interplay between Web Tables and Texts/KBs for Knowledge Extraction and Embedding,” 03/2020. **\$48,485**.
- [10] **Huan Sun (Sole PI)**. NSF, “III: Small: Towards Resolving Ad-hoc Concept Queries with Table Answers via Multi-source Data Mining,” 08/2018-07/2022. **\$499,001**.
- [11] **Huan Sun (Site PI)**. Patient-Centered Outcomes Research Institute (PCORI), “Unlocking Clinical Text in EMR by Query Refinement Using Both Knowledge Bases and Word Embedding,” 01/2018-03/2022. **\$1,060,000 (total); \$483,001 (my portion)**. In collaboration with Nationwide Children’s Hospital (PI: Dr. Simon Lin, Primary Contact).
- [12] **Huan Sun (Sole PI)**. Army Research Office (ARO), “Advancing Human and Machine Question Answering via Human-Machine Collaboration,” 09/2017-04/2021. **\$498,526**.

- [13] **Huan Sun (Sole PI)**. Fujitsu Lab of America, “Information Extraction for IoT Device APIs Based on Developer Documentation,” 05/2017. **\$50,000**.

Publications

(*: My advisee;
#: visiting/partly
funded student;
+: equal
contribution)

At OSU:

- [1] Zhen Wang*, Rameswar Panda, Leonid Karlinsky, Rogerio Feris, Huan Sun, Yoon Kim, “Multitask Prompt Tuning Enables Parameter-Efficient Transfer Learning,” *The 11th International Conference on Learning Representations (ICLR’23)*.
- [2] Boshi Wang*, Xiang Deng*, Huan Sun, “Iteratively Prompt Pre-trained Language Models for Chain of Thought,” *The 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP’22, long)*.
- [3] Bernal Jimnez Gutierrez, Nikolas McNeal*, Clay Washington, You Chen, Lang Li, Huan Sun, Yu Su, “Thinking about GPT-3 In-Context Learning for Biomedical IE? Think Again,” *Findings of the Conference on Empirical Methods in Natural Language Processing, 2022 (Findings of EMNLP’22, long)*.
- [4] Shijie Chen, Ziru Chen*, Huan Sun, Yu Su, “Error Detection for Interactive Text-to-SQL Semantic Parsing,” *The 2nd Workshop on Interactive Learning for Natural Language Processing at NeurIPS 2022 (InterNLP’22)*.
- [5] Shijie Chen⁺, Ziru Chen^{*,+}, Xiang Deng*, Ashley Lewis, Lingbo Mo*, Samuel Stevens, Zhen Wang*, Xiang Yue*, Tianshu Zhang*, Yu Su, Huan Sun, “Bootstrapping a User-Centered Task-Oriented Dialogue System,” *The 1st Proceedings of Alexa Prize TaskBot, 2022*. **Third-Place Winner. TaskBot deployed in Amazon Alexa.**
- [6] Lingbo Mo^{*,+}, Zhen Wang^{*,+}, Jie Zhao, Huan Sun, “Knowledge Transfer between Structured and Unstructured Sources for Complex Question Answering,” *The Workshop on Structured and Unstructured Knowledge Integration (SUKI) at NAACL 2022*.
- [7] Xiang Yue*, Ziyu Yao, Huan Sun, “Synthetic Question Value Estimation for Domain Adaptation of Question Answering,” *The 60th Annual Meeting of the Association for Computational Linguistics (ACL’22, long)*.
- [8] Lingbo Mo*, Ashley Lewis, Huan Sun, Michael White, “Towards Transparent Interactive Semantic Parsing via Step-by-Step Correction,” *Findings of the 60th Annual Meeting of the Association for Computational Linguistics (Findings of ACL’22, long)*.
- [9] Xiang Yue^{*,+}, Xinliang Frederick Zhang^{*,+}, Ziyu Yao*, Simon Lin, **Huan Sun**, “CliniQG4QA: Generating Diverse Questions for Domain Adaptation of Clinical Question Answering,” *The 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM’21, long)*. **Best Paper Award.**
- [10] Xiang Deng*, Yu Su, Alyssa Lees, You Wu, Cong Yu, **Huan Sun**, “ReasonBERT: Pre-trained to Reason with Distant Supervision,” *The 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP’21, long)*.
- [11] Xinliang Frederick Zhang*, Heming Sun*, Xiang Yue*, Simon Lin, **Huan Sun**, “COUGH: A Challenge Dataset and Models for COVID-19 FAQ Retrieval,” *The 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP’21, short)*.
- [12] Boyuan Pan^{#,+}, Yazheng Yang⁺, Cai Deng, **Huan Sun**, “TopNet: Learning from Neural Topic Model to Generate Long Stories,” *The 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD’21, research track, acceptance rate: ~15.4%)*

- [13] Xiang Yue^{*,+}, Minxin Du⁺, Tianhao Wang, Yaliang Li, **Huan Sun**, Sherman S. M. Chow, “Differential Privacy for Text Analytics via Natural Text Sanitization,” *Findings of the Joint Conference of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Findings of ACL-IJCNLP’21, long)*
- [14] Xiang Deng^{*}, Ahmed Hassan Awadallah, Christopher Meek, Oleksandr Polozov, **Huan Sun**, Matthew Richardson, “Structure-Grounded Pretraining for Text-to-SQL,” *The 2021 Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT’21, long)*.
- [15] Ziyu Yao^{*}, Frank F. Xu, Pengcheng Yin, **Huan Sun**, Graham Neubig, “Learning Structural Edits via Incremental Tree Transformations,” *The Ninth International Conference on Learning Representations 2021 (ICLR’21)*.
- [16] Xiang Deng^{*}, **Huan Sun**, Alyssa Lees, You Wu, Cong Yu, “TURL: Table Understanding through Representation Learning,” *47th International Conference on Very Large Data Bases (VLDB’21)*. **2022 ACM SIGMOD Research Highlight Award**.
- [17] Soheil Moosavinasab, Emre Sezgin, **Huan Sun**, Jeffrey Hoffman, Yungui Huang, Simon Lin, “DeepSuggest: Using Neural Networks to Suggest Related Keywords for a Comprehensive Search of Clinical Notes,” *ACI Open 2021*.
- [18] Zhen Wang^{*}, Bo Zong, **Huan Sun**, “Modeling Context Pair Interaction for Pairwise Tasks on Graphs,” *The 14th International Conference on Web Search and Data Mining (WSDM’21, acceptance rate: 18.6%)*.
- [19] Kaushik Mani^{*,+}, Xiang Yue^{*,+}, Bernal Jimenez Gutierrez, Yungui Huang, Simon Lin, **Huan Sun**, “Clinical Phrase Mining with Language Models,” *IEEE International Conference on Bioinformatics and Biomedicine 2020 (BIBM’20, short)*.
- [20] Ziyu Yao^{*}, Yiqi Tang, Wen-tau Yih, **Huan Sun**, Yu Su, “An Imitation Game for Learning Semantic Parsers from User Interaction,” *2020 Conference on Empirical Methods in Natural Language Processing (EMNLP’20, long)*.
- [21] Bernhard Kratzwald[#], Stefan Feuerriegel, **Huan Sun**, “Learning a Cost-Effective Annotation Policy for Question Answering,” *2020 Conference on Empirical Methods in Natural Language Processing (EMNLP’20, long)*.
- [22] Jie Zhao^{*}, **Huan Sun**, “Adversarial Training for Code Retrieval with Question-Description Relevance Regularization,” *Findings of 2020 Conference on Empirical Methods in Natural Language Processing (Findings of EMNLP’20, long)*[A new acceptance category of EMNLP; review scores: 4/4/4 with 5 being the highest].
- [23] Zhen Wang^{*}, Jennifer Lee, Simon Lin, **Huan Sun**, “Rationalizing Medical Relation Prediction from Corpus-level Statistics,” *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL’20, long)*.
- [24] Xiang Yue^{*}, Bernal Jimenez^{*}, **Huan Sun**, “Clinical Reading Comprehension: A Thorough Analysis of the emrQA Dataset,” *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL’20, long)*.
- [25] Jiankai Sun, Jie Zhao^{*}, **Huan Sun**, Srinivasan Parthasarathy, “EndCold: An End-to-End Framework for Cold Question Routing in Community Question Answering Services,” *The 29th International Joint Conference on Artificial Intelligence (IJCAI’20, acceptance rate: 12.6%)*.
- [26] Xiang Yue^{*}, Zhen Wang^{*}, Jingong Huang^{*}, Srinivasan Parthasarathy, Soheil Moosavinasab, Yungui Huang, Simon M. Lin, Wen Zhang, Ping Zhang, and **Huan Sun**, “Graph Embedding on Biomedical Networks: Methods, Applications, and Evaluations,” *Bioinformatics, Vol 36 Issue 4, 15 Feb 2020, Page 1241-1251*. (ESI Highly Cited Paper: among the **top 1%** highly cited papers of its academic field)

- [27] Ziyu Yao*, Yu Su, **Huan Sun**, Wen-tau Yih, “Model-based Interactive Semantic Parsing: A Unified Formulation and A Text-to-SQL Case Study,” *The 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP’19, long)*.
- [28] Xiang Deng*, **Huan Sun**, “Leveraging 2-hop Distant Supervision from Table Entity Pairs for Relation Extraction,” *The 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP’19, long)*.
- [29] Zhen Wang*, Xiang Yue*, Seyed Soheil Moosavinasab, Yungui Huang, Simon Lin, **Huan Sun**, “SurfCon: Synonym Discovery on Privacy-Aware Clinical Data,” *The 25th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD’19, acceptance rate: 14.2%)*.
- [30] Jie Zhao*, Ziyu Guan, **Huan Sun**, “Riker: Mining Rich Keyword Representations towards Explainable Product Question Answering,” *The 25th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD’19, acceptance rate: 14.2%)*.
- [31] Boyuan Pan#, Hao Li, Ziyu Yao*, Deng Cai, **Huan Sun**, “Reinforced Dynamic Reasoning for Conversational Question Generation,” *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (ACL’19, long)*.
- [32] Ziyu Yao*, Jayavardhan Reddy Peddamail*, **Huan Sun**, “CoaCor: Code Annotation for Code Retrieval with Reinforcement Learning,” *The Web Conference (former WWW Conference) 2019 (acceptance rate: 18%)*.
- [33] Ziyu Yao*, Xiujun Li, Jianfeng Gao, Brian Sadler, **Huan Sun**, “Interactive Semantic Parsing for If-Then Recipes via Hierarchical Reinforcement Learning,” *The AAAI Conference on Artificial Intelligence (AAAI’19, acceptance rate: 16.2%)*.
- [34] Teng Xiao, Jiaxin Ren, Zaiqiao Meng, **Huan Sun**, Shangsong Liang, “Dynamic Bayesian Metric Learning for Personalized Product Search,” *Proceedings of the 28th ACM International Conference on Information and Knowledge Management (CIKM’19)*.
- [35] Long Chen, Ziyu Guan, Wei Zhao, Wanqing Zhao, Xiaopeng Wang, Zhou Zhao, **Huan Sun**, “Answer Identification from Product Reviews for User Questions by Multi-task Attentive Networks,” *The AAAI Conference on Artificial Intelligence (AAAI’19, acceptance rate: 16.2%)*.
- [36] Wei Zhao, Ziyu Guan, Yuhui Huang, Tingting Xi, **Huan Sun**, Zhiheng Wang and Xiaofei He, “Discerning Influence Patterns with Beta-Poisson Factorization in Microblogging Environments,” *IEEE Transactions on Knowledge and Data Engineering (TKDE) 2019*.
- [37] Jayavardhan Reddy Peddamail*, Ziyu Yao*, Zhen Wang*, **Huan Sun**, “A Comprehensive Study of StaQC for Deep Code Summarization,” *SIGKDD 2018 Deep Learning Day*.
- [38] Ziyu Yao*, Daniel S. Weld, Wei-Peng Chen, **Huan Sun**, “StaQC: A Systematically Mined Question-Code Dataset from Stack Overflow,” *The Web Conference (former WWW Conference) 2018*.
- [39] Yu Su⁺, Honglei Liu⁺, Semih Yavuz, Izzeddin Gur, Huan Sun, Xifeng Yan. “Global Relation Embedding for Relation Extraction,” *The Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT’18, long)*.
- [40] Jie Zhao*, Yu Su, Ziyu Guan, **Huan Sun**, “An End-to-End Deep Framework for Answer Triggering with a Novel Group-Level Objective,” *The 2017 Conference on Empirical Methods in Natural Language Processing (EMNLP’17, short)*.

- [41] Yaliang Li, Nan Du, Chaochun Liu, Yusheng Xie, Wei Fan, Qi Li, Jing Gao, **Huan Sun**, “Reliable Medical Diagnosis from Crowdsourcing: Discover Trustworthy Answers from Non-Experts,” *The 10th ACM International Conference on Web Search and Data Mining (WSDM’17)*.
- [42] Yu Su, **Huan Sun**, Brian Sadler, Mudhakar Srivatsa, Izzeddin Gur, Zenghui Yan, Xifeng Yan, “On Generating Characteristic-rich Question Sets for QA Evaluation,” *The 2016 Conference on Empirical Methods in Natural Language Processing (EMNLP’16, long)*.
- [43] Chaochun Liu, **Huan Sun**, Nan Du, Shulong Tan, Hongliang Fei, Wei Fan, Tao Yang, Hao Wu, Yaliang Li, Chenwei Zhang, “Augmented LSTM Framework to Construct Medical Self-diagnosis Android,” *IEEE 16th International Conference on Data Mining (ICDM’16)*.

Before joining OSU:

- [44] **Huan Sun**, Hao Ma, Xiaodong He, Wen-tau Yih, Yu Su, Xifeng Yan, “Table Cell Search for Question Answering,” *The 25th International Conference on World Wide Web (WWW) 2016*.
- [45] Yang Li, Shulong Tan, **Huan Sun**, Jiawei Han, Dan Roth, Xifeng Yan, “Entity Disambiguation with Linkless Knowledge Bases,” *The 25th International Conference on World Wide Web (WWW) 2016*.
- [46] Fangqiu Han, Shulong Tan, **Huan Sun**, Xifeng Yan, Mudhakar Srivatsa, Deng Cai, “Distributed Representations of Expertise,” *SIAM International Conference on Data Mining (SDM), 2016*.
- [47] **Huan Sun**, Hao Ma, Wen-tau Yih, Chen-Tse Tsai, Jingjing Liu, Ming-Wei Chang, “Open Domain Question Answering via Semantic Enrichment,” *The 24th International Conference on World Wide Web (WWW) 2015*.
- [48] Yu Su, Shengqi Yang, **Huan Sun**, Mudhakar Srivatsa, Sue Kase, Michelle Vanni, Xifeng Yan, “Exploiting Relevance Feedback in Knowledge Graph Search,” *The 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD) 2015*.
- [49] Ziyu Guan, Shengqi Yang, **Huan Sun**, Mudhakar Srivatsa, Xifeng Yan, “Fine-Grained Knowledge Sharing in Collaborative Environments,” *IEEE Transactions on Knowledge and Data Engineering (TKDE) 2015*.
- [50] **Huan Sun**, Mudhakar Srivatsa, Shulong Tan, Yang Li, Lance Kaplan, Shu Tao, Xifeng Yan, “Analyzing Expert Behaviors in Collaborative Networks,” *The 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD) 2014*.
- [51] Feida Zhu, **Huan Sun**, Xifeng Yan, “Network Mining and Analysis for Social Applications,” *The 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD) 2014*, **Tutorial**.
- [52] Shengqi Yang, Yinghui Wu, **Huan Sun**, Xifeng Yan, “Schemaless and Structureless Graph Querying,” *The 40th International Conference on Very Large Data Bases (VLDB) 2014*.
- [53] Shengqi Yang, Yanan Xie, Yinghui Wu, Tianyu Wu, **Huan Sun**, Jian Wu, Xifeng Yan, “SLQ: A User-friendly Graph Querying System,” *ACM International Conference on Management of Data (SIGMOD) 2014, Demo*.
- [54] **Huan Sun**, Mudhakar Srivatsa, Lance Kaplan, Xifeng Yan, “Analyzing Expert Behaviors in Collaborative Networks,” *International School and Conference on Network Science (NetSci) 2014*.

- [55] Nan Li, **Huan Sun**, Kyle Chipman, Jemin George, Xifeng Yan, “A Probabilistic Approach to Uncovering Attributed Graph Anomalies,” *SIAM International Conference on Data Mining (SDM) 2014*.
- [56] **Huan Sun**, Alex Morales, Xifeng Yan, “Synthetic Review Spamming and Immunization,” *The 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD) 2013*. A shorter version appeared in *the 22nd International Conference on World Wide Web 2013 (WWW’13, poster)*
- [57] Shulong Tan, Yang Li, **Huan Sun**, Ziyu Guan, Xifeng Yan, Jiajun Bu, Chun Chen, Xiaofei He, “Interpreting the Public Sentiment Variations on Twitter,” *IEEE Transactions on Knowledge and Data Engineering (TKDE) 2014*.
- [58] **Huan Sun**, Gengxin Miao, Xifeng Yan, “Noise-Resistant Bicluster Recognition,” *IEEE 13th International Conference on Data Mining (ICDM) 2013*. Also appeared in *the 17th Annual International Conference on Research in Computational Molecular Biology 2013 (RECOMB’13, poster)*.

Students Supervised/Mentored

Students from underrepresented groups I have worked with:

- Female: Ziyu Yao (Ph.D., 2016-2021), Tianshu Zhang (Ph.D., since 2021), Ash Lewis (Ph.D. in Linguistics, since 2020), Alyssa Allen (Ph.D. in Linguistics, since 2022), Leslie Zhou (Master, 2020-2021), Momo Liu (Undergrad, 2016-2017), Yu Huo (Undergrad, since 2022)
- African American: Jaylen Jones (Ph.D., since 2022)
- Hispanic: Alex Morales (Master, 2012-2013)

Ph.D. Alumni:

- Ziyu Yao, Ph.D., AU2016-SU2021.
 - Awards: Presidential Fellowship (the most prestigious award given by the Graduate School at OSU), Graduate Student Research Award, Rising Stars in EECS (by UC Berkeley).
 - “On Advancing Natural Language Interfaces: Data Collection, Model Development, and User Interaction,” *Dissertation*.
 - Now at George Mason as a tenure-track assistant professor. She also received tenure-track offers in CS from University of Arizona and Emory University.
- Jie Zhao, Ph.D., AU2016-AU2020.
 - “Towards Building Trustworthy Automatic Question Answering Systems,” *Dissertation*.
 - First: Applied Scientist at Amazon; Now: Senior Applied Scientist at Microsoft.
- Zhen Wang, AU2017-AU2022.
 - Awards: Rising Stars in Data Science (by UChicago), Graduate Research Award 2022.
 - “Knowledge-Centric Natural Language Processing: Acquisition, Representation, Transfer, and Reasoning,” *Dissertation*.
 - To be a postdoc jointly hosted by UC San Diego/MBZUAI/CMU. He also got an offer from Johns Hopkins University.

Current Ph.D. Students (only those graduating are listed):

- Xiang Deng, AU2018–SU2023 (expected).
 - Awards: 2022 SIGMOD Research Highlight Award, Presidential Fellowship 2022

- First-author papers published in EMNLP’21, VLDB’21, NAACL’21, EMNLP’19.
- Summer Internships at Google Research, Microsoft Research (Redmond) and Amazon.
- Xiang Yue, AU2018-SU2023 (expected).
 - Awards: 2021 BIBM Best Paper Award
 - First-author papers published in ACL’22, ACL’21, ACL’20, Bioinformatics 2020 (ESI Highly Cited Paper: top 1% cited paper of its academic field).
 - Summer Internships at Microsoft Research (Redmond), Tencent AI.

Undergraduate & Master Alumni: (Only those who closely worked with me on research projects are listed.)

- Ziru “Ron” Chen, B.S., 2022.
 - Awards: University Fellowship for Ph.D. Study at OSU, Undergraduate Research Award
 - The only undergraduate student and one of the two student leaders on the award-winning OSU TacoBot Team (with 8 Ph.D. students and 9 students in total) for the Alexa TaskBot Challenge.
- Xinliang “Frederick” Zhang, B.S., 2021, joined UMichigan (Ph.D. program in EECS).
 - Awards: Undergraduate Research Award, Honorable Mention of CRA Outstanding Undergraduate Researcher Award
 - First-author paper (short) published in EMNLP’21.
- Heming Sun, B.S., 2021, joined Univ. of Southern California (M.S. program in CS).
- Kaushik Mani, M.S., 2020, now at Oracle.
 - One short paper published in BIBM’20.
- Jayavardhan Reddy Peddamail, M.S., 2019, now at Google.
 - One first-author paper in KDD 2018 Deep Learning Day; One second-author paper in WWW 2019.
- Dushyanta Dhyani, M.S., 2018, now at Amazon.
 - Thesis: “Boosting Supervised Neural Relation Extraction with Distant Supervision”

Supported (Visiting) Students:

- Bortik Bandyopadhyay, Autumn 2018, Spring 2019, Advisor: Prof. Srinivasan Parthasarathy.
 - Now: Machine Learning Engineer at Apple
- Boyuan Pan, Nov 2018 - May 2019.
 - Papers published in SIGKDD 2021, ACL 2019.
- Yuantong Li, Sep 2018 - Jan 2019, now at UCLA.

Before joining OSU:

- Alex Morales, Master student, 2012-2013, UCSB. Collaborator of my work published in SIGKDD 2013, “Synthetic Review Spamming and Defense.” Went on to pursue a Ph.D. degree at Univ. of Illinois at Urbana Champaign (with 5-year fellowship).
- Sang Nguyen, Undergraduate student from Oxnard College, Summer 2012, UCSB. “Perceptron for Classification: Analysis and Application.”

Teaching Experience

At OSU:

- Introduction to Data Mining (CSE5243, Mixed Undergraduate & Graduate Level)
Instructor, Autumn 2022, OSU.
- On the Opportunities and Risks of Pre-trained Language Models (CSE5539, Graduate Level)
Instructor, Spring 2022, OSU.
- Speech and Language Processing (CSE5525, Mixed Undergrad. & Graduate Level)
Instructor, Autumn 2021, OSU.
- Speech and Language Processing (CSE5525, Mixed Undergrad. & Graduate Level)
Instructor, Autumn 2020, OSU.
- Natural Language Question Answering: Conversational, Interactive, Interpretable (CSE5539, Graduate Level)
Instructor, Spring 2020, OSU.
- Introduction to Data Mining (CSE5243, Mixed Undergraduate & Graduate Level)
Instructor, Autumn 2019, OSU.
- Introduction to Data Mining (CSE5243, Mixed Undergraduate & Graduate Level)
Instructor, Autumn 2018, OSU.
- Human-Machine Collaboration Mechanisms in Data Mining, Database, and AI (CSE5249, Graduate Level)
Instructor, Spring 2018, OSU.
- Introduction to Data Mining (CSE5243, Mixed Undergraduate & Graduate Level)
Instructor, Autumn 2017, OSU.
- Natural Language Question Answering (CSE5539, Graduate Level)
Instructor, Autumn 2016, OSU.

Before joining OSU:

- Artificial Intelligence (CS165A)
Guest lecturer, Winter 2012, UCSB
- Special Topics in CS: Information Management (CS290D)
Guest lecturer, Spring 2014, UCSB

Invited Talks & Tutorials

- TDAI Foundations of Data Science and AI Community of Practice at OSU, 01/2023
- Table Representation Learning Workshop at NeurIPS, 12/2022
- Research Seminar at the Department of Biomedical Informatics at OSU, 09/2022
- Apple (Machine Learning, Apple Knowledge Platform), 09/2022
- Microsoft Semantic Machines, 07/2022
- Deep Learning on Graphs Workshop @ NAACL'22, 07/2022
- Fudan University, 06/2022
- Tutorial@TDAI Summer School, 06/2022
60 in-person & 100 online registered attendees
- Northeastern University (Distinguished Speaker), 04/2022
- Choose Ohio First Scholarship Program, Online, 11/2021
- Tutorial, SIGKDD, Online, 08/2021
- Nesh (startup), Texas AI community, Online, 07/2021
- Allen Institute for AI, Online, 07/2021

- Facebook AI Research, Online, 11/2020
- UIUC, Online, 10/2020
- Rulai (Startup), Online, 06/2020
- Google Research, New York City, 02/2020
- Facebook Conversational AI, Menlo Park, CA, 10/2019
- IEEE National Aerospace & Electronics Conference (NAECON), Dayton, Ohio, 07/2019
- OSU-Tsukuba Joint Linguistics Workshop, Columbus, Ohio, 02/2019
- Army Science Planning and Strategy Meeting, Austin, Texas, 01/2019
- City Brain Project, Hangzhou, China, 07/2018
- Fudan University, 07/2018
- Tsinghua University, 06/2018
- UMass Amherst, 02/2018
- Rice University, 02/2018
- University of Illinois, Chicago, 01/2018
- Penn State University, 01/2018
- Fujitsu Laboratories of America, 06/2017
- University of Science and Technology of China, Hefei, 01/2017
- Baidu Research, Sunnyvale, 09/2015
- University of Washington, 08/2015
- Army Research Lab, Silver Spring, 05/2015
- The Ohio State University, 03/2015
- Simon Fraser University, 03/2015
- ShanghaiTech University, China, 09/2014
- Tutorial @ SIGKDD, 08/2014

**DEI Related
Services and
Outreach
Activities
(Selected)**

- Talk at OSU AI Club (with undergrads and grads), 11/2022
- Diversity advocate, CSE Faculty Search Committee ([RAISE Initiative](#)), 2022-2023
- Guest speaker, CSE 6891: Departmental Research Seminar, 08/2022
- Guest speaker, [Choose Ohio First Scholarship program](#), CSE Overview (CSE 2193), 11/2021
- Attendee to promote diversity at OSU, Tapia 2021, CMDIT, 09/2021
- Faculty mentor, [OSU LSAMP program](#) (which aims to diversify the STEM workforce by increasing underrepresented minorities in STEM degrees), 2020-2021
- Guest speaker, ENGR 1195.02 Women in Engineering (WiE) Seminar Class, 09/2019
- Meeting with female students at Women in Engineering (WiE) Cafe, 08/2019
- Faculty mentor, OSU LSAMP program (which aims to diversify the STEM workforce by increasing underrepresented minorities in STEM degrees), 2019-2020
- Hosting two high school summer interns, 06-08/2018
 - Sounak Dey: Thomas Worthington High School
 - Jacob Mayhew: St. Charles Preparatory School

- Served as Judge in AI Club hackathon at Ohio State, 04/2018
Event theme: Weekend-long challenge on AI related projects
Participants: *Undergraduate/graduate students.*
- Enrolled in the Metro Early College High School Internship program, 10/2017 -
Program goal: *To host high school summer interns.*
- Served as Judge in ASA DataFest at Ohio State, 05/2017
Event theme: Weekend-long challenge on data science related projects
Participants: *Undergraduate students from 5 colleges.*
- Invited speaker at the OSU Innovation Summit, 05/2017
“Mining Big Data to Build Intelligent Assistants Who Can Speak Natural Language”
Event theme: Innovation and disruption in the next digital age
Audience: *Entrepreneurial-minded leaders, academics & students.*
- Invited speaker at High School I/O (the first high school hackathon), 04/2017
“How to Build Intelligent Assistants Who Can Speak Natural Language”
Audience: *High school students from Central Ohio.*
- Attendee of the 2017 Ohio Celebration of Women in Computing (OCWiC), 02/2017
Event theme: *Engaging and empowering Ohio women in computing.*
- Invited speaker at Center of Science and Industry (COSI), 10/2016
“Data Mining and To Be a Data Miner”
Audience: *Underrepresented high school students from Central Ohio.*

Academic Services

Chair Positions:

- 2023: ACL (Area Chair), SIGKDD (Area Chair), AAAI (Area Chair), Workshop on Knowledge Augmented Methods for NLP at AAAI (Co-Chair), SIGKDD Hands-on Tutorial (Co-Chair)
- 2022: Wise-Supervision@AKBC Workshop (Co-Chair), Table Representation Learning Workshop at NeurIPS (Panel Chair)
- 2021: NAACL (Area Chair), NLP for Programming Workshop at ACL (Co-chair), NLPCC (Area Chair), ACL (Birds of a Feather Session Chair)
- 2020: Natural Language Interface Workshop at ACL (Co-chair)
- 2019: SDM (publicity co-chair, session chair), SIGKDD (session chair)

Review/Topic Editor:

- 2022-now: Women in Data Mining and Management (Frontiers in Big Data)
- 2018-now: Frontiers in Big Data (Data Mining and Management Section)

Panelists and Program Committee Member in Recent Years:

- 2022: NSF CAREER Panel, ACL Rolling Review, EMNLP
- 2021: NAACL, EMNLP, ACL Rolling Review
- 2020: SIGKDD, ACL, EMNLP
- 2019: NSF Panel, SDM, NAACL, SIGKDD, EMNLP-MRQA
- 2018: ACL, ICDM, CoNLL
- 2017: NSF Panel, SIGKDD, ACL, EMNLP, WWW, CIKM (senior PC), IJCNLP, IJCAI BOOM workshop
- 2016: SIGKDD, EMNLP, ICDM, CIKM, BigMine, IJCAI BOOM workshop

Journal Reviewer:

- Transactions on the Web, since 2022

- Transactions on Information Systems, since 2021
- Transactions on Audio, Speech and Language Processing, since 2021
- IEEE/ACM Transactions on Computational Biology and Bioinformatics, since 2021
- Transactions of the Association for Computational Linguistics, since 2020
- ACM Transactions on Knowledge Discovery from Data (TKDD), since 2016
- Data Mining and Knowledge Discovery, since 2016
- IEEE Transactions on Knowledge and Data Engineering (TKDE), since 2015