

Nelson F. Liu

EDUCATION **Stanford University**, Stanford, California **2019 - Present**
Ph.D. in Computer Science

University of Washington, Seattle, Washington **2015 - 2019**
B.S. with Distinction in Computer Science, B.A. in Linguistics
Advisor: Noah A. Smith
Thesis: *Implicit Linguistic Knowledge of Neural Natural Language Processing Models*

The Center for Brains, Minds and Machines, Cambridge, Massachusetts **August 2019**
Brains, Minds and Machines Summer Course

FELLOWSHIPS,
AWARDS &
HONORS NSF Graduate Research Fellowship (Natural Language Processing), 2019
Finalist, CRA Outstanding Undergraduate Researcher Award, 2019
Outstanding Undergraduate Honors Thesis Award, University of Washington CSE, 2019
Outstanding Computer Science Senior Award, University of Washington CSE, 2019
Best Paper Award, ACL 2018 Workshop on Representation Learning for NLP (RepL4NLP 2018) [5]
Barry M. Goldwater Scholarship, 2018
Washington Research Foundation Fellowship, 2018
Phi Beta Kappa, 2017
Washington Research Foundation Fellowship, 2017
UW HPCC Cloud Credit Grant Award, 2017
Mary Gates Research Scholarship, 2016
SciPy Scholarship, 2016

PUBLICATIONS [1] *Quoref: A Reading Comprehension Dataset with Questions Requiring Coreferential Reasoning*
Pradeep Dasigi, **Nelson F. Liu**, Ana Marasović, Noah A. Smith and Matt Gardner.
Conference on Empirical Methods in Natural Language Processing & International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), 2019.

[2] *Barack’s Wife Hillary: Using Knowledge Graphs for Fact-Aware Language Modeling*
Robert L. Logan IV, **Nelson F. Liu**, Matthew E. Peters, Matt Gardner and Sameer Singh.
Annual Meeting of the Association for Computational Linguistics (ACL), 2019.

[3] *Linguistic Knowledge and Transferability of Contextual Representations*
Nelson F. Liu, Matt Gardner, Yonatan Belinkov, Matthew E. Peters, and Noah A. Smith.
North American Chapter of the Association for Computational Linguistics (NAACL), 2019.

[4] *Inoculation by Fine-Tuning: A Method for Analyzing Challenge Datasets*
Nelson F. Liu, Roy Schwartz, and Noah A. Smith.
North American Chapter of the Association for Computational Linguistics (NAACL), 2019.

[5] *LSTMs Exploit Linguistic Attributes of Data*
Nelson F. Liu, Omer Levy, Roy Schwartz, Chenhao Tan and Noah A. Smith.
ACL Workshop on Representation Learning for NLP (RepL4NLP), 2018.
Best Paper Award.

[6] *AllenNLP: A Deep Semantic Natural Language Processing Platform*
Matt Gardner, Joel Grus, Mark Neumann, Oyvind Tafjord, Pradeep Dasigi, **Nelson F. Liu**,
Matthew Peters, Michael Schmitz, and Luke Zettlemoyer.
ACL Workshop for Natural Language Processing Open Source Software (NLP-OSS), 2018.

[7] *Discovering Phonesthemes with Sparse Regularization*
Nelson F. Liu, Gina-Anne Levow, and Noah A. Smith.
NAACL Workshop on Subword and Character Level Models in NLP (SCLeM), 2018.

[8] *Crowdsourcing Multiple Choice Science Questions*
Johannes Welbl, **Nelson F. Liu**, and Matt Gardner.
EMNLP Workshop on Noisy User-generated Text (W-NUT), 2017.

UNPUBLISHED
MANUSCRIPTS [9] *Augmenting Statistical Machine Translation with Subword Translation of Out-of-Vocabulary Words*
Nelson F. Liu, Jonathan May, Michael Pust and Kevin Knight. arXiv:1808.05700. August 2018.

RESEARCH
EXPERIENCE

University of Washington Computer Science and Engineering Seattle, Washington
Research Assistant **September 2015 - June 2019**

- Characterizing challenge datasets to discern which require models to truly generalize and which simply exploit blind spots a model's particular training dataset [4].
- Investigated how the properties of natural language data affect RNN generalization; RNNs trained on non-linguistic tasks learn to pick up on linguistic features, and that such linguistic features aid generalization [5].
- Studied phonesthemes, a noted exception to the arbitrariness of language, by building models to extract them from text [7].
- Built a model for predicting Supreme Court case outcomes from the text of amicus briefs and other court documents. Analyzed the resultant clusters of key phrases that were indicative of each justice's decisions.
- Advisor: Professor Noah A. Smith

Allen Institute for Artificial Intelligence (AI2) Seattle, Washington
Research Intern (Part-Time) **September 2018 - June 2019**
Research Intern **June - September 2018**

- Augmenting language models with explicit knowledge graph information to enable fact-aware text generation [2].
- Analyzing and interpreting deep contextualized word representations to discover what linguistic knowledge they capture and what aspects they miss [3].
- Advisors: Dr. Matt Gardner and Professor Noah A. Smith

USC Information Sciences Institute Marina del Rey, California
Natural Language Group Research Intern **June 2017 - September 2017**

Worked on language-independent translation of rare and out-of-vocabulary words, particularly with neural sequence to sequence models [9].

- Advisors: Professor Kevin Knight and Professor Jonathan May

Allen Institute for Artificial Intelligence (AI2) Seattle, Washington
Research Intern **January - March 2017**

Conducted transfer learning experiments to improve generalization across domains in multiple choice and span-predicting reading comprehension models [8].

- Also built significant portions of `deep_qa`, a framework for machine comprehension research on top of Keras. This formed the basis of the AllenNLP project [6].
- Advisor: Dr. Matt Gardner

TEACHING
EXPERIENCE

Natural Language Processing (CSE 447 / CSE M 547) **January - March 2019**
Teaching Assistant

Assisting with course planning and development, leading a weekly discussion section, and holding weekly office hours. Developing new teaching material based on AllenNLP.

- Instructor: Professor Noah A. Smith

University of Washington CSE Direct Admit Seminar **September - December 2018**
Teaching Assistant

Organizing and leading a seminar for first-year computer science students interested in getting involved in undergraduate research and pursuing research-based careers.

Natural Language Processing Capstone (CSE 481N) **March - June 2018**
Teaching Assistant

Advised teams of senior undergraduates and masters students on the design and implementation of original NLP projects.

- Instructor: Professor Yejin Choi

Natural Language Processing (CSE 447 / CSE M 547) **January - March 2018**
Teaching Assistant

Assisted with course planning and development, led a weekly discussion section, and held weekly office hours. Developed and administered a new PyTorch-based SQuAD reading comprehension project assignment.

- Instructor: Professor Yejin Choi

PROFESSIONAL EXPERIENCE	<p>scikit-learn <i>Google Summer of Code Developer</i> May - August 2016 Contributed to the popular Python machine learning library scikit-learn. Implementing various enhancements to decision tree module, most notably adding a adding new impurity splitting criterion and tree pre-pruning.</p>
LEADERSHIP	<p>Machines Who Learn <i>Organizer</i> 2016 - Present <i>Officer</i> 2015 - 2016</p> <ul style="list-style-type: none"> • Organizing and leading a machine learning and data science student organization at the University of Washington. • Organizing a reading group for students interested in discussing recent and seminal papers in AI (mostly NLP, vision, and robotics). <p>University of Washington Undergraduate Research Leader 2017 - Present</p> <ul style="list-style-type: none"> • Helping undergraduates get involved in undergraduate research, and encouraging them to do so. • Leading workshops, speaking to classes, participating in panels, and designing outreach events to help current and prospective students learn more about undergraduate research.
PROFESSIONAL SERVICE	<p>Program Committee Member:</p> <ul style="list-style-type: none"> • Annual Meeting of the Association for Computational Linguistics (ACL): 2019 • Conference on Empirical Methods in Natural Language Processing (EMNLP): 2019 • AAAI Conference on Artificial Intelligence (AAAI): 2020 • Conference on Computational Natural Language Learning (CoNLL): 2019 • Workshop on Analyzing and Interpreting Neural Networks for NLP (BlackboxNLP): 2018, 2019 • ACL Student Research Workshop: 2019 • Workshop on Methods for Optimizing and Evaluating Neural Language Generation: 2019
SELECTED PERSONAL / OPEN SOURCE PROJECTS	<p>AllenNLP <i>Contributor</i> A software library with abstractions for NLP research, written on top of PyTorch. Also see [6].</p> <p>scikit-learn <i>Contributor</i> Popular machine learning toolkit for Python. I contribute Python and Cython patches to the project, answer questions on the project mailing list and issue tracker, and review contributions from other developers.</p>