

RYAN STEED

5000 Forbes Ave., Pittsburgh P.A., 15213

ryansteed@cmu.edu • rbsteed.com

EDUCATION

Carnegie Mellon University August 2020 - Present
Ph.D., Machine Learning & Public Policy GPA: 3.99
Advised by Alessandro Acquisti

George Washington University August 2017 - May 2020
B.S. Computational Economics, Minor in Philosophy, *summa cum laude* GPA: 3.98

RESEARCH EXPERIENCE

Carnegie Mellon University August 2020 – Present
Doctoral Student
Leading empirical research intersecting privacy, ethics, and algorithmic systems to inform public policy. Also collaborating with fellows at the Mozilla Foundation on the Open Source Audit Tooling project.

OpenAI March 2022 – April 2022
Consultant
Red-teamed state-of-the-art machine learning model for bias, stereotypes, and other ethical concerns.

Oracle Labs May 2021 – August 2021
Software Engineering Intern
Scoped, motivated and lead investigation of social bias transfer from pre-trained large language models to task-specific models through fine-tuning. First-authored a research publication and software package.

George Washington University September 2018 – September 2020
Research Assistant, Department of Computer Science
With Dr. Aylin Caliskan, designed and executed experiments to study embedded social bias in computer vision. Independently, developed heuristic-based preference aggregation for participatory machine learning. With Dr. Rahul Simha, estimated the effect of policy on innovation through patent citation networks. First-authored 3 peer-reviewed publications with software packages.

Duke University May 2017 – January 2018
Research Assistant & Lead Modeler, Social Science Research Institute
Lead design/implementation of novel graph models for social network analysis. Contributed to deep learning model for classify political actions reported in the news.

OTHER EXPERIENCE

SolveForGood June 2020 – September 2020
Volunteer Data Scientist
Built an algorithmic transparency tool for the 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 \$10,000 to design an environmental education app. 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.

Magic Number

May 2018 – August 2018

Machine Learning Engineer

Researched, tested, and implemented machine learning classification and clustering microservices to categorize patent documents into packaged sectors for clients.

HONORS, AWARDS, & FELLOWSHIPS

Meta Research PhD Fellowship—Tech Policy

August 2022 - Present

Meta Platforms, Inc.

Suresh Konda Best First Student Research Paper Award

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

2nd Prize, Politics & Economics

April 2019

GWU Research Showcase Poster Competition

Data MASTER Fellowship

May 2018 - May 2019

GWU Mathematics And Statistics Training, Education, and Research Program

Sigelman Undergraduate Research Enhancement Award

October 2018, 2019

George Washington University

PUBLICATIONS

Steed, R., Liu, T., Wu, Z. S., Acquisti, A., Policy impacts of statistical uncertainty and privacy. *Science*, 377(6609):928–931, Aug. 2022. DOI: 10.1126/science.abq4481. URL: <https://www.science.org/doi/10.1126/science.abq4481>.

Caliskan, A., **Steed, R.**, Managing the risks of inevitably biased visual artificial intelligence systems. *Brookings*, Sept. 2022. URL: <https://www.brookings.edu/blog/techtank/2022/09/26/managing-the-risks-of-inevitably-biased-visual-artificial-intelligence-systems/>.

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 (Volume 1: Long Papers)*, pages 3524–3542, Dublin, Ireland. Association for Computational Linguistics, 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, pages 701–713, Virtual Event, Canada. Association for Computing Machinery, 2021. DOI: 10.1145/3442188.3445932.

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. URL: <https://doi.org/10.1007/s43681-020-00035-y>.

Steed, R., Williams, B., Heuristic-based weak learning for automated decision-making. In *Workshop on Participatory Machine Learning at the 2020 International Conference on Machine Learning*, 2020.

PRESS

- M. Heikkilä. The viral AI avatar app Lensa undressed me—without my consent. *MIT Technology Review*, Dec. 2022. URL: <https://www.technologyreview.com/2022/12/12/1064751/the-viral-ai-avatar-app-lensa-undressed-me-without-my-consent/>.
- K. Errick. Statistical Uncertainty Could be Problematic for Evidence-Based Policies, Study Finds. *Nextgov*, Aug. 2022. URL: <https://www.nextgov.com/analytics-data/2022/08/statistical-uncertainty-could-be-problematic-evidence-based-policies-study-finds/376580/>.
- C. Hackerott. Carnegie-Mellon Researchers Published Report on Statistical Uncertainty in Census Bureau Data, Sept. 2022. URL: <https://www.jdsupra.com/legalnews/carnegie-mellon-researchers-published-9893819/>.
- K. Hao. An AI saw a cropped photo of AOC. It autocompleted her wearing a bikini. *MIT Technology Review*. URL: <https://www.technologyreview.com/2021/01/29/1017065/ai-image-generation-is-racist-sexist/>.
- T. Feathers. Sexist AI is Even More Sexist Than We Thought. URL: <https://www.vice.com/en/article/y3gj3v/sexist-ai-is-even-more-sexist-than-we-thought>.
- D. Gershgorn. Men Wear Suits, Women Wear Bikinis: Image Generating Algorithms Learn Biases ‘Automatically’, Jan. 2021. URL: <https://onezero.medium.com/men-wear-suits-women-wear-bikinis-image-generating-algorithms-learn-biases-automatically-eee3d8a56f2e>.

INVITED TALKS & PRESENTATIONS

[Lecture] University of California Riverside, Data Science Ethics	February 2023
[Lecture] Carnegie Mellon University, Foundations of Privacy	November 2021, 2022
[Poster] Carnegie Mellon University, CyLab Partners Conference	October 2022
Boston Data Privacy Group	October 2022
[Oral] 2nd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization (EAAMO ‘22)	October 2022
[Poster] Theory and Practice of Differential Privacy (TPDP) Workshop at the Thirty-ninth International Conference on Machine Learning (ICML ‘22)	July 2022
Columbia University, Privacy Chats	April 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:** FAccT 2023; NeurIPS Datasets & Benchmarks Track, 2021 – Present; *Management Science*, 2022; ICIS Cybersecurity, Privacy & Ethics Track, 2021; Workshop on Responsible Computer Vision at CVPR, 2021.
- Mentor:** CMU Research Experience for Undergraduates (REU), May 2022 – August 2022; CMU Undergraduate AI Mentoring Program, September 2022 – Present.