

Shi Feng *Curriculum Vitae*

CONTACT	shif@uchicago.edu ihsgnef.github.io scholar.google.com/shifeng	
ACADEMIC POSITIONS	University of Chicago Postdoctoral scholar Advisor: Chenhao Tan	2021 - 2023
EDUCATION	University of Maryland <i>Ph.D. in Computer Science</i> Advisor: Jordan Boyd-Graber Thesis: <i>Towards Human-AI Cooperation on Sequential Decision Making Problems</i>	2016 - 2021
	Shanghai Jiao Tong University <i>B.S. in Computer Science</i> ACM Honor Class	2012 - 2016
INDUSTRY EXPERIENCE	Salesforce Research <i>Research Intern</i> Advisor: Bryan McCann	June 2020 - Sept 2020 Remote
	Microsoft Research <i>Research Intern</i> Advisor: Jessica Lundin	June 2018 - Sept 2018 Redmond, Washington
	Microsoft Research Asia <i>Research Intern</i> Advisors: Shujie Liu, Mu Li	August 2015 - February 2016 Beijing, China
PUBLICATIONS	<ul style="list-style-type: none">[1] Measuring Inductive Biases of In-Context Learning with Underspecified Demonstrations Chenglei Si*, Dan Friedman*, Nitish Joshi, Shi Feng, Danqi Chen, He He <i>Association for Computational Linguistics (ACL)</i>, 2023.[2] Machine Explanations and Human Understanding Chacha Chen*, Shi Feng*, Amit Sharma, Chenhao Tan <i>Transactions on Machine Learning Research (TMLR)</i>, 2023 <i>ACM Conference on Fairness, Accountability, and Transparency (ACM FAccT)</i>, 2023 Best Paper, <i>ICML Workshop on Human-Machine Collaboration and Teaming (HMCaT)</i>, 2022.[3] Learning Human-Compatible Representations for Case-Based Decision Support Han Liu, Yizhou Tian, Chacha Chen, Shi Feng, Yuxin Chen, Chenhao Tan <i>International Conference on Learning Representations (ICLR)</i>, 2023.[4] Learning to Explain Selectively Shi Feng, Jordan Boyd-Graber <i>Empirical Methods in Natural Language Processing (EMNLP)</i>, 2022.[5] Active Example Selection for In-Context Learning Yiming Zhang, Shi Feng, Chenhao Tan <i>Empirical Methods in Natural Language Processing (EMNLP)</i>, 2022.[6] Calibrate Before Use: Improving Few-shot Performance of Language Models Tony Z. Zhao*, Eric Wallace*, Shi Feng, Dan Klein, Sameer Singh <i>International Conference on Machine Learning (ICML)</i>, 2021.[7] Concealed Data Poisoning Attacks on NLP Models Eric Wallace*, Tony Z. Zhao*, Shi Feng, Sameer Singh <i>North American Chapter of the Association for Computational Linguistics (NAACL)</i>, 2021.[8] Quizbowl: The Case for Incremental Question Answering Pedro Rodriguez, Shi Feng, Mohit Iyyer, He He, Jordan Boyd-Graber <i>Journal of Machine Learning Research (JMLR)</i>, 2021.[9] What can AI do for me: Evaluating Machine Learning Interpretations in Cooperative Play Shi Feng, Jordan Boyd-Graber <i>ACM Conference on Intelligent User Interfaces (ACM IUI)</i>, 2019.	

- [10] Universal Adversarial Triggers for Attacking and Analyzing NLP
Eric Wallace, **Shi Feng**, Nikhil Kandpal, Matt Gardner, Sameer Singh
Empirical Methods in Natural Language Processing (EMNLP), 2019.
- [11] Misleading Failures of Partial-input Baselines
Shi Feng, Eric Wallace, Jordan Boyd-Graber
Association for Computational Linguistics (ACL), 2019.
- [12] Understanding Impacts of High-Order Loss Approximations and Features in Deep Learning Interpretation
Sahil Singla, Eric Wallace, **Shi Feng**, Soheil Feizi
International Conference on Machine Learning (ICML), 2019.
- [13] Trick Me If You Can: Human-in-the-loop Generation of Adversarial Examples for Question Answering
Eric Wallace, Pedro Rodriguez, **Shi Feng**, Jordan Boyd-Graber
Transactions of the Association for Computational Linguistics (TACL), 2019.
- [14] Pathologies of Neural Models Make Interpretation Difficult
Shi Feng, Eric Wallace, Alvin Grissom II, Mohit Iyyer, Pedro Rodriguez, Jordan Boyd-Graber
Empirical Methods in Natural Language Processing (EMNLP), 2018.
- [15] Improving Attention Modeling with Implicit Distortion and Fertility for Machine Translation
Shi Feng, Shujie Liu, Nan Yang, Mu Li, Ming Zhou, Kenny Q. Zhu
International Conference on Computational Linguistics (COLING), 2016.

TEACHING
EXPERIENCE

Courses:

- Teaching assistant: UMD’s graduate level NLP course.
- Teaching assistant: UMD’s graduate level ML course.
- Teaching assistant: UMD’s undergraduate level Objective-oriented Programming course.
- Teaching assistant: SJTU’s undergraduate level Computation Theory and Automata course.
- Teaching assistant: SJTU’s undergraduate level Algorithms course.

Tutorials:

- NAACL, 2022. *Human-Centered Evaluation of Explanations*

MENTORING

Student Research Mentoring

- Chenglei Si (2022-2023), UMD Undergrad. Published [1]. Now PhD at Stanford.
- Yiming Zhang (2022-2023), UChicago Masters. Published [5] Now PhD at CMU.
- Chacha Chen (2022-2023), UChicago PhD. Published [2].
- Han Liu (2022-2023), UChicago PhD. Published [3]
- Eric Wallace (2018-2019), UMD Undergrad. Published [10, 13] Now PhD at UC Berkeley.
- Matthew Shu (2021-present), Highschool student. Now Undergrad at Yale.

PRESENTATIONS

Invited Talks & Presentations & External Visits

- GWU, 2023. *Evaluating AI: From Crowdsourcing Truths to Truth-finding Processes*
- UNT, 2023. *Pragmatic Interpretability*
- UAlbany, 2023. *Pragmatic Interpretability*
- USC ISI, 2023. *Pragmatic Interpretability*
- NEC Labs Europe, 2023. *Pragmatic Interpretability*
- UPenn, 2019. *Evaluating Interpretability of NLP Models*
- UCSD, 2019. *Evaluating Interpretability of NLP Models*
- UCI, 2019. *Evaluating Interpretability of NLP Models*
- NLP Highlights Podcast, 2019. *Pathologies of Neural Models Make Interpretation Difficult*

Conference Oral Presentations: ACL 2023 [1], FAccT 2023 [2] NAACL 2021 Virtual [7], EMNLP 2019 Hong Kong, [10], IUI 2019 Los Angeles, [9], EMNLP 2018 Brussels [14].

ACADEMIC
SERVICE

Program Committee Member

- Conferences: ACL (2019, 2020), NeurIPS (2020, 2021, 2022), EMNLP (2018, 2019, 2020, 2021), ACL Rolling Review (2021, 2022), ICLR (2021), NAACL (2022), FAccT (2023), CoNLL (2020)
- Workshops: Human-Centered Explainable AI (CHI 2023), Trust and Reliance in AI-Assisted Tasks (CHI 2023), Dynamic Adversarial Data Collection (NAACL 2022)

Awards

- **Best Reviewer:** EMNLP (2018, 2020), NeurIPS (2020)