

# Weiyan Shi (she/her/hers)

---

CONTACT	<a href="mailto:ws2634@columbia.edu">ws2634@columbia.edu</a> <a href="https://wyshi.github.io">wyshi.github.io</a> <a href="#">Google Scholar</a>	
RESEARCH INTERESTS	Natural Language Processing (dialogue systems, privacy-preserving NLP models, etc), Machine Learning, Artificial Intelligence	
EDUCATION	<b>Columbia University</b> Ph.D. in Computer Science Advisor: <a href="#">Zhou Yu</a>	2020 - 2023 (anticipated)
	<b>University of California, Davis</b> Ph.D. in Computer Science (transferred with my advisor) Advisor: <a href="#">Zhou Yu</a>	2018 - 2020
	<b>University of California, Berkeley</b> M.A. in Statistics	2015 - 2016
	<b>Renmin University of China</b> B.S. in Mathematics and Applied Mathematics	2011 - 2015
AWARDS & HONORS	Rising Star in Machine Learning, UMD <b>Best Paper Nomination, ACL</b> 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	<b>Cisco Research Grant</b> <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  <b>NSF SaTC Grant</b> <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	2022-2023
EMPLOYMENT HISTORY	<b>Research Intern, Meta AI Research</b> <i>Hosts: <a href="#">Jason Weston</a> and <a href="#">Jing Xu</a></i> <i>Project: <a href="#">Integrating user feedback into chatbots</a></i>  <b>Research Intern, Facebook AI Research</b> <i>Host: <a href="#">Mike Lewis</a></i> <i>Project: <a href="#">Dialogue nonsense detection in the game of Diplomacy</a></i>	2022.05 - 2022.09  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** 20. [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 member of the team. Authors listed alphabetically.)

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

Selected media coverage: [\[The Economist\]](#) [\[MIT Technology Review\]](#) [\[Forbes\]](#)

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

19. [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

18. [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

17. [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

16. [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

15. [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

14. [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

13. [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

12. **INSPIRED: Toward Sociable Recommendation Dialogue Systems**  
 Shirley Anugrah Hayati, Dongyeop Kang, Qingxiaoyang Zhu, Weiyan Shi, Zhou Yu  
*Empirical Methods in Natural Language Processing (EMNLP)*, 2020
  11. **Structured Attention for Unsupervised Dialogue Structure Induction**  
 Liang Qiu, Yizhou Zhao, Weiyan Shi, Yuan Liang, Feng Shi, Tao Yuan, Zhou Yu, Song-Chun Zhu  
*Empirical Methods in Natural Language Processing (EMNLP)*, 2020
  10. **Understanding User Resistance Strategies in Persuasive Conversations**  
 Youzhi Tian, Weiyan Shi, Chen Li, Zhou Yu  
*Findings of Empirical Methods in Natural Language Processing (EMNLP)*, 2020
  9. **Effects of Persuasive Dialogues: Testing Bot Identities and Inquiry Strategies**  
Weiyan Shi, Xuwei Wang, Yoo Jung Oh, Jingwen Zhang, Saurav Sahay, Zhou Yu  
*Conference on Human Factors in Computing Systems (CHI)*, 2020
  8. **End-to-End Trainable Non-Collaborative Dialogue System**  
 Yu Li, Kun Qian, Weiyan Shi, Zhou Yu  
*AAAI Conference on Artificial Intelligence (AAAI)*, 2020
  7. **How to Build User Simulators to Train RL-based Dialogue Systems**  
Weiyan Shi\*, Kun Qian\* (equal contribution), Xuwei Wang, Zhou Yu  
*Empirical Methods in Natural Language Processing (EMNLP)*, 2019
  6. **Persuasion for Good: Towards a Personalized Persuasive Dialogue System for Social Good (Best Paper Nomination)**  
 Xuwei Wang\*, Weiyan Shi\* (equal contribution), Richard Kim, Yoojung Oh, Sijia Yang, Jingwen Zhang, Zhou Yu  
*Annual Meeting of the Association for Computational Linguistics (ACL)*, 2019
  5. **Unsupervised Dialogue Structure Learning**  
Weiyan Shi, Tiancheng Zhao, Zhou Yu  
*North American Chapter of the Association for Computational Linguistics (NAACL)*, 2019
  4. **Sentiment Adaptive End-to-End Dialogue Systems**  
Weiyan Shi, Zhou Yu  
*Annual Meeting of the Association for Computational Linguistics (ACL)*, 2018
- PREPRINTS
3. **AutoReply: Detecting Dialogue Nonsense Introspectively with Discriminative Replies**  
Weiyan Shi, Emily Dinan, Adi Renduchintala, Daniel Fried, Athul Paul Jacob, Zhou Yu, Mike Lewis  
 arXiv, 2022
  2. **Social Influence Dialogue Systems: A Scoping Survey of the Efforts Towards Influence Capabilities of Dialogue Systems**  
 Kushal Chawla\*, Weiyan Shi\* (equal contribution), Jingwen Zhang, Gale Lucas, Zhou Yu, Jonathan Gratch  
 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	<i>Social Influence Dialogue Systems</i> . University of Maryland, Dec. 2022 <i>Selective Differential Privacy in Language Modeling</i> . Columbia NLP Seminar, Nov. 2021 <i>Reducing Repetition and Contradiction in Dialogue Systems</i> . Columbia NLP Seminar, Feb. 2021 <i>Sentiment-adaptive Dialogue Systems</i> . Cresta, Sept. 2019	
MENTORING	<b>Max Chen</b> Ph.D. student Project: <a href="#">Persuasive dialogue systems</a>	2021-Present Columbia University
	<b>Yu Li</b> Ph.D. student Project: <a href="#">Anti-scam dialogue systems</a> and <a href="#">dialogue evaluation</a>	2018-Present Columbia University
	<b>Payton Guo</b> Undergraduate student Project: Dialogue data denoising	2022-Present New York University
	<b>Ryan Shea</b> Master student Project: <a href="#">Membership inference attack on NLP models</a>	2022-Present Columbia University
	<b>Evan Li</b> Undergraduate student Project: <a href="#">Differential privacy and attack models in NLP</a>	2021 Columbia University
	<b>Chelsea Cui</b> Master student (Master at Columbia → Now at Amazon) Project: <a href="#">Differential privacy and dialogue simulation</a>	2021 Columbia University
	<b>Shirley Anugrah Hayati</b> Master student (Master at CMU → Now Ph.D. at University of Minnesota) Project: <a href="#">Sociable recommendation dialogues</a>	2019-2020 CMU
	<b>Jing Gu</b> Master student (Master at UC Davis → Now Ph.D. at UC Santa Cruz) Project: <a href="#">Pre-training models for task-oriented dialogue systems</a>	2019-2020 UC Davis
	<b>Xuwei Wang</b> Undergraduate student (Master at CMU → Now at Meta) Project: <a href="#">Persuasive dialogues</a> and <a href="#">dialogue user simulator</a>	2019-2020 Zhejiang University
TEACHING	<b>Co-instructor: COMS 6998, Conversational AI</b> , 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

**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