

RESEARCH INTERESTS

I study the ethical and social implications of language technologies, focusing on the complexities of language and language technologies in their social contexts, and on supporting NLP practitioners in their ethical work.

EDUCATION

University of Massachusetts Amherst

Ph.D., Computer Science	2021
M.S., Computer Science	2018

Wellesley College

B.A., Mathematics	2015
-------------------	------

RESEARCH EXPERIENCE

Senior Researcher. Microsoft Research, Montréal.	2022 –
Postdoctoral Researcher.	2020 – 2022
Fairness, Accountability, Transparency, & Ethics (FATE) Group.	

Research Intern. Microsoft Research, New York.	Summer 2019
Mentors: Solon Barocas, Hal Daumé III, Hanna Wallach.	

Graduate Research Assistant. University of Massachusetts Amherst.	2015 – 2020
Advisor: Brendan O'Connor.	

HONORS AND AWARDS

100 Brilliant Women in AI Ethics	2022
ACL Outstanding Reviewer Award	2021
EMNLP Outstanding Reviewer Award	2020
ACL Outstanding Reviewer Award	2020
NSF Computing Innovation Fellowship (declined)	2020
Rising Stars in EE/CS Workshop Participant (fully funded travel & arrangements)	2019
ICWSM Best Reviewer Award	2019
Microsoft Accomplishments in Search & Artificial Intelligence Award for UMass CICS	2019
UMass CICS Outstanding Synthesis Award	2019
New England Future Faculty Workshop Participant	2018
Passed Ph.D. Qualifying Exam (Portfolio) with Distinction	2018
NSF Graduate Research Fellowship	2017
W-NUT Best Paper Award	2017
CRA-Women Grad Cohort Workshop Participant (fully funded travel & arrangements)	2016
UMass CICS Paul Utgoff Memorial Scholarship (for one student in machine learning)	2016

Full Conference Publications (Peer-Reviewed)

17. ECBD: Evidence-Centered Benchmark Design for NLP.
Yu Lu Liu, **Su Lin Blodgett**, Jackie Chi Kit Cheung, Q. Vera Liao, Alexandra Olteanu, and Ziang Xiao. ACL. 2024.
16. Understanding the Impacts of Language Technologies’ Performance Disparities on African American Language Speakers.
Jay L. Cunningham, **Su Lin Blodgett**, Hal Daumé III, Christina Harrington, Hanna Wallach, and Michael A. Madaio. Findings of ACL. 2024.
15. “*One-size-fits-all?*” Examining Expectations about What Constitute “Fair” or “Good” NLG System Behaviors.
Li Lucy, **Su Lin Blodgett**, Milad Shokouhi, Hanna Wallach, and Alexandra Olteanu. NAACL. 2024.
14. The Perspectivist Paradigm Shift: Assumptions and Challenges of Capturing Human Labels.
Eve Fleisig, **Su Lin Blodgett**, Dan Klein, and Zeerak Talat. NAACL. 2024.
13. Responsible AI Considerations in Text Summarization Research: A Review of Current Practices.
Yu Lu Liu, Meng Cao, **Su Lin Blodgett**, Jackie Chi Kit Cheung, Alexandra Olteanu, and Adam Trischler. Findings of EMNLP. 2023.
12. FairPrism: Evaluating Fairness-Related Harms in Text Generation.
Eve Fleisig, Aubrie N. Amstutz, Chad Atalla, **Su Lin Blodgett**, Hal Daumé III, Alexandra Olteanu, Emily Sheng, Dan Vann, and Hanna Wallach. ACL. 2023.
11. It Takes Two to Tango: Navigating Conceptualizations of NLP Tasks and Measurements of Performance.
Arjun Subramonian, Xingdi Yuan, Hal Daumé III, and **Su Lin Blodgett**. Findings of ACL. 2023.
10. This Prompt is Measuring <MASK>: Evaluating Bias Evaluation in Language Models.
Seraphina Goldfarb-Tarrant, Eddie L. Ungless, Esma Balkır, and **Su Lin Blodgett**. Findings of ACL. 2023.
9. Taxonomizing and Measuring Representational Harms: A Look at Image Tagging.
Jared Katzman, Angelina Wang, Morgan Scheuerman, **Su Lin Blodgett**, Kristen Laird, Hanna Wallach, and Solon Barocas. AAAI. 2023.
8. Examining Responsibility and Deliberation in AI Impact Statements and Ethics Reviews.
David Liu, Priyanka Nanayakkara, Sarah Sakha, Grace Abuhamad, **Su Lin Blodgett**, Nicholas Diakopoulos, Jessica Hullman, and Tina Eliassi-Rad. AIES. 2022.
7. Deconstructing NLG Evaluation: Evaluation Practices, Assumptions, and Their Implications.
Kaitlyn Zhou, **Su Lin Blodgett**, Adam Trischler, Hal Daumé III, Kaheer Suleman, and Alexandra Olteanu. NAACL. 2022.
6. Stereotyping Norwegian Salmon: An Inventory of Pitfalls in Fairness Benchmark Datasets.
Su Lin Blodgett, Gilsinia Lopez, Alexandra Olteanu, Robert Sim, and Hanna Wallach. ACL. 2021.
5. A Survey of Race, Racism, and Anti-Racism in NLP.
Anjalie Field, **Su Lin Blodgett**, Zeerak Talat, and Yulia Tsvetkov. ACL. 2021.
4. Language (Technology) is Power: A Critical Survey of “Bias” in NLP.
Su Lin Blodgett, Solon Barocas, Hal Daumé III, and Hanna Wallach. ACL. 2020.
3. Twitter Universal Dependency Parsing for African-American and Mainstream American English.
Su Lin Blodgett, Johnny Wei, and Brendan O’Connor. ACL. 2018.

2. Monte Carlo Syntax Marginals for Exploring and Using Dependency Parses.
Katherine Keith, **Su Lin Blodgett**, and Brendan O'Connor. NAACL. 2018.
1. Demographic Dialectal Variation in Social Media: A Case Study of African-American English.
Su Lin Blodgett, Lisa Green, and Brendan O'Connor. EMNLP. 2016.

Workshop Publications (Lightly Peer-Reviewed)

4. Examining Political Rhetoric with Epistemic Stance Detection.
Ankita Gupta, **Su Lin Blodgett**, Justin H. Gross, and Brendan O'Connor. Workshop on Natural Language Processing and Computational Social Science (NLP+CSS). 2022.
3. A Dataset and Classifier for Recognizing Social Media English.
Su Lin Blodgett, Johnny Wei, and Brendan O'Connor. Workshop on Noisy User-generated Text (W-NUT). **W-NUT Best Paper Award**. 2017.
2. Racial Disparity in Natural Language Processing: A Case Study of Social Media African-American English.
Su Lin Blodgett and Brendan O'Connor. Fairness, Accountability, and Transparency in Machine Learning Workshop (FAT/ML). 2017.
1. Visualizing Textual Models with In-Text and Word-as-Pixel Highlighting.
Abram Handler, **Su Lin Blodgett**, and Brendan O'Connor. Workshop on Human Interpretability in Machine Learning (WHI). 2016.

Non-Archival Presentations (Lightly Peer-Reviewed)

6. How different mental models of AI-based writing assistants impact writers' interactions with them
Shalaleh Rismani, **Su Lin Blodgett**, Alexandra Olteanu, Q. Vera Liao, and AJung Moon. Workshop on Intelligent and Interactive Writing Assistants (In2Writing). 2024.
5. Measuring Machine Learning Harms from Stereotypes: Requires Understanding Who is Being Harmed by Which Errors in What Ways.
Angelina Wang, Xuechunzi Bai, Solon Barocas, and **Su Lin Blodgett**. Conference on Equity and Access in Algorithms, Mechanisms, and Optimization (EAAMO). 2023.
4. Representational Harms in Image Captioning and Labeling.
Jared Katzman, Solon Barocas, **Su Lin Blodgett**, Kristen Laird, and Hanna Wallach. Beyond Fairness: Towards a Just, Equitable, and Accountable Computer Vision Workshop. 2021.
3. Debunking Debiasing: A Critique of Bias Measurement in NLP.
Su Lin Blodgett, Solon Barocas, Hal Daumé III, and Hanna Wallach. Women in Machine Learning Workshop (WiML). 2019.
2. Debunking Debiasing: A Critique of Bias Measurement in NLP.
Su Lin Blodgett, Solon Barocas, Hal Daumé III, and Hanna Wallach. Conference on New Directions in Analyzing Text as Data (TADA). 2019.
1. Natural Language Processing Methods to Learn about Social Determinants of Health from Electronic Health Records.
David L. Chin, Vishal A. Shetty, **Su Lin Blodgett**, and Brendan T. O'Connor. AcademyHealth Annual Research Meeting. 2019.

Book Chapters

1. Beyond "Fairness": Structural Injustice Lenses On AI for Education.
Michael Madaio, **Su Lin Blodgett**, Elijah Mayfield, and Ezekiel Dixon-Román. Invited chapter in *The Ethics of Artificial Intelligence in Education: Current Challenges, Practices and Debates*, Wayne Holmes and Kaśka Porayska-Pomsta (Eds.), Routledge. 2022.

Other

4. Evaluating the Social Impact of Generative AI Systems in Systems and Society. Irene Solaiman, Zeerak Talat, William Agnew, Lama Ahmad, Dylan K. Baker, **Su Lin Blodgett**, Hal Daumé III, Jesse Dodge, Ellie Evans, Sara Hooker, Yacine Jernite, Alexandra Sasha Luccioni, Alberto Lusoli, Margaret Mitchell, Jessica Newman, Marie-Therese Png, Andrew Strait, and Apostol Vassilev. arXiv. 2023.
3. Fairness and Sequential Decision Making: Limits, Lessons, and Opportunities. Samer B. Nashed, Justin Svegliato, and **Su Lin Blodgett**. arXiv. 2023.
2. Risks of AI Foundation Models in Education. **Su Lin Blodgett** and Michael Madaio. arXiv. 2021.
1. How to Write a Bias Statement: Recommendations for Submissions to the Workshop on Gender Bias in NLP. Christian Hardmeier, Marta R. Costa-jussà, Kellie Webster, Will Radford, and **Su Lin Blodgett**. arXiv. 2021.

WORKSHOPS AND TUTORIALS ORGANIZED

6. Tutorial: Human-Centered Evaluation of Language Technologies. **Su Lin Blodgett**, Jackie Chi Kit Cheung, Q. Vera Liao, and Ziang Xiao. EMNLP. 2024.
5. Third Workshop on Bridging Human–Computer Interaction and Natural Language Processing. **Su Lin Blodgett**, Amanda Cercas Curry, Sunipa Dev, Michael Madaio, Ani Nenkova, Ziang Xiao, and Diyi Yang. NAACL. 2024.
4. Panel: Responsible Language Technologies: Foreseeing and Mitigating Harms. **Su Lin Blodgett**, Q. Vera Liao, and Alexandra Olteanu. CHI. 2022.
3. Second Workshop on Bridging Human–Computer Interaction and Natural Language Processing. **Su Lin Blodgett**, Hal Daumé III, Michael Madaio, Ani Nenkova, Brendan O’Connor, Hanna Wallach, and Qian Yang. NAACL. 2022.
2. First Workshop on Bridging Human–Computer Interaction and Natural Language Processing. **Su Lin Blodgett**, Michael Madaio, Brendan O’Connor, Hanna Wallach, and Qian Yang. EACL. 2021.
1. Tutorial: The Meaning and Measurement of Bias: Lessons from NLP. Abigail Z. Jacobs, **Su Lin Blodgett**, Solon Barocas, Hal Daumé III, and Hanna Wallach. FAccT. 2020.

INVITED TALKS AND PANELS

“Two Challenges for Equitable Language Technologies”. Responsible AI Seminar Series, Nokia Bell Labs	2024
“Ethical reasoning throughout the AI lifecycle”. Responsible AI & Human Rights Summer School, Mila	2024
“Two Challenges for Equitable Language Technologies”. Workshop on AI Driven Online Advertising: Market Design, Generative AI, and Ethics, WWW	2024
Invited panelist. Panel on the Future of Work in the Age of LLMs, Workshop on Trust and Reliance in Evolving Human-AI Workflows, CHI	2024
Invited panelist. Panel on Generative AI and Biases, Queer in AI Workshop, NeurIPS	2023
“Towards Equitable Language Technologies”. AI-GAP: Workshop on Algorithmic Biases in Artificial Intelligence from Interdisciplinary Perspectives, University of L’Aquila	2023

———. Social, Ethical and Legal Issues in Computing Lecture Series, Queens University	2023
———. Language Technologies Institute Colloquium, Carnegie Mellon University	2023
Invited panelist. Gender & Tech Event, University of Cambridge	2023
“Measurement and Disagreement in Equitable NLP”. Linguistics Colloquium, University of Québec at Montréal	2023
“Challenges for Subjectivity in NLP”. Keynote, 7th Workshop on Online Abuse and Harms (WOAH), ACL	2023
Invited panelist. Panel on algorithmic justice, SPUI25	2023
“Towards Equitable Language Technologies”. Keynote, Workshop on Algorithmic Injustice, University of Amsterdam	2023
Invited panelist. CRAFT Session on User Engagement in Algorithm Testing and Auditing: Exploring Opportunities and Tensions between Practitioners and End Users, FAccT	2023
“Examining How We Examine Language Technologies”. Software Engineering for Machine Learning Applications (SEMLA), Polytechnique Montréal	2023
Invited panelist. Panel on ChatGPT and Education: Hype versus Reality. University of Massachusetts Amherst	2023
“Towards Equitable Language Technologies”. School of Information Colloquium, University of Arizona	2022
———. NSERC Responsible AI Seminar Series, Toronto Metropolitan University	2022
“Prospects for Participatory Approaches in NLP”. Keynote, 1st Workshop on Perspectivist Approaches to NLP, LREC	2022
“Fairness-Related Harms in Language Technologies”. Emerging Problems: New Challenges in FAccT from Research, to Practice, to Policy, FAccT CRAFT Session	2022
“Towards Equitable Language Technologies”. Keynote, Second Workshop on Speech and Language Technologies for Equality, Diversity, and Inclusion (LT-EDI), ACL	2022
———. Computational Linguistics Lunch Seminar Series, University of Pennsylvania	2022
———. Young Achievers Symposium, Center for Socially Responsible Artificial Intelligence, Pennsylvania State University	2022
———. Linguistics Colloquium Series, University of Washington	2022
“Towards Building Equitable Language Technologies”. Cultural Assessment Working Group, Digital Library Federation	2021
“Towards Equitable Language Technologies”. Computer Science Colloquium, Williams College	2021
“Language and Justice: Social Implications of Language Technologies”. Bocconi University	2021
“Computational Approaches to Measuring Stereotyping in Language”. Workshop on Generics and Stereotypes in Language, Institute for Language, Logic and Computation, University of Amsterdam	2021
“Towards Building Equitable Language Technologies”. Responsible Data Science and AI Speaker Series, University of Illinois at Urbana-Champaign	2021
Invited panelist. Workshop on Foundation Models, Stanford Institute for Human-Centered AI (HAI)	2021

“Towards Building Equitable Language Technologies”. xAI Discussion Group, Aggregate Intellect	2021
Invited panelist, Ethics in AI Panel. Level 3 AI Assistant Conference (L3-AI)	2021
“Towards Building Equitable Language Technologies”. Annual Linguistics Student Association Spring Colloquium, San Diego State University	2021
———. Artificial Intelligence in the Liberal Arts, Amherst College	2021
“Language and Justice: Reconsidering Harms in Language Systems and Practices”. Seminar Series, Naver Labs	2021
———. NLP/Text-as-Data Speaker Series, New York University	2021
———. AI Ethics Group, MIT	2021
———. NLP Seminar, University of California, Berkeley	2021
Invited panelist, Voices of Data Science at UMass Amherst	2021
“Language and Justice: Reconsidering Harms in Language Systems and Practices”. NLP Seminar, Stanford University	2021
“Language and Justice: Understanding Harms in NLP Systems and Practices”. Computer Science & Engineering Colloquia Series, Washington University in St. Louis	2020
“Language and Justice: A Collection of Thoughts on the Social Implications of Language Technologies”. NLP with Friends	2020
“Language and Justice: Understanding Harms in NLP Systems and Practices”. Educational Testing Service	2020
“Sociolinguistically Driven Approaches for Justice in Natural Language Processing Systems”. Microsoft Research New York	2020
“Fairness and Justice in Natural Language Processing”. Womxn in Computer Science, Dartmouth College	2019
“Demographic Dialectal Variation in Social Media”. Computational Social Science Institute Seminar, University of Massachusetts Amherst	2019
———. Workshop on New Methods in Computational Sociolinguistics, Lorentz Center, Leiden University	2018
———. Summer Dialect Research Project, Department of Linguistics, University of Massachusetts Amherst	2018
———. Data Science Research Symposium, University of Massachusetts Amherst	2017
———. Computational Linguistics Workshop, Department of Linguistics, University of Massachusetts Amherst	2017
———. Computer Science Colloquium, Wellesley College	2016

TEACHING AND MENTORING

Instructor. Ethical Issues Surrounding Artificial Intelligence Systems and Big Data. Co-designed and co-taught the course, a computer science first year seminar.	Fall 2018
Teaching Assistant. Advanced Natural Language Processing (graduate).	Spring 2017

Guest Lectures

“Language, Society, and Language Technologies”. Carnegie Mellon University	2023
“Language, Language Technologies, and Justice”. Georgia Institute of Technology	2022
“Towards Equitable Language Technologies”. University of Massachusetts Amherst	2021
“Language and Justice”. University of Wisconsin-Madison	2020
“NLP and Social Good”. University of Massachusetts Amherst	2020
“Ethics in NLP”. University of Massachusetts Amherst	2020
———. Mount Holyoke College	2020
“Ethics in Machine Learning”. Mount Holyoke College	2020
“Variation in Social Media”. Mount Holyoke College	2017
———. University of Massachusetts Amherst	2017

Mentoring

Gaurav Verma, Georgia Institute of Technology, MSR research internship	2023
Yu Lu Liu, McGill University, Mila and MSR project	2022 –
Jules Barbe, McGill University, Mila and MSR project	2021 – 2022
Lucy Li, University of California, Berkeley, MSR research internship	2021
Kaitlyn Zhou, Stanford University, MSR research internship	2021
Jay Cunningham, University of Washington, MSR research internship	2021

Ph.D. Committees

Akshita Jha, Virginia Tech
Zhila Aghajari, Lehigh University
Marta Marchiori Manerba, University of Pisa

SERVICE AND OUTREACH

Committees and Reviewing

Co-Chair of the Student Program Committee, AIES 2023

Co-Chair of the Ethical Advisory Committee, ACL 2022

Member of the Advisory Board, Socially Responsible Language Modelling Research (SoLaR) Workshop 2023

Area Chair: FAccT 2024, EMNLP 2023, AACL 2022, Black in AI Workshop 2021

Conference/journal reviewer: ACL, ACL Rolling Review (ARR), CoNLL, EACL, EMNLP, NAACL, AIES, FAccT, ICWSM, The Web Conference, ACM Computing Surveys, Nature, PLOS One, TACL

Workshop reviewer: Women in Machine Learning, Workshop on Noisy User-generated Text, Teaching NLP, NLP for Positive Impact, Workshop on Online Abuse and Harms, Gender Bias in Natural Language Processing, Data Science with Human in the Loop

Ethics reviewer: ARR, EMNLP, NeurIPS

Other

Application Guidance Panelist, Women@Mila	2023
Podcast guest, Unhinged Collaboration Podcast	2023
Moderator, Fairness at the Intersections Tutorial Panel (NeurIPS)	2022
Panelist, UMass CICS CS Women	2022
Podcast guest, McGill AI Podcast	2022
Podcast guest, The Good Robot	2022
Mentoring group leader, EACL	2021
Podcast guest, Rasa Chats	2021
Podcast guest, Radical AI Podcast	2021
Mentoring group leader, EMNLP	2020
Ph.D. Applicant Support Program mentor, UMass CICS	2020
Volunteer, Women in Machine Learning Workshop (WiML)	2019
Senior Ph.D. Student Panelist, UMass CICS CS Women	2019
Succeeding in Graduate School Panelist, UMass Graduate Orientation	2018
Volunteer, ACL	2018
Co-organizer, Inaugural Male Allyship Workshop, UMass Computer Science	2017
Panelist, Mount Holyoke Graduate School Panel	2017
Social Media Chair, UMass CS Women	2016 – 2017
Mentoring Committee Member, UMass Graduate Women in STEM	2016 – 2017
Volunteer, EMNLP	2016
Volunteer, UMass Women in Engineering and Computing Career Day	2015