

# XIAOCHUANG HAN (Han)

xiaochuang.han@gmail.com ◇ 404-988-6956 ◇ xhan77.github.io

## EDUCATION

---

**Carnegie Mellon University**  
Master of Language Technologies  
Advisor: Yulia Tsvetkov

August 2019 - Present

**Georgia Institute of Technology**  
B.S. in Computer Science, Minor in Mathematics  
Advisor: Jacob Eisenstein

August 2015 - May 2019

## PUBLICATIONS

---

[1] *Fortifying Toxic Speech Detectors Against Veiled Toxicity*  
**Xiaochuang Han** and Yulia Tsvetkov.  
*EMNLP 2020 (short)*

[2] *Explaining Black Box Predictions and Unveiling Data Artifacts through Influence Functions*  
**Xiaochuang Han**, Byron C. Wallace, and Yulia Tsvetkov.  
*ACL 2020 (long)*

[3] *Unsupervised Domain Adaptation of Contextualized Embeddings for Sequence Labeling*  
**Xiaochuang Han** and Jacob Eisenstein.  
*EMNLP 2019 (long)*

[4] *No Permanent Friends or Enemies: Tracking Dynamic Relationships between Nations from News*  
**Xiaochuang Han**, Eunsol Choi, and Chenhao Tan.  
*NAACL 2019 (long)*

[5] *Mind Your POV: Convergence of Articles and Editors Towards Wikipedia's Neutrality Norm*  
Umashanthi Pavalanathan, **Xiaochuang Han**, and Jacob Eisenstein.  
*CSCW 2018*

[6] *Interactional Stancetaking in Online Forums*  
Scott Kiesling, Umashanthi Pavalanathan, Jim Fitzpatrick, **Xiaochuang Han**, and Jacob Eisenstein.  
*Computational Linguistics, September 2018*

[7] *Predicting the Suitability of Service Animals Using Instrumented Dog Toys*  
Ceara Byrne, Jay Zuerndorfer, Larry Freil, **Xiaochuang Han**, Andrew Sirolly, Scott Cilliland, Thad Starner, and Melody Jackson.  
*IMWUT, December 2017*

## RESEARCH EXPERIENCE

---

**CMU LTI TsvetShop**  
Graduate Research Assistant, with Prof. Yulia Tsvetkov

August 2019 - Present

- Fortify toxic language classifiers against veiled toxicity using interpretable ML methods. [1]
- Explore the interpretability of NLP models through the lens of training examples. [2]

- Build data and models to forecast scientific mispractice in medical papers by their preregistrations.

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

August 2017 - May 2019

*Undergraduate Research Assistant*, with Prof. Jacob Eisenstein

- Improved unsupervised domain adaptation of contextualized embeddings for sequence labeling. [3]
- Explored variational methods for geo-entity resolution.
- Analyzed the effect of Wikipedia's neutrality norm. [5]
- Worked on stance classifiers in a quantitative model of stancetaking in online forums. [6]

### **University of Colorado Boulder NLP and CSS Lab**

May 2018 - August 2018

*Research Intern*, with Prof. Chenhao Tan

- Built an unsupervised model to explore entity-to-entity relations in world news. [4]

## **TEACHING ASSISTANTSHIPS**

---

### **CMU 11-711: Algorithms for NLP**

Fall 2020

*Graduate Teaching Assistant*, with Prof. Emma Strubell, Yulia Tsvetkov, and Robert Frederking

- Adapted and redesigned homework assignments, gave a lecture on natural language inference and interpretability in neural NLP, led recitations and hosted office hours.

## **ACADEMIC SERVICE**

---

### **Program Committee Member**

- NAACL 2021, EACL 2021, EMNLP 2020, W-NUT 2020

## **STUDENT ORGANIZATIONS**

---

### **Georgia Tech Big Data Club**

August 2015 - May 2019

*President and Lecturer*

- Organized weekly meetings and gave lectures on big data and machine learning tools and algorithms.