

Sharon Levy

Ph.D. Candidate
University of California, Santa Barbara

Email: sharonlevy@cs.ucsb.edu
Web URL: <https://sharonlevy.github.io/>

Education

- **Ph.D., Computer Science - Natural Language Processing**
University of California, Santa Barbara, 2018 - Present
Advisor: William Wang
GPA: 3.92
- **M.S., Computer Science**
University of California, Santa Barbara, 2017 - 2018
Advisor: William Wang
GPA: 3.92
- **B.S., Computer Science**
University of California, Santa Barbara, 2013 - 2017
College of Creative Studies
GPA: 3.88

Research Interests

Natural language processing; Responsible AI; Text generation; Computational social science; Question-answering

Awards and Honors

- EECS Rising Star, 2022
- AWS AI/ML Grant Proposal, \$130K, 2022
- Amazon Alexa AI Fellowship, 2020-2022
- CS Outstanding Teaching Assistant, 2020
- CRA-WP Grad Cohort for Women, 2019
- Regents Fellowship, 2018-2019
- Holbrook Fellowship, 2019
- Grace Hopper Poster Presentation, 2018
- Highest Honors (Top 2.5%), 2017

Experience

- **University of California, Santa Barbara, CA** 12/2017 – Present.
Graduate Student Researcher
- **Amazon Web Services (AWS) AI**, New York City, NY 06/2022 – 09/2022.
Applied Scientist Intern (AWS Comprehend Team), Mentors: Neha Anna John and Ling Liu
 - Analyzed social biases across multiple languages
- **Facebook**, 06/2021 – 10/2021.
Facebook AI Applied Research Intern (AI Integrity Team), Mentor: Yi-Chia Wang
 - Analyzed the interpretability of conflicts in Facebook group conversations
 - Evaluated combination of user characteristics and text-based conversation dynamics
- **Pinterest**, 06/2020 – 08/2020.
Pinterest Labs Research Intern (Ph.D., Machine Learning), Mentor: Jacob Gao
 - Machine learning models for look-alike/act-alike targeting for ads
 - Analyzed existing ads-related data and features for modeling feasibility
- **Akamai Technologies**, Santa Clara, CA 06/2017 – 09/2017.
Security Engineer Intern, Mentor: Richard Lin
 - Created and developed algorithmic improvements to system defending against SQL injections
- **University of California, Santa Barbara, CA** 06/2014 – 09/2014.
Web Developer Intern
 - Developed teacher and student interfaces for an educational research project in a Ruby on Rails environment. Project goal is to enable programming education for minority students, 4th - 6th grade
- **KLA-Tencor**, Milpitas, CA 06/2012 – 12/2012.
Software Developer Intern
 - Developed a touch user interface on iPad for augmented reality capability of a new metrology system

Publications and Preprints

1. **Sharon Levy**, Emily Allaway, Melanie Subbiah, Lydia Chilton, Desmond Patton, Kathleen McKeown and William Yang Wang. “SafeText: A Benchmark for Exploring Physical Safety in Language Models”, to appear in Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP 2022), Long Paper, ACL.
2. Alex Mei*, Anisha Kabir*, **Sharon Levy**, Melanie Subbiah, Emily Allaway, John N. Judge, Desmond Patton, Bruce Bimber, Kathleen McKeown and William Yang Wang. “Mitigating Covertly Unsafe Text within Natural Language Systems”, to appear in Findings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP 2022).
3. Matthew Ho, Aditya Sharma, Justin Chang, Michael Saxon, **Sharon Levy**, Yujie Lu and William Yang Wang. “WikiWhy: Answering and Explaining Cause-and-Effect Questions”, 2022.

4. Samhita Honnavalli*, Aesha Parekh*, Lily Ou*, Sophie Groenwold*, **Sharon Levy**, Vicente Ordonez and William Yang Wang. “Towards Understanding Gender-Seniority Compound Bias in Natural Language Generation”, to appear in Proceedings of The 13th Language Resources and Evaluation Conference (LREC 2022).
5. Kai Nakamura, **Sharon Levy**, Yi-Lin Tuan, Wenhui Chen, William Yang Wang, “HybridDialogue: An Information-Seeking Dialogue Dataset Grounded on Tabular and Textual Data”, to appear in Findings of 60th Annual Meeting of the Association for Computational Linguistics (Findings of ACL 2022), long paper, Dublin, Ireland.
6. **Sharon Levy**, Robert E. Kraut, Jane A. Yu, Kristen M. Altenburger, Yi-Chia Wang, “Understanding Conflicts in Online Conversations”, to appear in Proceedings of the ACM Web Conference 2022 (WWW 2022), online, ACM.
7. Alon Albalak, **Sharon Levy**, William Yang Wang. “Addressing Issues of Cross-Linguality in Open-Retrieval Question Answering Systems For Emergent Domains”, 2022.
8. Michael Saxon, **Sharon Levy**, Xinyi Wang, Alon Albalak and William Yang Wang, “Modeling Disclosive Transparency in NLP Application Descriptions”, to appear in Proceedings of The 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP 2021), long paper, online, ACL.
9. **Sharon Levy**, Kevin Mo, Wenhan Xiong and William Yang Wang, “Open-Domain Question-Answering for COVID-19 and Other Emergent Domains”, to appear Proceedings of The 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP 2021), demos track, online, ACL.
10. **Sharon Levy**, Michael Saxon and William Yang Wang, “Investigating Memorization of Conspiracy Theories in Text Generation”, to appear in Findings of The Joint Conference of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Findings of ACL-IJCNLP 2021), long paper, online, ACL.
11. *Sophie Groenwold, *Lily Ou, *Aesha Parekh, *Samhita Honnavalli, **Sharon Levy**, Diba Mirza and William Yang Wang, “Investigating African-American Vernacular English in Transformer-Based Text Generation”, to appear in Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP 2020), short paper, ACL.
12. **Sharon Levy**, Wenhan Xiong, Elizabeth Belding, and William Yang Wang, “SafeRoute: Learning to Navigate Streets Safely in an Urban Environment”, to appear in ACM Transactions on Intelligent Systems and Technology (ACM TIST), journal paper, ACM, 2020.
13. ***Sharon Levy**, *Kai Nakamura, and William Yang Wang, “Fakeddit: A New Multimodal Benchmark Dataset for Fine-grained Fake News Detection”, to appear in Proceedings of 12th International Conference on Language Resources and Evaluation (LREC 2020), full paper, Marseille, France, May 11-16, 2020, ELRA.
14. **Sharon Levy** and William Yang Wang, “Cross-lingual Transfer Learning for COVID-19 Outbreak Alignment”, Presented at the 1st Workshop on NLP for COVID-19 at ACL 2020.
15. *Sophie Groenwold, *Samhita Honnavalli, *Lily Ou, *Aesha Parekh, **Sharon Levy**, Diba Mirza, William Yang Wang, “Evaluating Transformer-Based Multilingual Text Classification”, 2020.

Invited Talks

1. Johns Hopkins University CLSP, 2023

2. Stanford University, 2023
3. UT Austin NLP Group, 2022
4. UCSB CS 165B, *Introduction to Machine Learning*, 2022
5. Fakespeak Workshop, University of Oslo, 2021
6. UCSB INT 200, *Seminar in Information Technology & Society*, 2021
7. SBCC Computer Science Club, 2021.

High School & Undergraduate Student Mentoring

- Ksenia Zhizhimontova (Cornell BS, 2019)
- Kai Nakamura (High school/Caltech BS, 2019-Present)
- Sophie Groenwold (UCSB BS, 2019-2021)
- Samhita Honnavalli (UCSB BS, 2019-2021, CRA Outstanding Undergraduate Researcher Award Honorable Mention)
- Lily Ou (UCSB BS, 2019-2021)
- Aesha Parekh (UCSB BS, 2019-2021, Chancellor's Award in Undergraduate Research, CRA Outstanding Undergraduate Researcher Award Finalist)
- Kevin Mo (Princeton BS, 2021)
- Nga Ngo (UCSB BS, 2021-2022)
- Aditya Sharma (UCSB BS, 2021-2022)
- Justin Chang (UCSB BS, 2021-2022)
- Matthew Ho (UCSB BS, 2021-2022)
- Alex Mei (UCSB BS/MS, 2022-Present)
- Anisha Kabir (UCSB BS, 2022)
- John Judge (UCSB BS, 2022)

Teaching Experience

- UC Santa Barbara, Fall 2019
Teaching Assistant, *Introduction to Machine Learning: Upper-division*, 80 students.

Service

- SoCalNLP, 2022
- ACL Rolling Review, 2022