	Wesley Hanwen Deng	
408-893-1312	hanwend@cs.cr	mu.edu
$\begin{array}{c} {\bf Research} \\ {\bf Interests} \end{array}$	Responsible AI; Human-AI Interaction; Algorithm Auditing; Organizational Sc Crowdsourcing; Social Computing; Fairness, Accountability, and Transparency.	
Education	Carnegie Mellon University, Pittsburgh, PAPh.D. in Human-Computer Interaction, School of Computer Science2021Advisors: Ken Holstein & Motahhare Eslami.	1 - now
	University of California, Berkeley, Berkeley, CAB.A. in Computer ScienceHighest Distinction & EECS Research HonorResearch Advisors: Niloufar Salehi & Kimiko Ryokai.	7 - 2020
Honor and Awards	K&L Gates Ethics and AI Presidential Fellow \$60,000, K&L Gates Endowment for Ethics and Computational Technologies at O Mellon University	Carnegie 2025
	Microsoft Research AI & Society Fellow \$45,000, Microsoft Research with Ken Holstein, Motahhare Eslami, Jason I. Hong.	2024
	Best Paper Award on Generative AI Red-Teaming 7th AAAI/ACM Conference on Artificial Intelligence, Ethics, and Society, (AIE [P10], with Michael Feffer, Anusha Sinha, Zachary C. Lipton, Hoda Heidari	ES '24). 2024
	Best Paper Award on User-engaged AI Auditing 12th AAAI Conference on Human Computation and Crowdsourcing (HCOMP [P8], with Sara Kingsley, Jiayin Zhi, Sizhe Zhang, Ken Holstein, Motahhare I Tianshi Li, Jason I. Hong, Hong Shen.	
	Award Winner of The Notre Dame-IBM Tech Ethics Lab Call for Pre \$60,000, Norte Dame-IBM Technology Ethics Lab with Motahhare Eslami, Ken Holstein, Jason I. Hong, and Toby Jia-jun Li.	oposal 2024
	Best Project Runner Up, CMU AI Governance Course (2/31), Carnegie Mellon University Selected by Juries from the Block Center for Technology and Society	2024
	Special Recognitions for Outstanding Reviews Computer-Supported Cooperative Work and Social Computing (CSCW)	2024
	AI Ethics & Society Workshop Invited Participants \$1,500, Allen Institute for AI	2023
	Special Recognitions for Outstanding Reviews CHI Conference on Human Factors in Computing Systems (CHI)	2023
	CMU Graduate Student Travel Grant , \$750, Carnegie Mellon University	2022
	CERES (Connecting the EdTech Research EcoSystem) Fellowship through Jacobs Foundation, Carnegie Mellon University with Ken Holstein, Amy Ogan, Jeff Bigham	2021

	Special Recognitions for Outstanding Reviews Computer-Supported Cooperative Work and Social Computing (CSCW)	2021
	EECS Undergraduate Research Honor Electrical Engineering and Computer Science, UC Berkeley	2020
	Highest Distinction, School of Letter and Science, UC Berkeley	2020
	Dean's List, School of Letter and Science, UC Berkeley 2018, 2	019, 2020
Conference Publications Stringently	[P12] Wesley H. Deng, Wang Claire, Howard Z. Han, Jason I. Hong [*] , Ken H. Motahhare Eslami [*] . (2025). WeAudit: Scaffolding User Auditors and AI Pra in Auditing Generative AI. In Submission	

Peer-reviewed

[P11] Wesley H. Deng, Solon Barocas^{*}, Jennifer Wortman Vaughan^{*}. (2025). Supporting Industry Computing Researchers in Assessing, Articulating, and Addressing the Potential Negative Societal Impact of Their Work. *To appear in Proceedings of the ACM on Human-Computer Interaction (CSCW '25).*

[P10] Michael Feffer, Anusha Sinha, Wesley H. Deng, Zachary C. Lipton, Hoda Heidari (2024). Red-Teaming for Generative AI: Silver Bullet or Security Theater? In Proceedings of AAAI/ACM Conference on Artificial Intelligence, Ethics, and Society, (AIES '24)

[P9] Jordan Taylor, Wesley H. Deng, Ken Holstein, Sarah Fox, Haiyi Zhu. (2024). Carefully Unmaking the "Marginalized User:" A Diffractive Analysis of a Gay Online Community. In Proceedings of ACM Transactions on Computer-Human Interaction (TOCHI '24)

[P8] Sara Kingsley*, Jiayin Zhi*, **Wesley H. Deng**, Jaimie Lee, Sizhe Zhang, Motahhare Eslami*, Kenneth Holstein*, Joson I. Hong*, Tianshi Li*, Hong Shen*. (2024). Investigating What Factors Influence Users' Detection of Harmful Algorithmic Bias and Discrimination. In Proceedings of the 12th AAAI Conference on Human Computation and Crowdsourcing (HCOMP '24).

[P7] Wesley H. Deng, Nur Yildirim, Monica Chang, Motahhare Eslami, Ken Holstein, Michael Madaio. (2023). Investigating Practices and Opportunities for Crossfunctional Collaboration around AI Fairness in Industry Practice. In Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT '23)

[P6] Wesley H. Deng, Boyuan Bill Guo, Alicia DeVrio, Hong Shen, Motahhare Eslami*, Ken Holstein*. (2023). Understanding Practices, Challenges, and Opportunities for User-Engaged Algorithm Auditing in Industry Practice. *in Proceedings of the ACM CHI Conference on Human Factors in Computing Systems (CHI '23).*

[P5] Wesley H. Deng, Manish Nagireddy, Michelle Seng Ah Lee, Jatinder Singh, Zhiwei Steven Wu, Kenneth Holstein, Haiyi Zhu. (2022). Exploring How Machine Learning Practitioners (Try To) Use Fairness Toolkits. In Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22)

[P4] Hong Shen, Leijie Wang, **Wesley H. Deng**, Ciell, Ronald Velgersdijk, Haiyi Zhu. (2022). The Model Card Authoring Toolkit: Towards Community-centered, Deliberation-driven AI Design. In Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22)

[P3] Dan Liebling, Katherine Heller, Samantha Roberston, Wesley H. Deng. (2022).

Opportunities for Human-centered Evaluation of Machine Translation Systems In Proceedings of the 2022 North American Chapter of the Association for Computational Linguistics (NAACL '22)

[P2] Hong Shen, Wesley H. Deng, Aditi Chattopadhyay, Steven Wu, Xu Wang, Haiyi Zhu. (2021). Value Cards: An Educational Toolkit for Teaching Social Impacts of Machine Learning through Deliberation. In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21)

[P1] Jon Gillick, **Wesley H. Deng**, Kimiko Ryokai, David Bamman. Robust Laughter Detection in Noisy Environments. (2021) In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '21)

Workshops & [WT06] Wesley H. Deng, Glen Berman, Harmanpreet Kaur, Reva Schwartz, Memunat A. Ibrahim, Motahhare Eslami, Ken Holstein, Michael Madaio (2024). Collaboratively Designing and Evaluating Responsible AI Interventions. Workshop at the ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW'24). ACM.

[WT05] Wesley H. Deng, Mireia Yurrita, Mark Díaz, Jina Suh, Nick Judd, Lara Groves, Hong Shen, Motahhare Eslami, Kenneth Holstein. (2024) Responsible Crowd-sourcing for Responsible Generative AI: Engaging Crowds in AI Auditing and Evaluation. Workshop at the AAAI 12th AAAI Conference on Human Computation and Crowdsourcing (HCOMP '24)

[WT04] Hoda Heidari, **Wesley H. Deng**^{*}, Michael A. Feffer^{*}, Anusha Sinha^{*}; Supporting NIST's Development of Guidelines on Red-teaming for Generative AI. Workshop at CMU (2024)

[WT03] Ziang Xiao, Wesley H. Deng, Michelle Lam, Motahhare Eslami, Juho Kim, Mina Lee, Q. Vera Liao. (2024) Human Centered Evaluation and Auditing of Large Language Models. Workshop at the ACM CHI Conference on Human Factors in Computing Systems (CHI'24)

[WT02] Wesley H. Deng, Michelle Lam, Alex Cabrera, Danaë Metaxa, Motahhare Eslami, Ken Holstein. (2023). Supporting User Engagement in Testing, Auditing, and Contesting AI. Workshop at the ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW'23). ACM.

[WT01] Wesley H. Deng, Shivani Kapania, Ken Holstein, Motahhare Eslami. Panelist: Lauren Wilcox, Su Lin Blodgett, Danaë Metaxa, Nicholas Diakopoulos, Karrie Karahalios, Shubhanshu Mishra, Christo Wilson (2023). User Engagement in Algorithm Testing and Auditing: Exploring Opportunities and Tensions between Practitioners and End Users. Critiquing and Rethinking Fairness, Accountability, and Transparency (CRAFT) session at the 2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT '23)

Workshop[WP10] Glen Berman, Ned Cooper, Wesley H. Deng, Ben Hutchinson (2024). Trou-
bling Taxonomies in GenAI Evaluation. NeurIPS Workshop on Evaluating Evaluations:
Examining Best Practices for Measuring Broader Impacts of Generative AI (EvalEval,
NeurIPS '24)

[WP9] Claire Wang, Wesley H. Deng, Jason I. Hong, Kenneth Holstein, Motahhare Eslami. (2024). Designing a Crowdsourcing Pipeline to Verify Reports from User AI Audits. Work in Progress / Demo at the AAAI 12th AAAI Conference on Human Computation and Crowdsourcing (HCOMP '24)

[WP8] Matheus Kunzler Maldaner, **Wesley H. Deng**, Jason I. Hong, Kenneth Holstein, Motahhare Eslami. (2024). MIRAGE: Multi-model Interface for Reviewing and Auditing Generative Text-to-Image AI. Work in Progress / Demo at the AAAI 12th AAAI Conference on Human Computation and Crowdsourcing (HCOMP '24)

[WP7] Wesley H. Deng^{*}, Michael A. Feffer^{*}, Anusha Sinha^{*}, Hoda Heidari (2024). An Executive Summary of Comments Submitted to NIST in Response to RFI 88368. White paper for National Institute of Standards and Technology

[WP6] Wesley H. Deng, Jason Hong (2024). Organizing Crowd Audits to Detect Bias in Machine Learning. NSF Fairness in AI PI Meeting; 2024

[WP5] Wesley H. Deng, Motahhare Eslami, Kenneth Holstein. (2023) Towards "Anytime, Anywhere" Community Learning and Engagement around the Design of Public Sector AI. Workshop on AI Literacy: Finding Common Threads between Education, Design, Policy, and Explainability; CHI 2023

[WP4] Wesley H. Deng, Nikita Mehandru, Samantha Roberston, Niloufar Salehi. (2022) Beyond General Purpose Machine Translation: The Need for Context-specific Empirical Research to Design for Appropriate User Trust. *Workshop on Trust and Reliance in AI-Human Teams; CHI 2022*

[WP3] Wesley H. Deng, Manish Narigeddy, Zhiwei Steven Wu, Kenneth Holstein, Haiyi Zhu. (2021) Evaluating Fairness in Practice: Exploring How Machine Learning Practitioners Use Fairness Toolkits. *Human-Centered AI Workshop; NeurIPS 2021*

[WP2] Samantha Roberston, **Wesley H. Deng**, Timnit Gebru, Margaret Mitchell, Dan Liebling, Michal Lahav, Katherine Heller, Mark Diaz, Samy Bengio, Niloufar Salehi. (2021) Three Directions for the Design of Human-Centered Machine Translation. 1st HCI + NLP Workshop, NAACL 2021

[WP1]Kimiko Ryokai, Julia Park, **Wesley H. Deng**. (2020) Personal laughter archives: reflection through visualization and interaction. Adjunct Proceedings of the 2020 ACM International Conference on Ubiquitous Computing (UbiComp 2020)

Grant Writing &[G12] Understanding On-the-ground Usage of AI Risk Taxonomy2025PreparationAwarded (\$5,000); Google Research Gift Grant2025Co-PI; with Glen Berman and Ned CooperCooper

 [G11] Mixed-Initiative AI Red-teaming: Supporting Effective Human-AI Collaboration in Red-teaming Generative AI 2025
 Not recommended for funding; Notre Dame-IBM Technology Ethics Lab
 Co-PI; Co-written with Alice Qian Zhang, Jason I. Hong, Adam Perrer, Motahhare Eslami, Ken Holstein, Hong Shen

[G10] Supporting Participatory AI Auditing of Generative AI Systems
 2025
 Awarded (\$65,000); Seed Fund of Block Center for Technology and Society
 Co-PI; Co-written with Motahhare Eslami, Ken Holstein, Jason Hong.

[G9] Empowering Privacy Practitioners in Doing Responsible AI Work 2025 Not recommended for funding; Seed Fund of Block Center for Technology and Society **Co-PI**; Co-written with Sauvik Das, Hank Hao-ping Lee, Motahhare Eslami, Ken Holstein, Jason Hong.

[G8] AI Ethics & Social Impact: Thinking about impact with Generative AI EthicsPlaybook2024Not recommended for funding; Commitment Grant Presentation for Bill & MelindaGates Foundation

	Project Presenter ; Co-presented with Jessie J. Smith, Je Jesse Dodge, and Margaret Mitchell	on Ball, Nicole DeCario,
	 [G7] Engaging End-users in Surfacing Harmful AI Behaviors Awarded (\$60,000); Notre Dame-IBM Technology Ethics I PI; Co-written with Motahhare Eslami, Ken Holstein, and Jawith Toby J. Li (Norte Dame University) 	Lab
	 [G6] Searching for Less Discriminatory Algorithms: Mechanic Requirements Submitted; National Science Foundation; Planning Grant for velopment, and Deployment of Technologies (ReDDDoT) Student contributor; Co-written with Emily Black, Logar Miranda Bogen 	2024 • Responsible Design, De-
	 [G5] Generative AI Meets Responsible AI: Supporting User-Dr of Generative AI Systems Awarded (\$60,000); Google Research Scholar Program Student contributor; Co-written with Motahhare Eslami, A 	2024
	[G4] Bridging Between Regulation and On-the-Ground Pract Awarded (\$45,000); MSR AI & Society Challenge PI; Co-written with Ken Holstein, Motahhare Eslami, and Je	2023
	[G3] Generative AI Meets Responsible AI: Supporting User-Dr of Generative AI Systems Not recommended for funding; AWS Amazon Research Awar Student contributor; Co-written with Motahhare Eslami, K	2023 ds
	 [G2] Scaffolding Responsible AI Practice at the Earliest Sta Formulation and Project Selection Awarded (\$350,193); PwC, Digital Transformation and In Student contributor; Co-written with Ken Holstein, Mota izzi, Hoda Heidari, Christian Kaestner, Haiyi Zhu, and John 	2023 novation Center hhare Eslami, Jodi Forl-
	[G1] Public AI Engagement through Public Installation and T Not recommended for funding; <i>Remaking Learning Moonshot</i> PI ; <i>Co-written with Ken Holstein, Motahhare Eslami</i>	
Research Experience	AllenNLP, Allen Institute for AI Research Assistant (Research Intern offer deferred due to Vis Research supervisor: Jesse Dodge & Nicole DeCario Research Topics: Responsible AI, Generative AI, Impact Ass	,
	FATE , <i>Microsoft Research</i> Research Intern Research supervisors: Jenn Wortman Vaughan & Solon Baro Research Topics: Responsible AI, Impact Assessment.	Summer 2023
	4A Lab , <i>Carnegie Mellon University</i> Ph.D. Student Research supervisors: Motahhare Eslami Research Topics: Responsible AI ([P7], [P8]), Algorithm Aud	Summer 2021 - now liting [P7]
	CoALA Lab , <i>Carnegie Mellon University</i> Pre-Doctoral Research Assistant & Ph.D. Student	Fall 2020 - now

Research supervisors: Ken Holstein

	Research Topics: Responsible AI ([WP3], [P5], [P7], [P8]), Algorithm Audit	ting $([P7])$
	Social AI Group, Carnegie Mellon University Pre-Doctoral Research Assistant Research supervisors: Haiyi Zhu and Steven Wu Research Topics: Responsible AI ([P2], [P4], [P5]), ML Fairness ([WP3], Social Computing [P6]	- Fall 2021 [P5]) and
	To3 Group, University of California, Berkeley	- Fall 2020 P3])
		- Fall 2020
Invited Talk and Panelist	Supporting Responsible AI in Real-world Organizational Contexts: Practices, Challenges, and Opportunities in Industry Settings	: Current
	Invited talk at Tencent, Shanghai	05/2024
	Invited talk at Google, Shanghai	05/2024
	Invited talk at Salesforce, (Remote)	04/2024
	Invited talk at Capital One, Data Science Team	07/2023
	Invited talk at TU Delft, Faculty of Industrial Design Engineering	05/2023
	Invited talk at TU Delft, Faculty of Electrical Engineering, Mathematics and Science, Delft University of Technology	$\begin{array}{c} \text{Computer} \\ 05/2023 \end{array}$
	Understanding Practices, Challenges, and Opportunities for User- Algorithm Auditing in Industry Practice	-Engaged
	Invited talk at University of California, Berkeley, Algorithmic Fairness an (AFOG) Group	nd Opacity $10/2022$
	Exploring How ML Practitioners (Try To) Use Fairness Toolkits	
	Invited talk at PwC AI Governance Group	03/2022
	Invited talk at Google PAIR	08/2022
	Invited talk at IBM AIF360 developer team monthly meeting	07/2022
	Invited talk at Fairlearn developer team meeting	03/2021
	Other talks and panels	
	Invited panelist, "Flavor of HCI" hosted by Carnegie Mellon University, User Experience Association	11/2022
	Invited speaker, "Applying for Ph.D. programs in Human-Computer Intera hosted by UC Berkeley, Chinese Student Association	$\begin{array}{c} \text{ction"} \\ 07/2022 \end{array}$

	Invited panelist, Graduate School Application Panel, hosted by Carnegie Mellon University, REU Summer Program in HCII	06/2022
	Invited presenter, Design Innovation Lecture Series, hosted by UC Berkeley, Jacobs Institute for Design Innovation	08/2022
	Invited panelist, EECS Research Honor program panel, hosted by UC Berkeley, EECS Honor program	03/2021
	Invited presenter, Designing Emergent Technology, hosted by UC Berkeley, Jacobs Institute for Design Innovation	12/2019
Selected Press	12/2023, World Privacy Forum, Risky Analysis: Assessing and Improving ernance Tools. An international review of AI Governance Tools and sugge pathways forward	
	07/2023, Center for Advancing Safety of Machine Intelligence, AI eth at Chicago conference, precursor to CASMI's next workshop.	ics debate
	06/2023, Queer in AI, 3 things that AI ethics toolkits get wrong.	
	12/2022, WIRED , ChatGPT, Galactica, and the Progress Trap: When large models fall short, the consequences can be serious. Why is it so hard to ach that?	
	11/2021, CMU HCII , HCII Researchers To Study Improving Algorithm, Science Education as Part of CERES Network.	Computer
Academic Service	Associate Chair DIS 2025, AI and Design Subcommittee	
	Program Committee FAccT, Since 2022 Empathy-centric Design at Scale Workshop, CHI 2022	
	Conference Reviewer CHI, since 2021, with special recognition for outstanding reviews (2023) CSCW, since 2021, with special recognition for outstanding reviews (2021, 2 FAccT, since 2021 AIES, since 2023 ICASSP 2021 TEI 2023	2024)
Community Service	Organizer : CMU HCII Ph.D. Student Social Hour 2021, 2022, 2 Organize the bi-monthly social event to foster engagement and mutual support HCII Ph.D. students, visiting scholars, and post-doctors.	
	Organizer : CMU HCII Ph.D. & Faculty Off-campus Tea Hour 22 Launch and organize semester-wise off-campus tea hours at a local coffee shop community and collaboration among HCII Ph.D. students and faculty memb	
	Organizer : CMU Ph.D. Student CHI Conference Practice Talk 2022, 2	2023, 2024

	Coordinate the annually event for HCII Ph.D. students to practice their conference talk for the CHI conference on Human Factors in Computing Systems.
	Organizer : CMU HCII Summer REU and Ph.D. Social Event 2023 Coordinate social event for the REU program, enabling undergraduates to connect with HCII faculty and Ph.D. students, and gain insights into HCI research and doctoral life.
	Organizer : CMU Responsible AI Initiate Ph.D. Student Gathering 2022 Initiate and coordinate the inaugural CMU Ph.D. student gathering focused on Re- sponsible AI research, promoting cross-campus communication and collaboration among students engaged in this field.
	Organizer : CMU HCII Ph.D. Junior-Senior Coffee chat 2021 Coordinated the department-supported HCII Ph.D. student coffee chats to foster mutual support and mentorship between junior and senior HCII Ph.D. students.
	Co-organizer : CMU HCII Department Welcoming Event 2024 Co-organize the HCII department welcoming party for all HCII Ph.D. students, Master students, staffs, and faculty members.
	Co-organizer: CMU HCII Seminar Speaker Student Host 2021, 2022, 2023, 2024 Support the department to host HCII seminar speakers by organizing the student lunch.
	Participant: CMU School of Computer Science Ph.D. Application Assistancefor underrepresented applicants2022, 2023Support 3-5 underrepresented applicants each year to apply for CS Ph.D. program.
	Participant : Graduate School Application Panel for Summer REU Program 2022 Attending panels for REU students and provide overview and help of PhD application.
	Participants: CMU HCII Ph.D. Opening House Events2021, 2022, 2023, 2024Volunteer for welcoming accepted Ph.D. candidates through various activities.
Research Mentorship	Jordan Wheeler, CMU, Research Assistant through Research Experiences for Undergraduates (REU)Summer 2024 - NowMatheus Kunzler Maldaner, University of Florida, Research Assistant through Research Experiences for Undergraduates (REU)Summer 2024 - NowClaire Wang, University of California, Berkeley, Research Assistant through Research Experiences for Undergraduates (REU)Summer 2024 - NowYanwei Huang, Zhejiang University, Research AssistantSummer 2024 - NowYanwei Huang, Zhejiang University, Research AssistantSummer 2024 - NowHoward Han, CMU, Master Student Research AssistantSpring 2023 - Spring 2024 - NowNow CMU HCII Ph.D.Spring 2023 - Spring 2024Xiaofeng Yan, CMU, Master Student Research AssistantFall 2022 - Spring 2023Jieyu Zhou, CMU, Master Student Research AssistantSpring 2022 - Spring 2023Leijie Wang, Tsing-hua University, Undergraduate HCI Research AssistantSpring 2021 - Spring 2021Now UW CSE Ph.D.Spring 2021 - Spring 2022Bill Guo, CMU, Undergraduate HCI Research AssistantSummer 2021 - Fall 2022
Project Mentorship	Bhavya Jha, CMU, Master Student Research AssistantSpring 2024 - NowVicky Zhu, CMU, Undergraduate Independent StudySpring 2024Carson Piehl, CMU, Undergraduate Independent StudySpring 2024Yvvone Fang, CMU, Master Student Research Assistant Spring 2023 - Summer 2023Anthony Pan, CMU, Undergraduate Design Project

Christy Zo, CMU, Undergraduate Design Project

Fall 2022

Rebecca Jiang, CMU, Undergraduate HCI Independent Study	Summer 2022
Vivek Nadig, Valley Christian High School, High School student	Summer 2022
Andrew Sim, CMU, Master Student Independent Study	Spring 2022
Harnoor Dhingra, CMU, Master Student Independent Study	Spring 2022
Susan Huang, CMU, Undergraduate HCI Research Assistant	Fall 2021
Alice Tran, CMU, Undergraduate HCI Research Assistant	Fall 2021

Teaching

Human-AI Interaction, 05-710, CMU

Graduate TA; 2024

User Centered Research and Evaluation, 05-710, CMU	Graduate TA; 2023
Data Science, Data 100, UC Berkeley	Undergraduate TA; 2020
Discrete Mathematics and Probability Theory, CS70, UC	Berkeley Group Tutor; 2019
Data Structure, CS 61B, UC Berkeley	Lab Assistant; 2019
Introduction of Computing, CS 61A, UC Berkeley	Lab Assistant; 2018
Physics for Engineers, Physics 7A, UC Berkeley	One-on-one Tutor; 2018