

Rebecca Dorn

(301) 832-2668 | rdorn@usc.edu | rebedorn.github.io

EDUCATION

Pursuing Ph.D., Computer Science

University of Southern California

Advisors: Kristina Lerman and Fred Morstatter

Academic Focus: Machine Learning Fairness, Computational Social Science

August 2021-Present

GPA 3.54

B.S., Computer Science

University of California, Santa Cruz

September 2016-June 2020

GPA 3.47

RESEARCH EXPERIENCE

Community-Cross-Instruct: Unsupervised Alignment of LMs in Online Communities

January 2024-Present

University of Southern California, Information Science Institute

- Designing method for self-supervised instruction tuning for LLM alignment with online communities on Reddit and Twitter.
- Devising psychiatric survey to evaluate model alignment with pro-anorexia groups by working with public health experts.

Harmful Speech Detection by Language Models Exhibit Gender-Queer Dialect Bias

August 2023-April 2024

University of Southern California, Information Science Institute

- Designing 108 templates based on 1.1 million tweets to measure dialect bias in GPT-3.5, Llama2 and BERT toxicity scores.
- Directing team of 10 expert annotators for community-centric chain-of-thought prompting for post-hoc bias mitigation.

OATH-Frames: Characterizing Online Attitudes Towards Homelessness via LLM

January 2023-January 2024

University of Southern California, Computer Science Department

- Collaborating with social work department to make a robust homelessness-attitude framework with 4 core categories.
- Scraping 5 million tweets about homelessness, testing 50 versions of few-shot prompting to probe GPT-4 as an annotator.

Non-Binary Gender Expression in Online Interactions

August 2022-January 2023

University of Southern California, Information Science Institute

- Collected 1.1 million tweets from users with biography pronouns to examine pronoun group differences in twitter activity.
- Calculated that non-binary users receive less likes and followers online despite similar behavior to binary pronoun users.

Gender and Prestige Bias in Coronavirus News

August 2021-August 2022

University of Southern California, Information Science Institute

- Extracted 90,000 expert quotes from news articles using Named Entity Recognition (NER) and Dependency Parsing.
- Headed evaluation of gender representation, deducing that 1 woman is referenced as a coronavirus expert for every 3 men.

Expanding Healthcare Subsidies

December 2020

Families USA

- Integrated multiple Census datasets to examine demographics of uninsured populations receiving unemployment benefits.
- Led creation of informational graphics to translate technical findings for health policy experts in video calls and e-mails.
- ★ Utilized by Nancy Pelosi's office to add provisions to the American Relief Act (2021).

PUBLICATIONS

Lee, E, Baird, A., Young, L. & **Dorn, R.** *The Role of Public Facing Organizations in the Transgender Rights Debate: An Issue Management Framework*. Organizational Communication Research Escalator at ICA 2024.

Dorn, R., Kezar, L., Morstatter, F. & Lerman, K. *Harmful Speech Detection by Language Models Exhibit Gender-Queer Dialect Bias*. Poster Session, SoCalNLP 2023. Under review.

Ranjit, J., Joshi, B., **Dorn, R.**, Petry, L., Koumoundouros, O., Bottarini, J., Liu, P., Rice, E. & Swayamdipta, S. *OATH-Frames: Characterizing Online Attitudes Towards Homelessness via LLM*. Poster Session, SoCalNLP 2023. Under review.

Dorn, R., Mokhberian, N., Jiang, J. Abramson, J., Morstatter, F. & Lerman, K. *Non-Binary Gender Expression in Online Interactions*. Workshop Paper, SBP-BRIMS 2023. Under review.

Dorn, R., Ma, Y., Morstatter, F. & Lerman, K. *Gender and Prestige Bias in Coronavirus News Reporting*. Arxiv Preprint 2022.

INVITED TALKS

- Dorn, R. (2023, November). Gender-Queer Dialect Bias in Harmful Speech Detection by Large Language Models. University of Southern Information Science Institute Fairness and Bias Group. Marina Del Rey, CA.
- Dorn, R. (2023, March). Non-Binary Gender Expression in Online Interactions. University of Southern California Center for AI in Society. Los Angeles, CA.
- Dorn, R. (2022, October). Studying Gender Equity with Data Science. University of Southern California Women in Science and Engineering (WiSE) STEM Bytes. Los Angeles, CA.

TEACHING EXPERIENCE

- | | |
|---|---------------------------|
| Teaching Assistant, Machine Learning for Data Science | Summer 2024 |
| Teaching Assistant, Data Structures | Fall 2023 |
| Teaching Assistant, Fairness in Artificial Intelligence | Spring 2023 |
| Undergraduate Teaching Assistant, Ethics & Algorithms | Winter, Spring 2020 |
| Undergraduate Peer Tutor | Winter, Spring, Fall 2019 |

MEDIA COVERAGE

- Russell, Adam (Host). *“There is no end goal for AI ethics, there will always be something new to mitigate”*. In *AI/nsiders*, June 2024.

LEADERSHIP EXPERIENCE

- | | |
|--|---------------------|
| Computer Science Education Coordinator
University of Southern California, Information Science Institute | Dec. 2023-Present |
| ▪ Designed Computer Science lessons in high school summer STEM program serving historically disadvantaged communities. | |
| Education Director
Habonim Dror Camp Moshava | Jan. 2020-Aug. 2020 |
| ▪ Led over 100 staff members in building social justice educational curriculum for campers entering grades 3 through 11. | |

AWARDS

- SBP-BRIMS Scholarship, 2023
Inclusive STEM Educator Award, 2019

SERVICE

- Member:** QueerInAI (2023-present), Women in Science and Engineering (2021 – present)
External Reviewer: LREC (2024), IJHCI (2024), CSCW (2023), WebSci (2022), JMIR (2022)