

Neel Guha

Email: nguha@stanford.edu

EDUCATION

- PhD**, Computer Science August 2021 –
Stanford University
- JD**, Stanford Law School August 2020 –
- MS**, Machine Learning August 2018 – December 2019
Carnegie Mellon University
- BS with Honors**, Computer Science September 2014 – June 2018
Stanford University
- Visiting Student**, Corpus Christi College January 2017 – March 2017
Oxford University

PROFESSIONAL EXPERIENCE

- Graduate Student Fellow**, RegLab, Stanford Law School June 2019 –
Advised by Professor Daniel Ho
- I develop computational language models for legal corpora and law related tasks. Legal corpora differ from traditional language corpora in important ways, requiring significant methodological adjustments to traditional approaches. I also researched bias in large mobility datasets used by local governments in COVID-19 response.
- Visiting Student**, HazyResearch, Stanford Computer Science May 2019 –
Advised by Professor Christopher Re
- My research focuses on developing methods and applications of machine learning to knowledge graphs and other structured information sources. This resulted in DICE, a benchmark for Data Integration, Cleaning, and Extraction task (<http://dicekr.io>). I also worked on Bootleg, a self-supervised system for named entity disambiguation over tail entities (<http://hazyresearch.stanford.edu/bootleg/>)
- Research Assistant**, Magic Lab, Stanford Mechanical Engineering June 2018 –
Advised by Professor Arun Majumdar
- I work on applications of machine learning to the challenge of power grid optimization - a problem known as Optimal Power Flow (OPF). Solving OPF would significantly improve our ability to leverage renewable energy sources.
- Research Assistant**, School of Computer Science, CMU June 2018 – December 2019
Advised by Professor Virginia Smith
- I researched methods for applying algorithmic fairness on large scale datasets. I developed methods for learning statistical models over massively distributed computational networks (federated learning). These methods are ideal for privacy sensitive contexts, i.e data belonging to different hospitals.
- Software Engineering Intern**, Laserlike June – September 2017
- I worked on user personalization and interest modeling at Laserlike, a small startup focused on determining and delivering tailored content to users. Laserlike was acquired by Apple in 2019.

Policy Intern, Federal Communications Commission

April – June 2017

- I worked with policy experts, engineers, and economists at the Office of Economics and Analytics (formerly known as the Office of Strategic Planning) to help draft policy recommendations on broadband expansion and collect broadband pricing data. I helped author a policy memo submitted to the Chairman’s office.

Software Engineering Intern, Google

June – September 2016

- I developed algorithms to support automatic watchtime bidding for YouTube Video Ads and ran live experiments evaluating different strategies.

Research Assistant, Institute for Quantitative Social Science, Harvard

June – September 2015

Advised by Professor Gary King and Professor Jennifer Pan

- I was part of a broader project to characterize propaganda spread by the Chinese Government (the ‘50cent party’) on Weibo (the Chinese equivalent of Twitter). I developed infrastructure to collect propaganda posts and detect fake accounts via machine learning. A full paper incorporating my work is accessible at <http://gking.harvard.edu/50c>

PUBLICATIONS

- Laurel Orr, Megan Leszczynski, Simran Arora, Sen Wu, **Neel Guha**, Xiao Ling, Chris Ré. Bootleg: Chasing the Tail with Self-Supervised Named Entity Disambiguation. To be presented at CIDR 2021.
- Amanda Coston, **Neel Guha**, Lisa Lu, Derek Ouyang, Alexandra Chouldechova, Daniel Ho. Leveraging Administrative Data for Bias Audits: Assessing Disparate Coverage with Mobility Data for COVID-19 Policy. *Under review*.
- **Neel Guha**, Zhecheng Wang, Matt Wytock, Arun Majumdar. (2019). Machine Learning for AC Optimal Power Flow. Oral Presentation at Climate Change Workshop at ICML 2019. **Honorable Mention, Best Paper**.
- **Neel Guha**, Ameet Talwalkar, Virginia Smith. (2018). One-Shot Federated Learning. Oral Presentation at Machine Learning on the Phone and other Consumer Devices Workshop (MLPCD 2) at NeurIPS 2018.
- **Neel Guha**. (2018). Knowledge Aggregation via Epsilon Model Spaces. Presented at the Common Model Infrastructure Workshop at 2018 ACM SIGKDD.
- **Neel Guha**. (2017). Behavioral Tracing of Twitter Accounts. Presented at the Privacy Online Workshop at the 2017 International Semantic Web Conference (ISWC).
- **Neel Guha**, Errol Ozdalga, and Matt Wytock. (2015). Medtree: A Search Engine for Medical Professionals. In International Symposium on Web Algorithms.
- **Neel Guha**. (2014). Spy Watch: A Tool for Transparency in Web Tracking. CEUR Workshop Proceedings. Presented at Privacy Online Workshop (Privon) at ISWC 2013.

A full list of papers can be found at <https://scholar.google.com/citations?hl=en&user=YI5N4HQAAAAJ>

TALKS

- Climate Change Workshop, ICLR 2020
- Climate Change Workshop, ICML 2019
- Workshop on Machine Learning on the Phone and other Consumer Devices, NeurIPS 2018
- Common Common Infrastructure Workshop, KDD 2018

COVERAGE

- *Smartphone Location Data Can Leave Out Those Most Hit by Covid-19*, The Wall Street Journal (April 5, 2021)

HONORS

- 2020 Francis Robbins Upton Fellowship in Engineering (declined)
- Finalist, 2020 Knight Hennessy Scholarship
- Best Paper Honorable Mention, Climate Change Workshop at ICML 2019

SERVICE

I served on the program committee at:

- Climate Change Workshop, NeurIPs 2020
- AI + Humanitarian Assistance and Disaster Response Workshop, NeurIPs 2020
- Climate Change Workshop, NeurIPs 2019
- Knowledge Representation & Reasoning Meets Machine Learning Workshop, NeurIPs 2019
- Adaptive and Multitask Learning Workshop, ICML 2019

I have served on the following committees:

- ResX, Provostial Task Force, Stanford University (May – December 2018)
- Committee on Residential Learning, Stanford University (September 2017 – June 2018)

I also volunteer with the Stanford Social Security Disability Project.