

Weiyan Shi (she/her/hers)

CONTACT	ws2634@columbia.edu wyshi.github.io Google Scholar	
RESEARCH INTERESTS	Natural Language Processing (dialogue systems, privacy-preserving NLP models, etc), Machine Learning, Artificial Intelligence	
EDUCATION	Columbia University Ph.D. in Computer Science Advisor: Zhou Yu	2020 - 2023 (anticipated)
	University of California, Davis Ph.D. in Computer Science (transferred with my advisor) Advisor: Zhou Yu	2018 - 2020
	University of California, Berkeley M.A. in Statistics	2015 - 2016
	Renmin University of China B.S. in Mathematics and Applied Mathematics	2011 - 2015
AWARDS & HONORS	Rising Star in Machine Learning, UMD Best Paper Nomination, ACL Dean's Distinguished Ph.D. Fellowship, UC Davis Department Citation (top 1), UC Berkeley Speaker at Department Commencement (top 1), UC Berkeley National Scholarship, RUC Presidential Fellowship for Studying Abroad, RUC	2022 2019 2018-2023 2016 2016 2014 2013
GRANTS	Cisco Research Grant <i>Co-writer and primary student representative for Zhou Yu's group</i> "Selective Differential Privacy in Language Modeling" – \$40K PI: Zhou Yu, Ruoxi Jia	2022-2023
	NSF SaTC Grant <i>Co-writer and primary student representative for Zhou Yu's group</i> "Accurate and Scalable Privacy-Preserving Approaches for Natural Language Generation" – under revision PI: Zhou Yu, Ruoxi Jia	
EMPLOYMENT HISTORY	Research Intern, Meta AI Research <i>Hosts: Jason Weston and Jing Xu</i> <i>Project: Integrating user feedback into chatbots</i>	2022.05 - 2022.09
	Research Intern, Facebook AI Research <i>Host: Mike Lewis</i> <i>Project: Dialogue nonsense detection in the game of Diplomacy</i>	2021.05 - 2021.12

Research Intern, Facebook AI Research

2020.06 - 2020.12

Host: [Mike Lewis](#)

Project: [Strategic negotiation AI agent in the game of Diplomacy](#)

Full-time Data Scientist, [24]7.ai

2016 - 2018

Customer service chatbot development

PUBLICATIONS 18. [Human-Level Play in the Game of Diplomacy by Combining Language Models with Strategic Reasoning](#)

FAIR, Anton Bakhtin*, Noam Brown*, Emily Dinan*, Gabriele Farina, Colin Flaherty*, Daniel Fried, Andrew Goff, Jonathan Gray*, Hengyuan Hu*, Athul Paul Jacob*, Mojtaba Komeili, Karthik Konath, Minae Kwon, Adam Lerer*, Mike Lewis*, Alexander H. Miller*, Sasha Mitts, Adithya Renduchintala*, Stephen Roller, Dirk Rowe, [Weiyang Shi](#)*, Joe Spisak, Alexander Wei, David Wu*, Hugh Zhang*, Markus Zijlstra

(*A core team member. Authors listed alphabetically.)

Science, 2022. [[Meta AI blog post](#)] [[Science News](#)]

Selected media coverage: [[The New York Times front page \(01/22/2023\)](#)]

[[The Washington Post](#)] [[The Economist](#)] [[MIT Technology Review](#)] [[Forbes](#)]

The listed conferences are top-tier in NLP, HCI and AI with acceptance rates between 20%-25%.

17. [Social Influence Dialogue Systems: A Scoping Survey of the Efforts Towards Influence Capabilities of Dialogue Systems](#)

Kushal Chawla*, [Weiyang Shi](#)* (equal contribution), Jingwen Zhang, Gale Lucas, Zhou Yu, Jonathan Gratch

European Chapter of the Association for Computational Linguistics (EACL), 2023

16. [Just Fine-tune Twice: Selective Differential Privacy for Large Language Models](#)

[Weiyang Shi](#), Ryan Shea, Si Chen, Chiyuan Zhang, Ruoxi Jia, Zhou Yu

Empirical Methods in Natural Language Processing (EMNLP), 2022

15. [Seamlessly Integrating Factual Information and Social Content with Persuasive Dialogue](#)

Maximillian Chen, [Weiyang Shi](#), Feifan Yan, Ryan Hou, Jingwen Zhang, Saurav Sahay, Zhou Yu

Asia-Pacific Chapter of the Association for Computational Linguistics (ACL-IJCNLP), 2022

14. [Towards Socially Intelligent Agents with Mental State Transition and Human Utility](#)

Liang Qiu, Yizhou Zhao, Yuan Liang, Pan Lu, [Weiyang Shi](#), Zhou Yu, Song-Chun Zhu

Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL), 2022

13. [Selective Differential Privacy for Language Modeling](#)

[Weiyang Shi](#), Aiqi Cui, Evan Li, Ruoxi Jia, Zhou Yu

North American Chapter of the Association for Computational Linguistics (NAACL), 2022

12. [Refine and Imitate: Reducing Repetition and Inconsistency in Persuasion Dialogues via Reinforcement Learning and Human Demonstration](#)

[Weiyang Shi](#), Yu Li, Saurav Sahay, Zhou Yu

Findings of Empirical Methods in Natural Language Processing (EMNLP), 2021

11. **LEGOEval: An Open-Source Toolkit for Dialogue System Evaluation via Crowdsourcing**

Yu Li, Josh Arnold, Feifan Yan, Weiyang Shi, Zhou Yu

Demo of Annual Meeting of the Association for Computational Linguistics (ACL), 2021

10. **PRAL: A tailored pre-training model for task-oriented dialogue generation**

Jing Gu, Qingyang Wu, Chongruo Wu, Weiyang Shi, Zhou Yu

Annual Meeting of the Association for Computational Linguistics (ACL), 2021

9. **INSPIRED: Toward Sociable Recommendation Dialogue Systems**

Shirley Anugrah Hayati, Dongyeop Kang, Qingxiaoyang Zhu, Weiyang Shi, Zhou Yu

Empirical Methods in Natural Language Processing (EMNLP), 2020

8. **Structured Attention for Unsupervised Dialogue Structure Induction**

Liang Qiu, Yizhou Zhao, Weiyang Shi, Yuan Liang, Feng Shi, Tao Yuan, Zhou Yu, Song-Chun Zhu

Empirical Methods in Natural Language Processing (EMNLP), 2020

7. **Understanding User Resistance Strategies in Persuasive Conversations**

Youzhi Tian, Weiyang Shi, Chen Li, Zhou Yu

Findings of Empirical Methods in Natural Language Processing (EMNLP), 2020

6. **Effects of Persuasive Dialogues: Testing Bot Identities and Inquiry Strategies**

Weiyang Shi, Xuewei Wang, Yoo Jung Oh, Jingwen Zhang, Saurav Sahay, Zhou Yu

Conference on Human Factors in Computing Systems (CHI), 2020

5. **End-to-End Trainable Non-Collaborative Dialogue System**

Yu Li, Kun Qian, Weiyang Shi, Zhou Yu

AAAI Conference on Artificial Intelligence (AAAI), 2020

4. **How to Build User Simulators to Train RL-based Dialogue Systems**

Weiyang Shi*, Kun Qian* (equal contribution), Xuewei Wang, Zhou Yu

Empirical Methods in Natural Language Processing (EMNLP), 2019

3. **Persuasion for Good: Towards a Personalized Persuasive Dialogue System for Social Good (Best Paper Nomination)**

Xuewei Wang*, Weiyang Shi* (equal contribution), Richard Kim, Yoojung Oh, Sijia Yang, Jingwen Zhang, Zhou Yu

Annual Meeting of the Association for Computational Linguistics (ACL), 2019

2. **Unsupervised Dialogue Structure Learning**

Weiyang Shi, Tiancheng Zhao, Zhou Yu

North American Chapter of the Association for Computational Linguistics (NAACL), 2019

1. **Sentiment Adaptive End-to-End Dialogue Systems**

Weiyang Shi, Zhou Yu

Annual Meeting of the Association for Computational Linguistics (ACL), 2018

Weiyang Shi, Emily Dinan, Adi Renduchintala, Daniel Fried, Athul Paul Jacob, Zhou Yu, Mike Lewis
arXiv, 2022

1. **When Life Gives You Lemons, Make Cherryade: Converting Feedback from Bad Responses into Good Labels**

Weiyang Shi, Emily Dinan, Kurt Shuster, Jason Weston*, Jing Xu* (equal contribution)
arXiv, 2022

INVITED
TALKS

Social Influence Dialogue Systems. University of Maryland, Dec. 2022

Selective Differential Privacy in Language Modeling. Columbia NLP Seminar, Nov. 2021

Reducing Repetition and Contradiction in Dialogue Systems. Columbia NLP Seminar, Feb. 2021

Sentiment-adaptive Dialogue Systems. Cresta, Sept. 2019

MENTORING

Max Chen

Ph.D. student

Project: [Persuasive dialogue systems](#)

2021-Present
Columbia University

Yu Li

Ph.D. student

Project: [Anti-scam dialogue systems](#) and [dialogue evaluation](#)

2018-Present
Columbia University

Payton Guo

Undergraduate student

Project: Dialogue data denoising

2022-Present
New York University

Ryan Shea

Master student

Project: [Membership inference attack on NLP models](#)

2022-Present
Columbia University

Evan Li

Undergraduate student

Project: [Differential privacy and attack models in NLP](#)

2021
Columbia University

Chelsea Cui

Master student (Master at Columbia → Now at Amazon)

Project: [Differential privacy and dialogue simulation](#)

2021
Columbia University

Shirley Anugrah Hayati

Master student (Master at CMU → Now Ph.D. at University of Minnesota)

Project: [Sociable recommendation dialogues](#)

2019-2020
CMU

Jing Gu

Master student (Master at UC Davis → Now Ph.D. at UC Santa Cruz)

Project: [Pre-training models for task-oriented dialogue systems](#)

2019-2020
UC Davis

Xuwei Wang

Undergraduate student (Master at CMU → Now at Meta)

Project: [Persuasive dialogues](#) and [dialogue user simulator](#)

2019-2020
Zhejiang University

- TEACHING**
- Co-instructor: COMS 6998, Conversational AI, Columbia** Fall 2022
- *Topics:* various topics on dialogue systems (task-oriented and open-domain dialogue systems, knowledge-enriched dialogue systems, dialogue systems under low-resource settings, dialogue evaluations, and multimodal dialogues, etc).
 - Co-designed (with Prof. Zhou Yu) the curriculum; co-led in-class discussions; co-supervised course projects; co-hosted guest speakers; held office hours and graded assignments.
- Guest Lecturer: COMS 4705, Natural Language Processing, Columbia** Spring 2022
- *Topic:* dialogue systems (spoken and text-based)
 - Prepared and delivered lectures to a 100-student upper-division undergraduate introduction course to NLP.
- Teaching Assistant: COMS 4156, Advanced Software Engineering, Columbia** Fall 2021
- *Topics:* modern software engineering skills and practices
 - Supervised course projects; held office hours and graded assignments
- Teaching Assistant: COMS 6998, Conversational AI, Columbia** Spring 2021
- *Topics:* various topics on dialogue systems
 - Co-designed (with Prof. Zhou Yu) the curriculum; co-supervised course projects; held office hours and graded assignments.
- SERVICE**
- Organization Committee:** [Social Influence in Conversations](#), ACL 2023
- Diversity & Publicity Chair:** [5th Workshop for Conversational AI](#), ACL 2023
- Area Chair:** [17th Women in Machine Learning Workshop](#) (WiML 2022)
- Publicity Chair:** [4th Workshop for Conversational AI](#), ACL 2022
- Program Committee/Conference Reviewer:** ACL 2019, ACL 2020, *SEM 2020, ICLR 2021, AAAI 2021, EACL 2021, NAACL 2021, ACL 2021, *SEM 2021, NeurIPS 2021, EMNLP 2021, NLPCC 2021, ICLR 2022, AAAI 2022, ICML 2022, ACL 2022, CHI 2022, SIGDIAL 2022, HCI+NLP workshop at NAACL 2022, ACL Rolling Review 2021-2022
- Journal Reviewer:** ACM Transactions on Human-Robot Interaction, Neurocomputing, ACM Transactions on Information Systems
- CS Graduate Admission Committee:** Columbia University, 2022