

Molly Lewis

Research Scientist (PhD, Psychology) with 10+ years using quantitative methods to understand and predict human behavior. Seeking to apply my expertise in machine learning, natural language processing, and experimental design in a researcher/data scientist role. Passionate about AI ethics and safety.

Contact

- ✉ mollyllewis@gmail.com
- ☎ +1 816 419 5764
- 🏠 mllewis.github.io
- 🔄 mllewis

Skills

Programming

- R / tidyverse / Shiny
- SQL
- HTML / CSS / javascript

Tools

- Experimental design
- NLP
- Redshift
- Sagemaker
- Literate coding (e.g., R markdown)
- Git / Github

Data Analysis

- Unsupervised text analysis (e.g., topic models; LLMs)
- Network analysis (e.g., igraph)
- Frequentist statistics
- Supervised ML (e.g., regression)
- Unsupervised ML (clustering, PCA)
- Data visualization (e.g., ggplot2)

Experience

Sr. Research Scientist (Quant. Researcher)

2022 - present

Amazon, Alexa Marketing Science

remote

- Lead project examining the predictors of Alexa device success after launch
- Extracted and wrangled massive dataset (>1B rows) of user behavior using Redshift
- Analyzed unstructured text data (customer reviews) to identify themes related to customer engagement and sales

Research Scientist

2019 - 2022

Carnegie Mellon University, Social and Decision Sciences

Pittsburgh, PA

- Lead research projects, managed teams, and supervised undergraduate researchers
- Published papers on language learning and presented work as senior author
- Worked with Google's *QuickDraw* dataset to examine cross-cultural semantic variability
- Developed open source teaching materials for R-based research methods course and taught course twice [course website]

Post-Doctoral Scholar

2017-2019

University of Chicago (Advisor: James Evans, Sociology)

Chicago, IL

- Used multilingual large-scale language models to understand the principles governing cross-linguistic variability in semantic structure
- Collaborated with industry partners on analysis of 1M word corpus
- Communicated research to technical and non-technical audiences

Education

Stanford University

PhD, 2016

Psychology, Cognitive/Developmental (Advisor: Michael C. Frank)

Reed College

BA, 2009

Linguistics

Other coursework: Computational Social Science Summer School (Princeton, NJ, 2017); Machine Learning Summer School (Tübingen, Germany, 2013)

Selected Publications

I have authored 25+ publications in peer-reviewed journals and conference proceedings and given over 30 invited talks (see [website](#) for full list).

- **Gender stereotypes are reflected in the distributional structure of 25 languages** (Lewis & Lupyan, 2020; *Nature Human Behavior*) [publication, repository]
 - Through analysis of language models trained on cross-linguistic text, I found that languages with stronger gender biases in their distributional statistics tend to have speakers with stronger gender biases, as measured behaviorally. Publication featured in media outlets such as *Scientific American*.
- **What might books be teaching young children about gender?** (Lewis, et al., 2021; *Psych. Science*) [publication, repository, app]
 - Examined gender representations in corpus of popular children's books through word embedding models, behavioral data, and other NLP techniques. Publication featured in *The Capital Times* and *The KC Star*.
- **Social biases in word embeddings and their relation to human cognition.** (Caliskan & Lewis, 2021; *Atlas of Language Analysis in Psychology*) [publication]
- **Characterizing variability in shared meaning through millions of sketches** (Lewis, Balamurugan, Zheng & Lupyan, 2021). [publication]