

# Sarah Wiegreffe

Atlanta, GA  
swiegreffe6@gatech.edu  
sarahwie.github.io

## EDUCATION

---

**Georgia Institute of Technology (Georgia Tech)**  
**PhD in Computer Science**  
**Passed qualifying exams: Spring 2019**

**August 2017 - present**

Research interest in natural language processing and machine learning with clinical applications. Specific interests include machine reading and information extraction, hierarchical representations of text, and text generation/summarization. Interested in temporal/hierarchical modeling of hospital clinical notes for the prediction of patient diagnosis and treatment.

**Honors College at the College of Charleston**  
**Bachelor of Science in Data Science, Summa Cum Laude**  
Minors in Mathematics and International Studies

**August 2013 - May 2017**

**University of Tartu, Estonia**  
Visiting Student, Faculty of Mathematics and Computer Science

**January - June 2015**

## PUBLICATIONS

---

**Wiegreffe, S.**, Flores, G., Choi, E., and Dai, A. *Learning Bi-Directional Clinical Event Representations: a Comparison of Architectures*. In submission at NeurIPS ML4Health Workshop 2019.

**Wiegreffe, S.** and Pinter, Y. (equal contribution). *Attention is Not Not Explanation*. EMNLP. **Oral presentation**. Hong Kong, November 2019.

**Wiegreffe, S.**, Choi, E., Yan, S., Sun, J. and Eisenstein, J. *Clinical Concept Extraction for Document-Level Coding*. ACL BioNLP Workshop. Florence, Italy, July 2019.

Mullenbach, J., **Wiegreffe, S.**, Duke, J., Sun, J. and Eisenstein, J. *Explainable Prediction of Medical Codes from Clinical Text*. NAACL-HLT. **Oral Presentation**. New Orleans, LA, June 2018.

**Wiegreffe, S.**, Anderson, P. and Obeid, J. *Can Classifications of Publications by Translational Categories be Automated?*. Proceedings of the American Medical Informatics Association (AMIA) Joint Summits on Translational Science. San Francisco, CA, March 2017.

## RESEARCH

---

**Research Intern, Google AI Health**

**May - August 2019**

Working with Dr. Edward Choi and Gerardo Flores on the Medical Records team to improve on outcome prediction for clinical time-series using unsupervised pretraining techniques.

### **Computational Linguistics Lab at Georgia Tech**

**August 2017 - present**

Projects include tying convex optimization to word embeddings to learn under hierarchical constraints as well as domain-knowledge grounding for end-to-end learning of effective representations of clinical text via deep learning.

### **Research and Development Intern, Sutter Heath**

**May 2018 - December 2018**

Worked with the Research, Development, and Dissemination group and Professor Jimeng Sun (Georgia Tech) to develop deep learning methodology for disease prediction from clinical text.

### **Anderson Lab at the College of Charleston**

**January 2016 - May 2017**

Researched extensions to Google's Word2Vec algorithm used to generate word embeddings for variable-length documents. Investigated performance of the algorithm when used directly as a classifier, and whether this technique, along with similarly created ensemble methods, could outperform benchmark preprocessing and machine learning pipelines on topic recognition tasks.

## **PROFICIENCIES**

---

<b>Daily Use</b>	<b>Languages:</b> Python ( <b>Pytorch/Tensorflow</b> /sklearn/pandas/nltk/numpy/multiproc), Bash. <b>Tools:</b> Git, TeX.
<b>Past Use</b>	<b>Languages:</b> R, Java, SQL, SAS, Octave. Fluency in French. <b>Tools:</b> Dynet, Oracle RDBMS, MongoDB, Tableau.

## **REVIEWING & SERVICE**

---

**EMNLP** 2019

**ACL** 2018 (subreviewer), 2019

**NeurIPS Machine Learning for Healthcare Workshop** 2017, 2018, 2019

**AMIA Informatics Summit** 2018, 2019

**Student Volunteer** ACM FAT\* 2019, NAACL 2018

## **AWARDS**

---

**School of Interactive Computing Travel Grant** (2019). Georgia Tech.

**Computing Research Association (CRA-W) Grad Cohort for Women Attendee** (2018, declined 2019).

**Phi Kappa Phi Graduate Fellowship Recipient** (2017). \$5,000.

**Data Science Major of the Year, Departmental Honors** (2017). College of Charleston.

**Grace Hopper Scholar**, the Anita Borg Institute (2016). Grace Hopper Celebration of Women in Computing Attendee (2015, 2016).

**Crosby Computer Science Award** (2014). Awarded by professorial nomination to the most promising student in an introductory computer science course at the College of Charleston.

**William Aiken Fellow** (2013-2017). A fellowship representing the top 1% of students at the College of Charleston. Recipient of associated scholarships totaling \$92,000.

## TEACHING AND SERVICE

---

### **Graduate Teaching Assistant, Georgia Institute of Technology**

Machine Learning (advanced undergraduate).

Deep Learning (graduate).

**Spring 2019**

**Fall 2019**

### **Graduate Research Assistant, Georgia Institute of Technology**

**Fall 2017 - Fall 2019**

### **Teaching Assistant, Charleston Digital Corridor**

MongoDB database course for adult professionals.

**Fall 2014**

### **Women in Computing Club at the College of Charleston**

**2014 - 2017**

President (2017), Vice-president (2016) and Treasurer (2015-16). Worked to promote diversity in computer science, host professional development workshops, and conduct community outreach.