Lucy Li

lucy3_li@berkeley.edu | 651-600-5458 | lucy3.github.io

Research interests: natural language processing (NLP) & computational social science.

Education

University of California, Berkeley

PhD. School of Information. Aug 2019 – Present

- Advisor: David Bamman.

Stanford University

MS. Computer Science. Depth area in Artificial Intelligence. Sept 2017 – June 2019

- Coursework: artificial intelligence, computer vision, computational biology, social networks, algorithms, natural language understanding, databases, logic, decision making under uncertainty.

BS. Symbolic Systems with distinction. Concentration in Natural Language. Sept 2014 – June 2018

- Study abroad at University of Oxford, Winter 2017, studied theoretical neuroscience.
- Coursework: natural language processing with deep learning, spoken language processing, mining massive datasets, semantics & pragmatics, syntax, social psychology, psycholinguistics, human behavioral biology.

Peer-reviewed Papers

* indicates equal contribution

1. **Li Lucy***, Dora Demszky*, Patricia Bromley, and Dan Jurafsky. Content Analysis of Textbooks via Natural Language Processing: Novel Findings on Gender, Race, and Ethnicity in Texas U.S. History Textbooks. *Under review,* 2020. [Finalist for best paper award at American Educational Research Association (AERA) Educational Data Science Conference.]
2. Emma Lurie, **Li Lucy**, Masha Belyi, Sofia Dewar, Daniel Rincón, John Baldwin, and Rajvardhan Oak. Investigating Causal Effects of Instructions in Crowdsourced Claim Matching. *Computation + Journalism Symposium (C+J)*, 2020. [non-archival.]
3. **Li Lucy** & Julia Mendelsohn. Using sentiment induction to understand variation in gendered online communities. *Society for Computation in Linguistics (SCiL)*, 2019.
4. **Li Lucy** & Jon Gauthier. Are distributional representations ready for the real world? Evaluating word vectors for grounded perceptual meaning. *Language Grounding for Robotics (RoboNLP) Workshop at the Association for Computational Linguistics (ACL)*, 2017.

Presentations

Content Analysis of Textbooks via Natural Language Processing. Text as Data (TADA). Stanford, CA. Oct 2019.

Grants & Awards

National Science Foundation Graduate Research Fellowships Program. April 2019
K. Jon Barwise Award for Distinguished Contributions to the Symbolic Systems Program. June 2018
Stanford Undergraduate Advising & Research (UAR) Small Grant. $1,500. May - June 2018
Symbolic Systems Grants for Education and Research (GEAR). $1,145. Aug 2017
Phi Beta Kappa, elected as junior. May 2017

Experience

Stanford Computer Science | Research Assistant | Stanford, CA Jan 2018 – Dec 2019

- Investigated the framing and representation of underrepresented groups in history textbooks with linguistics PhD student Dora Demszky.
- Advised by linguistics professor Dan Jurafsky and education professor Patricia Bromley.

École Polytechnique Fédérale de Lausanne | Research Intern | Lausanne, Switzerland July 2018 – Sept 2018

- Interned at Summer@EPFL, an international research program, with Professor Robert West in the Data Science Lab.
- Operationalized and analyzed behavioral trends in a political quote dataset using Apache Spark, emotion lexicons, Stanford CoreNLP parsers, and social networks.

Stanford Computer Science | Research Assistant | Stanford, CA April 2017 – June 2018

- Collaborated on computational social science projects with Stanford postdoctoral researcher David Jurgens, now a University of Michigan assistant professor, in the Stanford Network Analysis Project (Jure Leskovec) and NLP group (Dan Jurafsky).
- Used language and social network features to classify fictional and real relationships with scikit-learn, NLTK, and Keras.
- Built a corpus of television dialogue with 131 million words, crawled for ~1 billion tweets, and crowdsourced for annotations.

Teaching & Advising

Stanford Computer Science | Course Assistant | Stanford, CA April 2019 – June 2019

- Assisted the teaching of a graduate-level course on natural language understanding (CS 224U) and awarded a bonus for being in the top 5% of course assistants in our department.
- Crafted and taught tutorials for Jupyter notebooks and NumPy, and mentored students' final projects.

Stanford Symbolic Systems Program | Advising Fellow | Stanford, CA Sept 2016 - June 2017, Jan 2019 – June 2019

- Acted as the liaison between the program's administration and 140+ students.
- Advised prospective and current students for a total of four academic quarters.
- Planned program events, held office hours, and provided feedback on the major's curriculum requirements.

Stanford Electrical Engineering | Course Assistant | Stanford, CA Sept 2017 - Dec 2017

- Assisted the teaching of a 200-student applied linear algebra course (EE/CME 103) with professors Stephen Boyd and David Tse, and received positive student evaluations.
- Taught a weekly section covering vectors, matrices, and least squares, and their applications such as control, time-series prediction, and document analysis.

Service

Berkeley AI4ALL | Project Developer & TA Aug 2019
Stanford AI4ALL | Research Mentor June 2019 – July 2019
Girls Teaching Girls to Code | NLP Track Mentor, Lead April 2018, April 2019

Skills

Computer Languages: Python, Julia, C++, SQL
Natural Languages: English, Mandarin Chinese, French.
Tools: NLTK, Stanford CoreNLP, SpaCy, scikit-learn, Apache Spark, MTurk, Figure Eight, Keras, TensorFlow, PyTorch.