Yao Dou

douy@gatech.edu
https://yao-dou.github.io/

Education

2021 Georgia Institute of Technology
Ph.D. in Computer Science   Aug 2021 – Aug 2026
Advisor: Prof. Wei Xu

2017 University of Washington
B.S. in Computer Science   Sep 2017 – Jun 2021
Graduated at age 20
Advisor: Prof. Yejin Choi

Experience

2024 Microsoft Research, Deep Learning Group
Research Intern, May 2024 – April 2025
Mentor: Michel Galley and Jianfeng Gao

2021 Allen Institute for AI, Mosaic Team
Research Intern, Jan 2021 – Mar 2021
Mentor: Maxwell Forbes and Yejin Choi

2020 Noonum, Inc
Data Science Intern, Jun 2020 – Sep 2020

Publications

EMNLP 2025 SimulatorArena: Are User Simulators Reliable Proxies for Multi-Turn Evaluation of AI Assistants?
Yao Dou, Michel Galley, Baolin Peng, Chris Kedzie, Weixin Cai, Alan Ritter, Chris Quirk, Wei Xu, Jianfeng Gao

COLM 2025 Evaluating LLMs on Chinese Idiom Translation
Cai Yang, **Yao Dou**, David Heineman, Xiaofeng Wu, Wei Xu

CSCW 2025 Measuring, Modeling, and Helping People Account for Privacy Risks in Online Self-Disclosures with AI
Isadora Krsek, Anubha Kabra, **Yao Dou**, Tarek Namus, Laura A Dabbish, Alan Ritter, Wei

	Xu, Sauvik Das
ICML 2025	CollabLLM: From Passive Responders to Active Collaborators Shirley Wu, Michel Galley, Baolin Peng, Hao Cheng, Gavin Li, **Yao Dou**, Weixin Cai, James Zou, Jure Leskovec, Jianfeng Gao **Outstanding Paper Award**
AAAI 2025	CROSSNEWS: A Cross-Genre Authorship Verification and Attribution Benchmark Marcus Ma, Duong Minh Le, Junmo Kang, **Yao Dou**, John Cadigan, Dayne Freitag, Alan Ritter, Wei Xu
ArXiv	TemporalBench: Benchmarking Fine-grained Temporal Understanding for Multimodal Video Models Mu Cai, Reuben Tan, Jianrui Zhang, Bocheng Zou, Kai Zhang, Feng Yao, Fangrui Zhu, Jing Gu, Yiwu Zhong, Yuzhang Shang, **Yao Dou**, Jaden Park, Jianfeng Gao, Yong Jae Lee, Jianwei Yang
EMNLP 2024	Improving Minimum Bayes Risk Decoding with Multi-Prompt David Heineman, **Yao Dou**, Wei Xu
EMNLP 2024	GPT-4 Jailbreaks Itself with Near-Perfect Success Using Self-Explanation Govind Ramesh, **Yao Dou**, Wei Xu
ACL 2024	Reducing Privacy Risks in Online Self-Disclosures with Language Models **Yao Dou**, Isadora Krsek, Tarek Naous, Anubha Kabra, Sauvik Das, Alan Ritter, Wei Xu
EMNLP 2023	Thresh: A Unified, Customizable and Deployable Platform for Fine-grained Text Evaluation (Demo) David Heineman, **Yao Dou**, Wei Xu
EMNLP 2023	Dancing Between Success and Failure: Edit-Level Simplification Evaluation using SALSA David Heineman, **Yao Dou**, Mounica Maddela, Wei Xu
ACL 2023	LENS: A learnable Evaluation Metric for Text Simplification Mounica Maddela*, **Yao Dou***, David Heineman, Wei Xu
EMNLP 2022	Improving Large-scale Paraphrase Acquisition and Generation **Yao Dou**, Chao Jiang, Wei Xu
ACL 2022	Is GPT-3 Text Indistinguishable from Human Text? Scarecrow: A Framework for Scrutinizing Machine Text **Yao Dou***, Maxwell Forbes*, Rik Koncel-Kedziorski, Noah A. Smith, Yejin Choi
AAAI 2021	MultiTalk: A Highly-Branching Dialog Testbed for Diverse Conversations **Yao Dou**, Maxwell Forbes, Ari Holtzman, Yejin Choi

Tutorials

ACL 2024	Automatic and Human-AI Interactive Text Generation **Yao Dou***, Philippe Laban*, Claire Gardent, Wei Xu

Honors and Awards

Microsoft Scholarship, 2020 − 2021
Dean's List, 2017 − 2021

Teaching

Georgia Institute of Technology
  TA, Advanced NLP (CS 8803), Fall 2023; NLP (CS 4650), Spring 2025
  Prof. Wei Xu

  TA, Machine Learning (CS 4641), Fall 2022
  Prof. Mahdi Roozbahani

University of Washington
  TA, Machine Learning (CSE 446), Spring 2020 and Fall 2020
  Prof. Kevin Jamieson, Prof. Jamie Morgenstern

Undergraduate Mentees

Cai Yang
Vinayak Athavale
Govind Ramesh
David Heineman -> PYI at Allen Institute for AI