

# RYAN STEED

ryansteed@cmu.edu • rbsteed.com

## EDUCATION

---

<b>Carnegie Mellon University</b> Ph.D., Machine Learning & Public Policy Advised by Alessandro Acquisti	August 2020 – Present GPA: 3.99
<b>Carnegie Mellon University</b> M.S., Machine Learning Research	August 2020 – May 2024 GPA: 3.99
<b>George Washington University</b> B.S. Computational Economics, Minor in Philosophy, <i>summa cum laude</i>	August 2017 – May 2020 GPA: 3.98

## RESEARCH EXPERIENCE

---

<b>Carnegie Mellon University</b> <i>Doctoral Candidate</i>	August 2020 – Present
Leading empirical research focused on evaluating the accountability and privacy of algorithmic systems to inform tech policy.	
<ul style="list-style-type: none"><li>• Published award-winning research in <i>Science</i> and other top academic publications.</li><li>• Featured in <i>MIT Tech Review</i>, <i>Vice</i>, <i>Nextgov</i>, and other outlets; cited in the <i>Fairness and Machine Learning</i> textbook and federal policy reports on AI accountability.</li><li>• Co-authored \$595k NSF grant on “privacy-preserving” tech; hired, managed, and mentored 10 research assistants across 4 projects.</li><li>• Delivered invited talks to academics, practitioners, students, and policymakers; presented to the Federal Trade Commission and Census Bureau.</li></ul>	

<b>Mozilla Foundation, Open Source Audit Tooling Project</b> <i>Collaborator</i>	July 2022 – Present
---	---------------------

Contributing to research on the practice of AI auditing. Produced two academic publications and a report; submitted 3 public comments on AI accountability to government agencies in the U.S. and E.U.

<b>Oracle Labs</b> <i>Software Engineering Intern, Ph.D. Research</i>	May 2021 – August 2021
--	------------------------

Scoped, motivated, and led investigation of social bias transfer from pre-trained large language models to fine-tuned models. First-authored a research publication.

<b>George Washington University</b> <i>Research Assistant, Department of Computer Science</i>	September 2018 – September 2020
--	---------------------------------

Led research on embedded social bias in computer vision, preference elicitation for participatory machine learning, and the effects of patent policy on innovation. First-authored 3 academic publications.

## OTHER EXPERIENCE

---

<b>Wilson Center</b> <i>Pathway to AI Policy Fellow</i>	October 2024 – Present
--	------------------------

Providing technical expertise on AI to U.S. federal policymakers.

<b>OpenAI</b> <i>Consultant, DALL-E 2 Red Team</i>	March 2022 – April 2022
---	-------------------------

Evaluated state-of-the-art image generation model DALL-E 2 for stereotyping and other ethical issues.

**SolveForGood** June 2020 – September 2020  
*Volunteer Data Scientist*  
 Built an algorithmic transparency tool for the property Assessor’s Office in Cook County, Chicago.

**GW Innovation Center** September 2017 – May 2020  
*Virtual Jane Project Manager & Technology Fellow*  
 Recruited and managed a team of 10 student VR animators/developers and fundraised over \$10k to design a VR app for environmental education. Hosted Dr. Jane Goodall for motion capture scan.

**Capital One** June 2019 – August 2019  
*Data Engineering Intern*  
 Designed a novel cluster sampling algorithm for active learning in analyst-in-the-loop cybersecurity.

## GRANTS & AWARDS

---

**National Science Foundation Grant #2319919 (\$595k)** October 2023 – Present  
*Understanding the Impact of Privacy-Preserving Analytics on the Ground, with Alessandro Acquisti*

**George Duncan Award for Best Second Research Paper** May 2024  
*Heinz College, Carnegie Mellon University*

**Distinguished Paper Award** April 2024  
*IEEE Conference on Secure and Trustworthy Machine Learning*

**Meta Research PhD Fellowship—Tech Policy** August 2022 – May 2024  
*Meta Platforms, Inc.*

**Suresh Konda Award for Best First Research Paper** May 2022  
*Heinz College, Carnegie Mellon University*

**TCS Presidential Fellowship** August 2021 – May 2022  
*Carnegie Mellon University*

**Social Good Fellowship** September 2017 – May 2020  
*GWU Innovation Center*

**Data MASTER Fellowship** May 2018 – May 2019  
*GWU Mathematics And Statistics Training, Education, and Research Program*

## REFEREED PUBLICATIONS

---

Birhane, A., **Steed, R.**, Ojewale, V., Vecchione, B., Raji, I., *AI auditing: The broken bus on the road to AI accountability*. In *2024 IEEE Conference on Secure and Trustworthy Machine Learning (SaTML)*, Apr. 2024. **Distinguished Paper Award**.

Ojewale, V., **Steed, R.**, Vecchione, B., Birhane, A., Raji, I. D., *Towards AI Accountability Infrastructure: Gaps and Opportunities in AI Audit Tooling*. Feb. 27, 2024.

**Steed, R.**, Liu, T., Wu, Z. S., Acquisti, A., “Policy impacts of statistical uncertainty and privacy”. *Science*, 377(6609), Aug. 26, 2022. **Award for Best First Research Paper at Heinz College. Cited by the AEA Committee on Economic Statistics.**

**Steed, R.**, Panda, S., Kobren, A., Wick, M., *Upstream Mitigation Is Not All You Need: Testing the Bias Transfer Hypothesis in Pre-Trained Language Models*. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics*. ACL 2022, May 2022.

**Steed, R.**, Caliskan, A., *Image Representations Learned With Unsupervised Pre-Training Contain Human-like Biases*. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, FAccT ’21*, Mar. 3, 2021. **Cited in *Fairness and Machine Learning*.**

**Steed, R.**, Caliskan, A., “A set of distinct facial traits learned by machines is not predictive of appearance bias in the wild”. *AI and Ethics*, 2021.

## OTHER PUBLICATIONS & WORKING PAPERS

---

**Steed, R.**, Schnadower Mustri, E., Acquisti, A., *Impacts of Differential Privacy on Social Science Research*. 2024.

**Steed, R.**, Acquisti, A., *Adoption of ‘Privacy-Preserving’ Analytics: Drivers, Designs, & Decoupling*. Feb. 7, 2024. **Award for Best Second Research Paper at Heinz College.**

**Steed, R.**, Qing, D., Wu, Z. S., *Quantifying Privacy Risks of Public Statistics to Residents of Subsidized Housing*. July 5, 2024.

Acquisti, A., **Steed, R.**, “Learning to Live with Privacy-Preserving Analytics”. *Communications of the ACM*, 66(7), June 22, 2023.

Caliskan, A., **Steed, R.**, *Managing the risks of inevitably biased visual artificial intelligence systems*. Brookings. Sept. 26, 2022.

**Steed, R.**, Williams, B., *Heuristic-Based Weak Learning for Automated Decision-Making*. In Workshop on Participatory Machine Learning at the International Conference on Machine Learning, Dec. 2, 2020.

## PUBLIC COMMENTS

---

Raji, D., Birhane, A., Vecchione, B., **Steed, R.**, Ojewale, V., *Comment on NIST-2023-0309*. Request for Information (RFI) Related to NIST’s Assignments Under Sections 4.1, 4.5 and 11 of the Executive Order Concerning Artificial Intelligence, National Institute of Standards and Technology, Feb. 2, 2024.

Raji, D., Vecchione, B., Birhane, A., **Steed, R.**, Ojewale, V., *Comment on NTIA-2023-0005*. NTIA AI Accountability Request for Comment, National Telecommunications and Information Administration, June 15, 2023. **Cited in the NTIA AI Accountability Policy Report.**

Raji, D., Vecchione, B., Birhane, A., **Steed, R.**, Ojewale, V., *Feedback from Mozilla Open Source Audit Tooling (OAT) Project*. Delegated Regulation on data access provided for in the Digital Services Act F3423931, European Commission, May 31, 2023.

## PRESS

---

M. Heikkilä. “The viral AI avatar app Lensa undressed me—without my consent”. *MIT Technology Review*, Dec. 12, 2022.

K. Errick. “Statistical Uncertainty Could be Problematic for Evidence-Based Policies, Study Finds”. *Nextgov*, Aug. 2022.

C. Hackerott. *Carnegie-Mellon Researchers Published Report on Statistical Uncertainty in Census Bureau Data*. JD Supra. Sept. 12, 2022.

K. Hao. “An AI saw a cropped photo of AOC. It autocompleted her wearing a bikini.” *MIT Technology Review*, Jan. 29, 2021.

T. Feathers. “Sexist AI is Even More Sexist Than We Thought”. *Vice*, Feb. 22, 2021.

D. Gershgorn. “Men Wear Suits, Women Wear Bikinis: Image Generating Algorithms Learn Biases ‘Automatically’”. *OneZero*, Jan. 29, 2021.

## SELECTED TALKS & PRESENTATIONS

---

Harvard Privacy Tools Seminar	October 2024
MITRE, Modeling & Analysis Innovation Center Brown Bag	September 2024
OpenDP Community Meeting, DP Beyond Algorithms Workshop	August 2024
[Poster] Theory and Practice of Differential Privacy (TPDP) Workshop	July 2022, August 2024
Mozilla Foundation, MozWeek	August 2024
NBER Conference on Data Privacy Protection and the Conduct of Applied Research	May 2023, 2024
Federal Trade Commission, Office of Technology Journal Club	April 2024
IEEE Conference on Secure & Trustworthy Machine Learning (SaTML)	April 2024
Mozilla Foundation, AI Learning Sessions	April 2024
USENIX Conference on Privacy Engineering Practice and Respect	September 2023
[Workshop] Privacy Law Scholars Conference	June 2023
Hot Topics in the Science of Security Symposium (HoTSoS)	April 2023
Boston Data Privacy Group	October 2022
2nd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization (EAAMO '22)	October 2022
U.S. Census Bureau, Statistical Research Division	December 2021

## TEACHING & SERVICE

---

**Teaching Assistant:** Machine Learning in Practice & Machine Learning for Public Policy, August 2022 – December 2022; Privacy in the Digital Age, August 2021 – May 2022.

**Reviewer:** *Journal of Privacy and Confidentiality*, 2024; ICLR, 2024; ACM Conference on Fairness, Accountability, and Transparency (FAccT) 2023–2024; NeurIPS Datasets & Benchmarks Track, 2021–2024; *Journal of Cybersecurity*, 2023–2024; *Journal of Official Statistics*, 2023; *Science*, 2023; International Conference on Information Systems (ICIS), 2021, 2023; *Management Science*, 2022; Workshop on Responsible Computer Vision at CVPR, 2021.

**Mentor:** CMU Undergraduate AI Mentoring Program, September 2022 – May 2023; CMU Research Experience for Undergraduates (REU), May 2022 – August 2022.