

Hua Shen Ph.D. (she/her)

Postdoctoral Scholar
Information School
University of Washington
Seattle, WA, 98195, USA

+1 484-929-4568
<http://hua-shen.org>
huashen@uw.edu
[@huashen218](https://twitter.com/huashen218)

RESEARCH INTERESTS

My research focuses on achieving **human-centered AI alignment**, i.e., helping humans explain and communicate with AI technology via interactions and, moreover, enhancing AI ethics by aligning their values and behavior with humans and society. The research keywords include bidirectional human-AI alignment, human-centered AI ethics, human-AI interaction, and conversational AI explanation.

EDUCATION

- 2024.09 - **Postdoctoral Scholar**
Responsible AI Systems and Experiences (RAISE) Center
Information School
University of Washington, Seattle, USA
- 2023.08 - 2024.08 **Postdoctoral Research Fellow**
School of Information, and Computer Science and Engineering Division
University of Michigan, Ann Arbor, USA
- 2019.08 - 2023.07 **Ph.D. in Informatics**
College of Information Sciences and Technology
Penn State University, PA, USA (*GPA: 3.93/4.0*)
- 2013.09 - 2016.06 **M.S. in Management Science and Engineering**
School of Information
Renmin University of China, Beijing, China (*Admission Exam Exemption*)
- 2009.09 - 2013.06 **B.S. in Information Security**
School of Computer and Communication Engineering
University of Science and Technology Beijing, China (*Top: 2/35*)

INDUSTRY EXPERIENCE

- 2022.09 - 2023.01 **Google Research**
Research Intern, Speech Intelligence Team, Seattle, USA
Mentors: Vicky Zayats, Johann Rocholl, Dan Walker, Dirk Padfield
[MultiTurnCleanup: A Benchmark for Multi-turn Spoken Conversational Transcript Cleanup](#)
(*EMNLP 2023, Oral*)
- 2022.05 - 2022.08 **Amazon Science**
Applied Scientist Intern, Alexa AI, Language Modeling Team, Seattle, USA
Mentors: Jari Kolehmainen, Yi Gu, Qi Luo, Aditya Gourav, Ivan Bulyko
Senior Principal Scientists: Ankur Gandhe, Denis Filimonov, Ariya Rastrow, Andreas Stolcke
Distilling Generative Large Language Models
(*Internal Presentation at Amazon Science*)

2021.05 - 2021.08 **Amazon Science**
 Applied Scientist Intern, Alexa AI, SpeakerID Team, Sunnyvale, USA
Mentors: Yuguang Yang, Guoli Sun, Ryan Langman, Eunjung (Christine) Han
Senior Principal Scientist Advisors: Jasha Droppo, Andreas Stolcke
[Improving Fairness in Speaker Verification via Group-adapted Fusion Network](#)
(ICASSP 2022)

SELECTED HONORS AND AWARDS

2024	Future Faculty Diversity Program , (Acceptance Rate < 20%), Virginia Tech
2024	Best Paper , (Top 1% , 1/76), AIED Conference
2024	Best Interactive Event Award , (Top 1% , 1/76), AIED Conference
2024	Carnegie Funds for Faculty Development (\$5,000) , UMich, School of Information
2023	Best Demo (Top 6% , 1/16), CSCW Conference
2023	Rising Stars of Data Science , UC San Diego, UChicago and Stanford University
2023	Google Research Science Conference Scholarships (\$3000) , Google Research
2023	Best Paper Honorable Mention (Top 5% , 3/65), IUI Conference
2020, 2022	Graduate Student Travel Grant (Virtual Conferences) (\$500), College of IST, PennState
2019	CRA-W Grad Cohort Award (\$2,000) , Grad Cohort Workshop for Women
2018 - 2019	Lehigh University Fellowship (\$57,752) , Lehigh University
2015	Outstanding Master Degree Candidate (Top 12.5% , 5/40), RenMin University of China
2013	Outstanding Merit Student (Top 0.2% , 1/550), Beijing
2013	Excellent Graduate (Top 0.2% , 1/550), Beijing
2013	Outstanding-Dissertation Award (Top 0.5% , 3/550), USTB
2011, 2012	National Endeavor Fellowship (¥10,000 , Top 2%, 5/550),
2012	Alumni Funding Scholarship (¥5,000 , Top 0.3%, 8/3,000), USTB

PUBLICATIONS

Preprints Under Review

2025	P.26	Hua Shen , Tiffany Kneareem, Reshmi Ghosh, Kenan Alkiek, Kundan Krishna, Yachuan Liu, Ziqiao Ma, Savvas Petridis, Yi-Hao Peng, Li Qiwei, Sushrita Rakshit, Chenglei Si, Yutong Xie, Jeffrey P. Bigham, Frank Bentley, Joyce Chai, Zachary Lipton, Qiaozhu Mei, Rada Mihalcea, Michael Terry, Diyi Yang, Meredith Ringel Morris, Paul Resnick, David Jurgens. Towards Bidirectional Human-AI Alignment: A Systematic Review for Clarifications, Framework, and Future Directions . arXiv:2406.09264.
	P.25	Hua Shen , Nicholas Clark, and Tanushree Mitra. Mind the Value-Action Gap: Do LLMs Act in Alignment with Their Values? arXiv:2501.15463.
	P.24	Hua Shen , Tiffany Kneareem, Reshmi Ghosh, Yu-Ju Yang, Tanushree Mitra, and Yun Huang. ValueCompass: A Framework of Fundamental Values for Human-AI Alignment . arXiv:2409.09586.
	P.23	Qianou Christina Ma, Weirui Peng, Hua Shen , Kenneth Koedinger, Tongshuang Wu. What You Say = What YouWant? Teaching Humans to Articulate Requirements for LLMs arXiv:2409.08775.

Archival Conference & Journal Papers

- 2020-2025 P.22 Tuc Nguyen, James Michels, **Hua Shen**, Thai Le. NoMatterXAI: Generating “No Matter What” Alterfactual Examples for Explaining Black-Box Text Classification Models. 39th Annual AAAI Conference on Artificial Intelligence.
- *AAAI 2025*
- P.21 Snehal Prabhudesai, Ananya Prashant Kasi, Anmol Mansingh, Anindya Das Antar, **Hua Shen**, Nikola Banovic. “Here the GPT made a choice, and every choice can be biased”: How Students Critically Engage with LLMs through End-User Auditing Activity. 2025 Conference on Human Factors in Computing Systems.
- *CHI 2025*
- P.20 Qianou Christina Ma, **Hua Shen**, Kenneth Koedinger, Tongshuang Wu. How to Teach Programming in the AI Era? Using LLMs as a Teachable Agent for Debugging. The 25th International Conference on Artificial Intelligence in Education.
- *AIED 2024* 🏆 **Best Paper** 🏆 **Best Interactive Event Award**
- P.19 **Hua Shen**, Vicky Zayats, Johann Rocholl, Dan Walker, and Dirk Padfield. MultiTurn-Cleanup: A Bench-mark for Multi-Turn Spoken Conversational Transcript Cleanup. The 2023 Conference on Empirical Methods in Natural Language Processing.
- *EMNLP 2023 (Oral)* 🏆 **Google Research Science Conf. Scholarships**
- P.18 **Hua Shen**, Chieh-Yang Huang, Tongshuang Wu, Ting-Hao (Kenneth) Huang. ConvXAI: Delivering Heterogeneous AI Explanations via Conversations to Support Human-AI Scientific Writing. The 26th ACM Conference On CSCW - Demonstrations.
- *CSCW 2023 Demo* 🏆 **Best Demo**
- P.17 Tongshuang Wu, **Hua Shen**, Daniel S Weld, Jeffrey Heer, Marco Tulio Ribeiro. ScatterShot: Interactive In-context Example Curation for Text Transformation. The ACM Conference on Intelligent User Interfaces.
- *IUI 2023* 🏆 **Best Paper Honorable Mention**
- P.16 **Hua Shen***, Adaku Uchendu*, Jooyoung Lee*, Thai Le, Ting-Hao’ Kenneth’ Huang, and Dongwon Lee. Does Human Collaboration Enhance the Accuracy of Identifying Deepfake Texts? The 11th AAAI Conference on Human Computation and Crowdsourcing.
- *AAAI HCOMP 2023*
- P.15 **Hua Shen**, Tongshuang Wu, Wenbo Guo, Ting-Hao (Kenneth) Huang. Are Shortest Rationales the Best Explanations For Human Understanding? 60th Annual Meeting of the Association for Computational Linguistics.
- *ACL 2022*
- P.14 **Hua Shen***, Yuguang Yang*, Guoli Sun, Ryan Langman, Eunjung Han, Jasha Droppo, Andreas Stolcke. Improving Fairness in Speaker Verification via Group-adapted Fusion Network. 2022 IEEE International Conference on Acoustics, Speech and Signal Processing.
- *ICASSP 2022* ★ **Amazon AI Applied Scientist Intern**
- P.13 **Hua Shen**, Ting-hao (Kenneth) Huang. How Useful Are the Machine-Generated Interpretations? A Human Evaluation on Guessing the Wrongly Predicted Labels. The 8th AAAI Conference on Human Computation and Crowdsourcing.
- *AAAI HCOMP 2020*
- P.12 Mina Lee, Katy Llonka Gero, John Joon Young Chung, Simon Buckingham Shum, Vipul Raheja, **Hua Shen**, Subhashini Venugopalan, Thiemo Wambsganss, (+28 authors). A Design Space for Intelligent and Interactive Writing Assistants. The ACM CHI conference on Human Factors in Computing Systems.
- *CHI 2024*
- P.11 Binfeng Xu, Xukun Liu, **Hua Shen**, Zeyu Han, Yuhan Li, Murong Yue, Zhiyuan Peng, Yuchen Liu, Ziyu Yao, Dongkuan Xu. Gentopia.AI: A Collaborative Platform for Tool-Augmented LLMs. The 2023 Conference on EMNLP - System Demonstration.
- *EMNLP 2023 Demo*
- P.10 Kai-Wei Chang, Haibin Wu, Yu-Kai Wang, Yuan-Kuei Wu **Hua Shen**, Wei-Cheng Tseng, Iu-thing Kang, Shang-Wen Li, Hung-yi Lee. SpeechPrompt: Prompting Speech Language Models for Speech Processing Tasks.
- *Transactions on Audio, Speech and Language Processing (TASLP 2024)*.

- P.9 Xinyang Zhang, Ningfei Wang, **Hua Shen**, Shouling Ji, Ting Wang. [Interpretable Deep Learning under Fire](#). Proceedings of the 29th USENIX Security Symposium.
- *USENIX 2020*
- P.8 Ren Pang, **Hua Shen**, Xinyang Zhang, Shouling Ji, Yevgeniy Vorobeychik, Xiapu Luo, Alex X. Liu, Ting Wang. [The Tale of Evil Twins: Adversarial Inputs versus Poisoned Models](#). The 27th ACM Conference on Computer and Communications Security.
- *ACM CCS 2020*
- 2015-2018 P.7 **Hua Shen**, Xun Liang. [A Quantitative Analysis Decision System Based on Deep Learning and NSGA-II for FX Portfolio Prediction](#). The 31st International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems.
- *IEA-AIE 2018*
- P.6 **Hua Shen**, Xun Liang. [A Time Series Forecasting Model Based on Deep Learning Integrated Algorithm with Stacked Autoencoders and SVR for FX Prediction](#). The 25th International Conference on Artificial Neural Networks.
- *ICANN 2016*
- P.5 **Hua Shen**, Xun Liang, Mingming Wang. [Emergency Decision Support Architectures for Bus Hijacking Based on Massive Image Anomaly Detection in Social Networks](#). The 2015 IEEE International Conference on Systems, Man, and Cybernetics.
- *IEEE SMC 2015*

Non-Archival Workshop Papers

- 2020-2024 P.4 **Hua Shen**, Tongshuang Wu. [Parachute: Evaluating Interactive Human-LM Co-writing Systems](#). The In2Writing Workshop at The 2023 ACM CHI Conference.
- *CHI 2023 In2Writing Workshop*
- P.3 **Hua Shen**, Ting-hao (Kenneth) Huang. [Explaining the Road Not Taken](#). The 2021 ACM CHI Workshop on Operationalizing Human-Centered Perspectives in Explainable AI.
- *CHI 2021 HCXAI Workshop*
- P.2 Shih-Hong Huang, Chieh-Yang Huang, Yuxin Deng, **Hua Shen**, Szu-Chi Kuan, and Ting-Hao' Kenneth' Huang. [Too Slow to Be Useful? On Incorporating Humans in the Loop of Smart Speakers](#). The WiP/Demo track of the 10th AAAI Conference on Human Computation and Crowdsourcing.
- *AAAI HCOMP 2022 WiP/Demo*
- P.1 Jiaqi Wang, **Hua Shen**, Chacha Chen, Frank E. Ritter. [Are Learners Satisfied with Their MOOC Experiences? Assessing and Improving Online Learners' Interactions](#).
- *Asian CHI 2021 Symposium*

Book

- 2014 B.1 Xun Liang, Xiaoping Yang, **Hua Shen**. [Social Commerce Theory and Practice](#)[M], Tsinghua University Press, 2014.
ISBN No. 9787302381129

Patents

- 2018-2015 PT.2 Xun Liang, **Hua Shen**, Run Cao. [An innovative Emergency Discovery Method for Microblogging](#), Archived by State Intellectual Property Office of The P.R.C (SIPO), Patent Publication.
No. CN103577404A

- PT.1 Xiaofei Li, Xun Liang, Xiaoping Zhou, Xiaojing Shi, **Hua Shen**, Haiyan Zhang. [A Cross-Platform Microblog Community Account Matching Method](#), Archived by State Intellectual Property Office of The P.R.C (SIPO), Patent Publication.
No. CN104765729A

RESEARCH COMMUNITY SERVICE

Program Committee / Associate Chair

HCI	CHI 2025 (Associate Chairs). CHI LBW 2024 (Associate Chair). CHI'24 HCXAI Workshop. CHI'23 In2Writing Workshop.
NLP	ACL 2023, 2021. EMNLP 2023, 2022. ICML 2023 AI & HCI Workshop.
AI	ICML 2023 AI & HCI Workshop.

Paper Reviewing

NLP	ACL 2024, NAACL 2024, EMNLP 2023, ACL RollingReview (ARR) 2023, 2022, 2021.
HCI	CSCW 2024, 2025, CHI 2024, TIIS 2024, CHI Late-Breaking Work 2023, 2022
Speech	ICASSP 2024
AI	NeurIPS 2022 HEGM Workshop. KDD 2020, 2019. ICDM 2019. ASONAM 2019. CIKM 2018. BigData 2018. TIST Transactions on Intelligent Systems and Technology 2019. TIS Transactions on Information Systems 2019.

RESEARCH TALKS

2024.11	Cornell University. Towards Bidirectional Human-AI Alignment via Interaction
2024.11	Columbia University. Towards Bidirectional Human-AI Alignment via Interaction
2024.11, 10	University of Washington. Towards Bidirectional Human-AI Alignment via Interaction
2024.10	Google. Towards Human-Centered Alignment in Speech Language Models
2024.10	Virginia Tech. Towards Bidirectional Human-AI Alignment via Interaction
2024.7	Microsoft, Responsible & Open AI Research Towards Bidirectional Human-AI Alignment via Interaction
2024.7	University of Illinois Urbana-Champaign. Towards Bidirectional Human-AI Alignment via Interaction
2023.12	EMNLP 2023. Oral talk. A Benchmark for Multi-Turn Spoken Conversational Transcript Cleanup
2023.11, 10	University of Michigan - Ann Arbor. Towards Human-Centered AI Alignment via Interaction
2023.11	Rising Star of Data Science. UChicago & UC San Diego. Towards Useful AI Interpretability via Interactive AI Explanations
2023.07	Carnegie Mellon University. Towards Useful AI Interpretability via Interactive AI Explanations
2023.05	Princeton University. Towards Useful AI Interpretability via Interactive AI Explanations


2023.04	CHI 2023 In2Writing Workshop. Parachute: Evaluating Interactive Human-LM Co-writing Systems
2023.03	IUI 2023. Virtual. ScatterShot: Interactive In-context Example Curation for Text Transformation
2023.01	Google Research. A Benchmark for Multi-Turn Spoken Conversational Transcript Cleanup
2022.09	University of Notre Dame. Useful XAI via Interactive AI Explanations
2022.08, 2021.08	Amazon Alexa AI. Distilling Auto-Regressive Large Language Models Improving Fairness in Speaker Verification Models
2021.05	CHI 2021 HCXAI Workshop. A Survey on Human-Centered Explainable NLP

MENTORING EXPERIENCE

PhD	Qianou (Christina) Ma (Carnegie Mellon University) How to Teach Programming in the AI Era? Using LLMs as a Teachable Agent for Debugging Teaching Humans to Articulate Requirements for LLMs Snehal Prabhudesai (University of Michigan) How Students Critically Engage with LLMs through End-User Auditing Activity Yu-Ju (Marisa) Yang (University of Illinois Urbana-Champaign) A Framework of Fundamental Values for Human-AI Alignment Kenan Alkiek (University of Michigan) Spatial Understanding of Large Language Models Jooyoung Lee (Penn State) Human collaborative evaluation on LLM-generated deepfake texts
Master	Anmol Mansingh (University of Michigan) How Students Critically Engage with LLMs through End-User Auditing Activity Ruchi Panchanadikar (Penn State; Incoming PhD at Clemson University) Developing Interactive AI Explanation System
Bachelor	Yuxin Deng (Penn State; Master at CMU) Human Study of Interactive AI Explanation System Ananya Kasi (University of Michigan) How Students Critically Engage with LLMs through End-User Auditing Activity Reuben Lee (Penn State; TDP at Capital One) Developing Interactive AI Explanation System Sushirita Rakshit (University of Michigan) Understanding of Large Language Models Yijun Pan (University of Michigan) Understanding of Large Language Models Daniel Tian (University of Michigan) Understanding of Large Language Models Omkar Yadav (University of Michigan) Stylometric Authorship Discernment and Interpretation Haotian Zhang (University of Michigan) Stylometric Authorship Discernment and Interpretation Michael Jiang (University of Michigan) Stylometric Authorship Discernment and Interpretation

TEACHING ASSISTANT EXPERIENCE

Penn State University

- 2021  **Graduate Student Online Teaching Certificate**
- 2023 SP, 2021 FA [DS 330] **Visual Analytics for Data Sciences**
Undergraduate course, 70 students (Instructor: Luke Zhang, Sharon Huang)
- 2022 SP [DS 320] **Data Integration**
Undergraduate course, 50 students (Instructor: Yasser EL-Manzalawy)
- 2021 SP, 2020 FA [DS 397] **Algorithmic Methods**
Undergraduate course, 25 students (Instructor: Hadi Hosseini)
- 2023, 2021 [IST 261] **Application Development Studio I**
Undergraduate course, 50 students (Instructor: Kenneth Huang)

Renmin University of China

- 2015 FA **Calculus II**
Undergraduate course, 50 students (Instructor: Wenlin Huang)
- 2015 SP **Calculus I**
Undergraduate course, 50 students (Instructor: Wenlin Huang)
- 2014 FA **Linear Algebra**
Undergraduate course, 40 students (Instructor: Qijiang Song)
- 2013 FA, 2014 SP **Probability Theory and Mathematical Statistics**
Undergraduate course, 40 students (Instructor: Qijiang Song)

MEDIA AND CONFERENCE COVERAGE

- NAACL 2024 **Tutorial on Human-AI Interaction in the Age of LLMs.**
Featuring Papers: [ScatterShot](#), [HypoCompass](#), [ConvXAI](#)
Instructor: [Diyi Yang](#), [Sherry Tongshuang Wu](#), [Marti Hearst](#)
- ACL 2024 **Tutorial on Automatic & Human-AI Interactive Text Generation.**
Featuring Papers: [Parachute](#), [Design Space for In2Writing](#)
Instructor: [Yao Dou](#), [Philippe Laban](#), [Claire Gardent](#), [Wei Xu](#)
- EMNLP 2023 **Tutorial on Designing, Evaluating, and Learning from Humans Interacting with NLP Models.**
Featuring Paper: [Parachute](#)
Speaker: [Tongshuang Wu](#), [Diyi Yang](#) and [Sebastin Santy](#)
- PSU News 2020 **Users don't understand computer explanations for image labeling errors.**
[PSU News](#), [TechXplore News](#), [ClickTech News](#)

OUTREACH AND LEADERSHIP

Service and Leadership

- 2025 **Initiator & Organizer**, [Bidirectional Alignment Special Interest Group @CHI'25](#)
Propose the workshop to foster interdisciplinary collaboration and submissions to support human agency, empower understanding and critical thinking in AI use, foster effective human-AI collaboration, and promote societal adaptations to maximize the AI benefits.

2025	Initiator & Organizer, Bidirectional Alignment Workshop @ICLR'25 Propose this workshop to foster interdisciplinary collaboration and encourage AI-centered researchers to perceive alignment as a dynamic and mutual process, encourage submissions to explore diverse methods to integrate human values to train, steer, and customize AI.
2024	MichiganAI Seminar Tsar, Michigan AI Lab. Coordinated MichiganAI seminars, including inviting speakers, coordinating and broadcasting seminars, and handled seminar logistics in July and August 2024.
2024	MichiganAI Social Media Take-Over, Michigan AI Lab. Took over Michigan AI Lab social media, i.e., Twitter and Facebook, in July 2024.
2024	Co-Organizer, The Michigan Interactive and Social Computing (MISC) research group. Co-organized the research events in 2024-2025 that connect researchers of HCI, CSCW across The University of Michigan from June to August, 2024.
2024	Co-organizer of Michigan UMSI Job Navigation Panels, School of Information. Initiated and organized two panels to invite senior and junior professors to provide experience and advice on academic job search for students.
2022	Student contributor, JSALT 2022. Contributed in the JSALT Leveraging Pre-training Models for Speech Processing Team.
2021	Invited lecturer, Penn State IST 501, Host: Dr. Mary Beth Rosson. 1 of the 3 representatives of 2019 PhDs to mentor 2020 PhDs in qualification exam.
2021	Course design lecturer, Penn State DS 330, Host: Dr. Sharon Huang. Served in Visual analytics for data sciences course project lecture.
2016-2018	Leader, CMBC Management Trainee Program. Served in head office future bankers program of China Minsheng Bank.
2014	Liaison officer, Singapore deputy prime minister delegation in APEC. Served in APEC (Asia-Pacific Economic Cooperation) finance ministerial meeting.
2012	Vice Chairman, USTB student union presidium. Served in the 6th presidium of the 22nd session of USTB student union.

TECHNICAL AND METHOD SKILLS

DL/ML Programming	PyTorch, TensorFlow, PyTorch-Transformers, S3PRL Speech, Scikit-Learn
System Development	Full-stack (frontend + ML-based backend) interactive system development.
Web and Visualization	React.js, HTML/CSS, JavaScript, D3.js, Bootstrap, Matplotlib
HCI/UX Research	Human Interview and Qualitative Coding, Human Studies and and Statistical Analysis
Database	Microsoft SQL Server, MySQL, PostgreSQL
Miscellaneous	Amazon MTurk Crowdsourcing, Adobe Illustrator, LaTeX
Operational Systems	Ubuntu, Linux/Unix, Shell Script, MacOS, Windows