

Sarah A. Wiegreffe

Atlanta, GA
✉ saw@gatech.edu
📁 sarahwie.github.io

Education

- 2017–May **Georgia Institute of Technology**, *Ph.D. in Computer Science*.
2022 Advisor: Taetle Chair Professor Mark Riedl.
(expected) Thesis Committee: Professors Alan Ritter, Wei Xu, Noah Smith (University of Washington), and Sameer Singh (University of California Irvine).
- 2017–2020 **Georgia Institute of Technology**, *M.S. in Computer Science*.
Specialization: Machine Learning.
Relevant coursework: Computational Statistics, Statistical Machine Learning, Deep Learning, Natural Language Processing.
- 2013–2017 **Honors College at the College of Charleston**, *B.S. in Data Science*.
Summa Cum Laude.
Awarded Data Science Major of the Year and Departmental Honors.
Minors in Mathematics and International Studies.
- 2015 **University of Tartu**, *Estonia*.
Visiting student in the Faculty of Mathematics and Computer Science.
Coursework: Cryptology, Computational Neuroscience, Advanced French (European scale B2→C1).

Publications

Acceptance rates listed where known. * denotes equal contribution.

Preprints

11. **Sarah Wiegreffe**, Jack Hessel, Swabha Swayamdipta, Mark Riedl, Yejin Choi. *Reframing Human-AI Collaboration for Generating Free-Text Explanations*. 2021.

Peer-reviewed, archival

10. **Sarah Wiegreffe***, Ana Marasović*. *Teach Me to Explain: A Review of Datasets for Explainable Natural Language Processing*. Conference on Neural Information Processing Systems (NeurIPS) Datasets and Benchmarks Track 2021. Online. Acceptance rate 38%.
9. **Sarah Wiegreffe**, Ana Marasović, Noah A. Smith. *Measuring Association Between Labels and Rationales*. Conference on Empirical Methods in Natural Language Processing (EMNLP) 2021. Punta Cana, Dominican Republic. Oral Presentation. Acceptance rate 23.4% (8.8% oral presentations).
8. Xiangyu Peng*, Siyan Li*, **Sarah Wiegreffe**, Mark Riedl. *Inferring the Reader: Guiding Automated Story Generation with Commonsense Reasoning*. Narrative Understanding Workshop at the North American Chapter of the Association for Computational Linguistics (NAACL) 2021. Online.

7. Sarthak Jain, **Sarah Wiegrefe**, Yuval Pinter, Byron C. Wallace. *Learning to Faithfully Rationalize by Construction*. Annual Meeting of the Association for Computational Linguistics (ACL) 2020. Online. Acceptance rate 22.7%.
6. **Sarah Wiegrefe***, Yuval Pinter*. *Attention is not not Explanation*. Conference on Empirical Methods in Natural Language Processing and the International Joint Conference on Natural Language Processing (EMNLP-IJCNLP) 2019. Hong Kong. Acceptance rate 24% (7% oral presentations).
5. **Sarah Wiegrefe**, Edward Choi, Sherry Yan, Jimeng Sun, Jacob Eisenstein. *Clinical Concept Extraction for Document-Level Coding*. Biomedical Natural Language Processing Workshop (BioNLP) at the Annual Meeting of the Association for Computational Linguistics (ACL) 2019. Florence, Italy.
4. James Mullenbach, **Sarah Wiegrefe**, Jon Duke, Jimeng Sun, Jacob Eisenstein. *Explainable Prediction of Medical Codes from Clinical Text*. Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT) 2018. New Orleans, LA. Acceptance rate 31% (oral presentation).

[Peer-reviewed, non-archival \(poster presentations\)](#)

3. Xiangyu Peng*, Siyan Li*, **Sarah Wiegrefe**, Mark Riedl. *Improving Neural Storytelling with Commonsense Inferences*. Women in Machine Learning (WiML) workshop at the Conference on Neural Information Processing Systems (NeurIPS) 2020. Online.
2. **Sarah Wiegrefe***, Yuval Pinter*. *Attention is not not Explanation*. Women in Machine Learning (WiML) workshop at the Conference on Neural Information Processing Systems (NeurIPS) 2019. Vancouver, Canada.
1. **Sarah Wiegrefe**, Jihad Obeid, Paul Anderson. *Can Classification of Publications by Translational Categories be Automated?* American Medical Informatics Association (AMIA) Translational Bioinformatics Summit 2017. San Francisco, CA.

Honors and Awards

- 2019-2021 **Travel Award**, *Georgia Tech Graduate Student Government Organization*.
Funded attendance at EMNLP 2019, ACL 2020, EMNLP 2020, ICLR 2021.
- 2020 **Outstanding Intern**, *Allen Institute for Artificial Intelligence*.
Gift of \$10,000 and returning offer. Awarded to 2-3 interns per year by research mentor nomination.
- 2019-2020 **Travel Award**, *Women in Machine Learning (WiML) Organization*.
Funded attendance at NeurIPS 2019, NeurIPS 2020.
- 2019 **Travel Award**, *Neural Information Processing (NeurIPS) Organization*.
Funded attendance at NeurIPS 2019.
- 2019 **Travel Award**, *Georgia Tech School of Interactive Computing*.
Specially funded attendance at EMNLP 2019.
- 2018-2019 **Travel Award**, *Georgia Tech College of Computing*.
Funded attendance at NAACL 2018, EMNLP 2019.

- 2018 **Graduate Cohort Member**, *ACM Computing Research Association*.
Sponsored to attend the Association for Computing Machinery (ACM)'s national workshop for female computing PhD students.
- 2017 **Graduate Fellowship**, *Phi Kappa Phi Honor Society*.
Gift of \$5,000. Awarded to 51 students nationwide beginning doctoral studies.
- 2017 **Data Science Major of the Year**, *College of Charleston*.
One student selected per academic year.
- 2017 **Departmental Honors**, *College of Charleston*.
Awarded for completion of a minimum of 12 credit hours of "exceptionally fine work" in the Data Science major, including an exceptional 6 credit hour bachelor's essay.
- 2016 **Grace Hopper Scholar**, *Anita Borg Institute*.
Sponsored by Intel to attend the largest annual conference for women in computing.
- 2015 **Harry and Reba Huge Scholarship**, *College of Charleston*.
Gift of \$5,400 towards study abroad. Awarded to 3 students per year. Additional \$500 given by the College of Charleston's School of Languages, Cultures and World Affairs.
- 2015 **Diploma of French Language Studies: Level B2**, *French Ministry of Education*.
Passed standardized oral and written exams. Recognized as having obtained fluency by the French government, sufficient for enrollment in French universities.
- 2014 **Crosby Computer Science Endowment Award**, *College of Charleston*.
Given to the most promising students in an introductory computer science course by professor nomination. Awarded to 2 students per year.
- 2013-2017 **Charleston Fellow**, *College of Charleston*.
Gift of \$92,000 toward tuition and fees. Awarded by competitive interview process to less than 0.01% of students at the university.
- 2013-2017 **International Scholar**, *College of Charleston*.
10 students selected from the Charleston Fellows. Gift of \$10,000 towards study abroad.

Selected Talks

- 2021 **Can Large Language Models Explain their Predictions?**, *Allen Institute for AI Internal Company-Wide Meeting*.
- 2021 **Measuring Association Between Labels and Free-Text Rationales (pre-recorded video)**, *EMNLP 2021*.
- 2021 **Measuring Association Between Labels and Free-Text Rationales (live; recorded)**, *NLP with Friends seminar*, Online.
- 2020 **BlackBoxNLP: What are we looking for, and where do we stand? (live; recorded)**, *NLP/ISI seminar*, University of Southern California.
- 2019 **Attention is not not Explanation (live; recorded)**, *EMNLP 2019*.
- 2019 **Transformers and Natural Language Applications**, *Guest lecture*, graduate deep learning course at Georgia Tech.
- 2019 **Self Attention for Universal Representations of Clinical Events**, *Final internship presentation*, Google AI.

Professional Experience

Nonprofits and Industry

- Summer-Fall 2021 **Research Intern, Allen Institute of Artificial Intelligence.**
Hosted by Drs. Jack Hessel and Swabha Swayamdipta, and Professor Yejin Choi. Worked on few-shot explanation generation and effective human evaluation.
- Summer-Fall 2020 **Research Intern, Allen Institute of Artificial Intelligence.**
Hosted by Dr. Ana Marasović and Professor Noah Smith. Worked on interpretability of deep learning models for NLP. **Awarded outstanding intern award.**
- Summer 2019 **Research Intern, Google AI Health (formerly/now Google Brain).**
Hosted by Dr. Edward Choi (now assistant professor at KAIST), Gerardo Flores, and Dr. Andrew Dai. Improved outcome prediction for clinical time-series data using unsupervised pretraining. Resulted in unpublished short paper *Learning Bi-Directional Clinical Event Representations: a Comparison of Architectures* (available upon request).
- Summer-Fall 2018 **Research Intern, Sutter Health.**
Hosted by Dr. Sherry Yan and Professor Jimeng Sun (Georgia Tech). Worked on deep learning methodology for disease prediction from clinical text.

Academia

- Spring 2020-Present **Research Assistant, Entertainment Intelligence/Human-Centered AI Lab, Georgia Tech.**
Advised by Professor Mark Riedl on research problems centered around interpreting NLP systems with applications to text generation and commonsense reasoning.
- Fall 2017-Spring 2019 **Research Assistant, Computational Linguistics Lab, Georgia Tech.**
Advised by Professor Jacob Eisenstein on problems such as convex optimization for incorporating lexical semantics in word embeddings and representation learning for clinical notes.
- Spring 2016-Spring 2017 **Research Assistant, Anderson Lab, College of Charleston.**
Advised by Professor Paul Anderson on word embeddings used directly as document-level classifiers. Resulted in Bachelor's Essay "Word2Vec Inversion Methods in Topic Recognition Tasks" presented in May 2017.

Teaching

Assistantships

- Fall 2021 **Natural Language Processing (CS 7643), Georgia Tech,** 91 students.
- Spring 2021 **Deep Learning (CS 4803/7643), Georgia Tech,** 170 students.
- Fall 2019 **Deep Learning (CS 4803/7643), Georgia Tech,** 215 students.
- Spring 2019 **Machine Learning (CS 4641), Georgia Tech,** 110 students.

Mentoring

- Spring 2021-present **Kaige Xie, Machine Learning PhD student at Georgia Tech.**
- Fall 2020-Spring 2021 **Xiangyu Peng, Machine Learning PhD student at Georgia Tech.**
and Siyan Li, undergraduate student at Georgia Tech → now M.S. student at Stanford.
Met weekly. Resulted in an extended abstract at a NeurIPS workshop (2020) and a full paper submission (2021).

Academic Service

Organization

- Workshop Organizer: *BlackBoxNLP 2022*
- Publicity Co-Chair: *NAACL 2021*
- Student Volunteer: *EMNLP 2019, FAT* 2019, NAACL 2018*

Conference Reviewing

- ARR: *November 2021 - present; monthly*
- NAACL: *2021*
- EMNLP: *2019, 2020, 2021*
- ACL: *2018 (subreviewer), 2019, 2020*
- AMIA Informatics: *2018, 2019*

Workshop Program Committees

- Deep Learning Approaches for Low-Resource NLP (NAACL): *2022*
- Commonsense Representation and Reasoning (ACL): *2022*
- BlackBoxNLP (EMNLP): *2020, 2021*
- Machine Learning for Healthcare (NeurIPS): *2017, 2018, 2019*

Other

- Reviewer, Georgia Tech PhD Application Support Program for underrepresented applicants: *2021*