# Weiyan Shi

*New York, NY*
✉ *ws2634@columbia.edu*
🖥 *https://wyshi.github.io/*

━━━━━ Education

2021-Present **Columbia University, New York, NY**
*Ph.D.* in Computer Science, *Advisor*: Zhou Yu
*Research interests*: Natural Language Processing, specifically, Dialogue Systems
**Google Scholar**

2018-2021 **University of California, Davis, CA**
*Ph.D.* in Computer Science, *Advisor*: Zhou Yu

2015-2016 **University of California, Berkeley, CA**
*Master of Arts* in Statistics

2011-2015 **Renmin University of China, Beijing, China**
*Bachelor of Science* in Mathematics and Applied Mathematics

━━━━━ Awards and Honors

2019 **Best Paper Nomination**, ACL 2019
2018-2023 Dean's Distinguished Graduate Fellowship, UC, Davis
2016 Department Citation in the Statistics Master's Program (top 1), UC, Berkeley
2016 MA Speaker at Department Commencement, UC, Berkeley
2013 National Scholarship for Excellent Academic Performance, RUC
2013 Presidential Fellowship for Studying Abroad, RUC

━━━━━ Publications

In submission **Selective Differential Privacy for Language Modeling**
*Weiyan Shi*, Aiqi Cui, Evan Li, Ruoxi Jia, Zhou Yu
arXiv preprint arXiv:2108.12944, in submission

In submission **What Does the Bot Do? Detecting Repetition and Contradiction in Dialogues by Asking and Answering Questions**
*Weiyan Shi*, Mary Williamson, Jason Weston, Mike Lewis, Zhou Yu

In submission **DEUX: An Attribute-Guided Framework for Sociable Recommendation Dialog Systems**
Yu Li, Shirley Anugrah Hayati, *Weiyan Shi*, Zhou Yu
arXiv preprint arXiv:2105.00825, in submission

In submission **Towards Socially Intelligent Agents with Mental State Transition and Human Utility**
Liang Qiu, Yizhou Zhao, Yuan Liang, Pan Lu, *Weiyan Shi*, Zhou Yu, Song-Chun Zhu
arXiv preprint arXiv:2103.07011, in submission

| | |
|---|---|
| EMNLP 2021 Findings | **Refine and Imitate: Reducing Repetition and Inconsistency in Persuasion Dialogues via Reinforcement Learning and Human Demonstration** <br> *Weiyan Shi*, Yu Li, Saurav Sahay, Zhou Yu <br> arXiv preprint arXiv:2012.15375, EMNLP 2021 Findings |
| ACL 2020 Demo | **LEGOEval: An Open-Source Toolkit for Dialogue System Evaluation via Crowdsourcing** <br> Yu Li, Josh Arnold, Feifan Yan, *Weiyan Shi*, Zhou Yu <br> The 2021 Annual Meeting of the Association for Computational Linguistics. |
| EMNLP 2020 | **INSPIRED: Toward Sociable Recommendation Dialog Systems** <br> Shirley Anugrah Hayati, Dongyeop Kang, Qingxiaoyang Zhu, *Weiyan Shi*, Zhou Yu <br> The 2020 Conference on Empirical Methods in Natural Language Processing. |
| EMNLP 2020 | **Structured Attention for Unsupervised Dialogue Structure Induction** <br> Liang Qiu, Yizhou Zhao, *Weiyan Shi*, Yuan Liang, Feng Shi, Tao Yuan, Zhou Yu, Song-Chun Zhu <br> The 2020 Conference on Empirical Methods in Natural Language Processing. |
| EMNLP 2020 Findings | **Understanding User Resistance Strategies in Persuasive Conversations** <br> Youzhi Tian, *Weiyan Shi*, Chen Li, Zhou Yu <br> The 2020 Conference on Empirical Methods in Natural Language Processing. |
| CHI 2020 | **Effects of Persuasive Dialogues: Testing Bot Identities and Inquiry Strategies** <br> *Weiyan Shi*, Xuewei Wang, Yoo Jung Oh, Jingwen Zhang, Saurav Sahay, Zhou Yu <br> Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. |
| AAAI 2020 | **End-to-End Trainable Non-Collaborative Dialog System.** <br> Yu Li, Kun Qian, *Weiyan Shi*, Zhou Yu <br> Proceedings of the 2020 AAAI Conference on Artificial Intelligence. |
| EMNLP-IJCNLP 2019 | **How to Build User Simulators to Train RL-based Dialog Systems** <br> *Weiyan Shi\**, Kun Qian\* (equal contribution), Xuewei Wang, Zhou Yu <br> The 2019 Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing. |
| ACL 2019 | **Persuasion for Good: Towards a Personalized Persuasive Dialogue System for Social Good** <br> Xuewei Wang\*, *Weiyan Shi\** (equal contribution), Richard Kim, Yoojung Oh, Sijia Yang, Jingwen Zhang, Zhou Yu <br> The 57th Annual Meeting of the Association for Computational Linguistics. <br> <span style="color:red">**Best Paper Nomination**</span> (28/2905, 1%). |

| | |
|---|---|
| NAACL 2019 | **Unsupervised Dialog Structure Learning**<br>*Weiyan Shi*, Tiancheng Zhao, Zhou Yu<br>The 17th Annual Conference of the North American Chapter of the Association for Computational Linguistics. |
| ACL 2018 | **Sentiment Adaptive End-to-End Dialog Systems**<br>*Weiyan Shi*, Zhou Yu<br>The 56th Annual Meeting of the Association for Computational Linguistics. |
| RepL4NLP workshop at ACL 2018 | **Exploiting Common Characters in Chinese and Japanese to Learn Cross-Lingual Word Embeddings via Matrix Factorization**<br>Jilei Wang*, Shiying Luo* (equal contribution), *Weiyan Shi*, Tao Dai, Shu-Tao Xia<br>The Third RepL4NLP workshop of the Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics. |

## ▬▬▬ Professional Experience

| | |
|---|---|
| 2021.5–<br>Present | **Facebook AI Research, Seattle, WA (Remote)**<br>*Research Intern*, *Host:* Michael Lewis<br>• Work on zero-shot nonsense detection. |
| 2020.6–<br>2020.12 | **Facebook AI Research, Seattle, WA (Remote)**<br>*Research Intern*, *Host:* Michael Lewis<br>• Worked on strategic dialogue generation. |
| 2016–2018 | **[24]7.ai, San Jose, CA**<br>*Full-time Data Scientist*<br>• Worked on real-world natural-language-related projects in the customer service domain, such as audio data collection, intent classification, dialogue workflow extraction, dialogue intent transition, and chatbot building with deep learning techniques.<br>• Implemented and applied models from academic papers to real-world natural-language data, and presented results to team members during weekly research meetings. |

## ▬▬▬ Service

| | |
|---|---|
| 2021 | Program Committee/Reviewer in ICLR, AAAI, EACL, NAACL, Neurocomputing, ACL, *SEM, NeurIPS, EMNLP |
| 2020 | Program Committee/Reviewer in ACL, THRI, *SEM |
| 2019 | Program Committee/Reviewer in ACL |