

EMILY SHENG

ewsheng at gmail · <https://ewsheng.github.io>

EDUCATION

PH.D. IN COMPUTER SCIENCE

Aug 2015 – Aug 2021

University of Southern California

Advisors: Dr. Prem Natarajan and Dr. Nanyun Peng

Research interests: fairness in natural language processing, language generation, dialogue systems

M.S. IN COMPUTER SCIENCE

Aug 2015 – May 2017

University of Southern California

B.A. IN COMPUTER SCIENCE, COGNITIVE SCIENCE

Aug 2010 – May 2014

University of California, Berkeley

PROFESSIONAL EXPERIENCE

SR SCIENTIST

May 2022 –

Microsoft Research

Seattle, WA

Identify, measure, and mitigate fairness-related harms in language technologies.

MACHINE LEARNING ENGINEER

July 2021 – May 2022

Twitter

Los Angeles, CA

Improve content health & fairness of the platform.

RESEARCH INTERN

May 2019 – April 2020

Google

Mountain View, CA

Evaluate and mitigate biases towards different demographics in a poetry composition system.

SOFTWARE ENGINEERING INTERN

May 2018 – Aug 2018

Google

Mountain View, CA

Evaluate textual similarity across text lengths for bag-of-words and convolutional neural network model variants.

SOFTWARE ENGINEER

July 2014 – July 2015

Expect Labs

San Francisco, CA

Prototype classifier for domain-specific named entity recognition to improve a natural language understanding system, and full-stack development of developer platform tools.

SELECTED PUBLICATIONS

Dev, S. *, Sheng, E. *, Zhao, J. *, Amstutz, A. *, Sun, J., Hou, Y., Sanseverino, M., Kim, J., Peng, N., Chang, K. -W. (2022). On Measures of Biases and Harms in NLP. In *Proceedings of AACL-IJCNLP-Findings 2022*.

Sheng, E. (Aug 2021). Fairness in Natural Language Generation. *PhD Thesis*.

Sheng, E., Chang, K.-W., Natarajan, P., Peng, N. (2021). Societal Biases in Language Generation: Progress and Challenges. In *Proceedings of ACL 2021*.

Sheng, E., Chang, K.-W., Natarajan, P., Peng, N. (2021). “Nice Try, Kiddo”: Investigating Ad Hominems in Dialogue Responses. In *Proceedings of NAACL 2021*.

- Sheng, E., Arnold, J., Yu, Z., Chang, K.-W., Natarajan, P., & Peng, N. (2021). Revealing Persona Biases in Dialogue Systems. *arXiv preprint arXiv:2104.08728*.
- Sheng, E., Chang, K.-W., Natarajan, P., Peng, N. (2020). Towards Controllable Biases in Language Generation. In *Proceedings of EMNLP-Findings 2020*.
- Sheng, E., & Uthus, D. (2020). Investigating Societal Biases in a Poetry Composition System. In *Proceedings of the 2nd Gender Bias in NLP Workshop*.
- Sheng, E., Chang, K.-W., Natarajan, P., Peng, N. (2019). The Woman Worked as a Babysitter: On Biases in Language Generation. In *Proceedings of EMNLP 2019*.

RESEARCH EXPERIENCE

- USC / ISI: Natural Language Processing** Aug 2015 – Aug 2021
Graduate Research Assistant at University of Southern California / Information Sciences Institute
 Evaluating and mitigating biases in language generation.
- UC Berkeley / ICSI: Resolving prepositional phrase attachment ambiguity** Jan 2014 – May 2014
Research project at University of California, Berkeley / International Computer Science Institute
 Survey of lexical, semantic, and contextual methods to resolve ambiguity (adv. Dr. Jerome Feldman).
- UC Berkeley / Walker Lab: Sleep study** June 2012 – May 2013
Research Assistant at University of California, Berkeley
 EEG, MRI, and behavioral tests to study effect of sleep on adolescents.
- UC Berkeley / Concepts and Cognition Lab: Yahoo Answers study** Aug 2011 – May 2013
Research Assistant at University of California, Berkeley
 Extract features of up-voted Yahoo Answers to find those favored in "good" explanations.

PROFESSIONAL ACTIVITY

- PUBLICITY CO-CHAIR** July 2021 – July 2022
NAACL 2022
- SEMINAR COORDINATOR** Sept 2019 – Aug 2020
USC Information Sciences Institute Natural Language Seminar
Marina del Rey, CA
- STUDENT CO-CHAIR** Sept 2019
2019 SoCal NLP Symposium
Los Angeles, CA
- REVIEWER**
*ACL, EMNLP, NAACL, AAI, *SEM, SoCal NLP Symposium, TrustNLP Workshop, RAI@KDD*

TEACHING EXPERIENCE

- TEACHING ASSISTANT** Aug 2015 – May 2016
Introduction to Computing course
University of Southern California