

# Swabha Swayamdipta

POSTDOCTORAL INVESTIGATOR, ALLEN INSTITUTE FOR AI

✉ swabhas@allenai.org | 🏠 swabhs.com | 🌐 swabhs | 🐦 @swabhaz

## Research Interests

- Inductive Biases for Structured Prediction and Unsupervised Learning
- Interpretability in Natural Language Processing
- Spurious and Social Biases in Datasets and Models

## Education

### Carnegie Mellon University

Pittsburgh, PA, USA

PHD IN LANGUAGE AND INFORMATION TECHNOLOGIES

2013 - 2019

- **Advisors:** Noah A. Smith, Chris Dyer
- **University of Washington**, Seattle, WA — *Visiting PhD student* Fall 2015 - Summer 2019
- **Thesis:** Syntactic Inductive Biases in NLP

### Columbia University

City of New York, NY, USA

MASTERS IN COMPUTER SCIENCE

2011 - 2012

- **Advisors:** Owen Rambow, Michael Collins
- **Specialization:** Natural Language Processing

### National Institute of Technology

Calicut, Kerala, India

BACHELORS OF TECHNOLOGY IN COMPUTER SCIENCE AND ENGINEERING

2006 - 2010

## Professional Experience

### Allen Institute for AI

Seattle, WA, USA

POSTDOCTORAL INVESTIGATOR

Aug 2019 - PRESENT

- **Advisor:** Yejin Choi
- **Project Team:** MOSAIC

### Allen Institute for AI

Seattle, WA, USA

RESEARCH INTERN, **Host:** MATTHEW E. PETERS

Summer-Fall 2018

- **Project:** Shallow Syntactic Priming of Large-scale Language Models

### Google Inc.

New York, NY, USA

RESEARCH INTERN, **Host:** ANKUR P. PARIKH

Summer 2017

- **Project:** Machine Comprehension for Google Neon

### Columbia University

New York, NY, USA

RESEARCH ASSISTANT, CENTER FOR COMPUTATIONAL LEARNING SYSTEMS

Jan 2013 - May 2013

### Google Inc.

Mountain View, CA, USA

SOFTWARE ENGINEER INTERN

Summer 2012

- Information retrieval evaluation for Google TV search

### Oracle Server Technologies

Bangalore, India

MEMBER TECHNICAL STAFF

2010 - 2011

- **Project:** Unified messaging for Fusion Middleware

### Microsoft Corporation

Hyderabad, India

SOFTWARE DEVELOPMENT INTERN

Summer 2009

- **Project:** Search functionality for customer hierarchies

## Teaching

### TUTORIALS

- |      |   |        |
|------|---|--------|
| 2019 | <b>Transfer Learning in Natural Language Processing</b> , Attendees: 350+                 | NAACL  |
| 2018 | <b>Frame Semantics across Languages: Towards a Multilingual FrameNet</b> , Attendees: 30+ | COLING |

## GUEST LECTURES

Winter 2019	<b>Phrase-Structure Parsing</b> , Host: Noah A. Smith	<i>UW CSE 447/547M: NLP</i>
Spring 2018	<b>Minimum Bayes Risk Decoding</b> , Host: Waleed Ammar	<i>UW CSE 599 D1: Advanced Topics in NLP</i>
Spring 2018	<b>Dependency Parsing</b> , Host: Noah A. Smith	<i>UW CSEP 517: NLP</i>
Fall 2017	<b>Unsupervised Learning</b> , Host: Noah A. Smith	<i>UW CSE 446: Machine Learning</i>
Fall 2014	<b>Dependency Parsing with Chu-Liu-Edmonds</b> , Hosts: Chris Dyer, Alon Lavie, Bob Frederking	<i>CMU 11-271: Algorithms for NLP</i>

## TEACHING ASSISTANTSHIP

Spring 2015	<b>CMU 11-411/611: Natural Language Processing</b> , Instructors: Chris Dyer, Alan W Black, Shomir Wilson	<i>Head TA</i>
-------------	---	----------------

## Mentorship

Fall 2020 - PRESENT	<b>Alon Jacovi</b> , PhD, Bar Ilan University	<i>AI2 Intern</i>
Summer 2020 - PRESENT	<b>Jenny Liang</b> , BS, UW CSE	<i>AI2 Intern</i>
Summer 2020 - PRESENT	<b>Jize Cao</b> , BS, UW CSE	<i>AI2 Intern</i>
Spring 2020 - PRESENT	<b>Xuhui Zhou</b> , CLMS, UW Linguistics	
Fall 2019 - Spring 2020	<b>Yiben Yang</b> , PhD, Northwestern University	
Fall 2019 - Spring 2020	<b>Chaitanya Malaviya</b> , PYI, Allen Institute for AI	
2018 - 2019	<b>Ron Fan</b> , MS, UW CSE	
2018 - 2019	<b>Karishma Mandyam</b> , BS, UW CSE	

## Invited Talks

Nov 02, 2020	<b>Responsible AI: Addressing Biases in Datasets and Models</b> , E+D Product Leads	<i>Microsoft</i>
Oct 30, 2019	<b>Biases and Learning Challenges in Natural Language Processing</b> , Rising Stars in EECS	<i>UIUC</i>
Oct 18, 2019	<b>Sprucing up a Dataset: Adversarially Filtering Dataset Artifacts</b> , Linguistics Colloquium	<i>UW Linguistics</i>
Apr 08, 2019	<b>Learning Challenges in Natural Language Processing</b> , AI2 Seminar	<i>Allen Institute for AI</i>
Apr 04, 2019	<b>Learning Challenges in Natural Language Processing</b> , NLP Seminar	<i>Stanford University</i>
Mar 13, 2019	<b>Learning Challenges in Natural Language Processing</b> , CLSP Seminar	<i>Johns Hopkins University</i>
Mar 11, 2019	<b>Learning Challenges in Natural Language Processing</b> , CILVR Seminar	<i>New York University</i>
Jan 22, 2019	<b>Learning Challenges in Natural Language Processing</b> , MSR Seminar	<i>Microsoft Research, New York</i>
Apr 27, 2018	<b>Syntactic Scaffolds for Semantic Structures</b> , NorthWest NLP	<i>Microsoft Research, Redmond</i>

## Awards and Honors

2020	<b>Best Paper: Honorable Mention</b> , Don't Stop Pretraining: Adapt Language Models to Domains and Tasks	<i>ACL</i>
2009	<b>Sun Campus Ambassador Scholarship</b> , Bangalore, India	<i>Sun Microsystems</i>
2003	<b>National Talent Search Scholarship</b> , Rourkela, India	<i>NCERT, India</i>

## Publications

### CONFERENCE PAPERS

Dataset Cartography: Mapping and Diagnosing Datasets with Training Dynamics	<i>EMNLP 2020</i>
• <b>S. Swayamdipta</b> , R. Schwartz, N. Lourie, Y. Wang, H. Hajishirzi, N. A. Smith, Y. Choi	
Generative Data Augmentation for Commonsense Reasoning	<i>Findings of EMNLP 2020</i>
• Y. Yang, C. Malaviya, J. Fernandez, <b>S. Swayamdipta</b> , R. LeBras, J. Wang, C. Bhagavatula, Y. Choi, and D. Downey	
Adversarial Filters of Dataset Biases	<i>ICML 2020</i>
• R. LeBras, <b>S. Swayamdipta</b> , C. Bhagavatula, R. Zellers, M. E. Peters, A. Sabharwal, and Y. Choi	
The Right Tool for the Job: Matching Model and Instance Complexities	<i>ACL 2020</i>
• R. Schwartz, G. Stanovsky, <b>S. Swayamdipta</b> , J. Dodge, and N. A. Smith	

Don't Stop Pretraining: Adapt Language Models to Domains and Tasks	<a href="#">ACL 2020</a>
<ul style="list-style-type: none"> <li>S. Gururangan, A. Marasović, <b>S. Swayamdipta</b>, K. Lo, I. Beltagy, D. Downey, and N. A. Smith</li> <li><b>Best Paper Honorable Mention</b></li> </ul>	
Syntactic Scaffolds for Semantic Structures	<a href="#">EMNLP 2018</a>
<ul style="list-style-type: none"> <li><b>S. Swayamdipta</b>, S. Thomson, K. Lee, L. Zettlemoyer, C. Dyer, N. A. Smith.</li> </ul>	
Learning Joint Semantic Parsers from Disjoint Data	<a href="#">NAACL 2018</a>
<ul style="list-style-type: none"> <li>H. Peng, S. Thomson, <b>S. Swayamdipta</b>, N. A. Smith</li> </ul>	
Annotation Artifacts in Natural Language Inference Data	<a href="#">NAACL 2018</a>
<ul style="list-style-type: none"> <li>S. Gururangan*, <b>S. Swayamdipta</b>*, O. Levy, R. Schwartz, S. Bowman, and N. A. Smith</li> <li>* equal contribution</li> </ul>	
Polyglot Semantic Role Labeling	<a href="#">ACL 2018</a>
<ul style="list-style-type: none"> <li>P. Mulcaire, <b>S. Swayamdipta</b>, and N. A. Smith</li> </ul>	
Multi-Mention Learning for Reading Comprehension with Neural Cascades	<a href="#">ICLR 2018</a>
<ul style="list-style-type: none"> <li><b>S. Swayamdipta</b>, A. Parikh, T. Kwiatkowski</li> </ul>	
Greedy, Joint Syntactic and Semantic Parsing with Stack LSTMs	<a href="#">CoNLL 2016</a>
<ul style="list-style-type: none"> <li><b>S. Swayamdipta</b>, M. Ballesteros, C. Dyer, N. A. Smith</li> </ul>	
A Dependency Parser for Tweets	<a href="#">EMNLP 2014</a>
<ul style="list-style-type: none"> <li>L. Kong, N. Schneider, <b>S. Swayamdipta</b>, A. Bhatia, C. Dyer, N. A. Smith</li> </ul>	
The Pursuit of Power and its Manifestation in Written Dialog	<a href="#">ICSC 2012</a>
<ul style="list-style-type: none"> <li><b>S. Swayamdipta</b>, O. Rambow</li> </ul>	

## WORKSHOP PAPERS

Multi-Task Learning for Incremental Parsing using Stack LSTMs	<a href="#">WiML @ NeurIPS 2016</a>
<ul style="list-style-type: none"> <li><b>S. Swayamdipta</b>, M. Ballesteros, C. Dyer, N. A. Smith</li> </ul>	
CMU: Arc-Factored, Discriminative Semantic Dependency Parsing	<a href="#">SemEval 2014</a>
<ul style="list-style-type: none"> <li>S. Thomson, D. Bamman, J. Dodge, <b>S. Swayamdipta</b>, N. Schneider, C. Dyer, N. A. Smith</li> </ul>	
The CMU Machine Translation Systems	<a href="#">WMT 2014</a>
<ul style="list-style-type: none"> <li>A. Matthews, C. Dyer, A. Lavie, G. Hanneman, W. Ammar, A. Bhatia, <b>S. Swayamdipta</b>, E. Schlinger, Y. Tsvetkov</li> </ul>	

## WORKING PAPERS

Shallow Syntax in Deep Water	<a href="#">arXiv:1908.11047</a>
<ul style="list-style-type: none"> <li><b>S. Swayamdipta</b>, M. Peters, B. Roof, C. Dyer and N. A. Smith</li> </ul>	
Frame-Semantic Parsing with Softmax-Margin Segmental RNNs and a Syntactic Scaffold	<a href="#">arXiv:1706.09528</a>
<ul style="list-style-type: none"> <li><b>S. Swayamdipta</b>, S. Thomson, C. Dyer and N. A. Smith</li> </ul>	
DyNet: The Dynamic Neural Network Toolkit	<a href="#">arXiv:1701.03980</a>
<ul style="list-style-type: none"> <li>G. Neubig, C. Dyer, Y. Goldberg, A. Matthews, W. Ammar, A. Anastasopoulos, M. Ballesteros, D. Chiang, D. Clothiaux, T. Cohn, K. Duh, M. Faruqi, C. Gan, D. Garrette, Y. Ji, L. Kong, A. Kuncoro, G. Kumar, C. Malaviya, P. Michel, Y. Oda, M. Richardson, N. Saphra, <b>S. Swayamdipta</b>, P. Yin</li> </ul>	

## Professional Service

---

### Area Chair

- NAACL 2021: Sentence-Level Semantics
- EACL 2021: Sentence-Level Semantics
- ACL 2020: Semantics (Long)

## Reviewer: Conferences

• ACL	2015-2020
• NAACL	2015-2020
• EMNLP	2015-2020
• NeurIPS	2018-2020
• ICML	2015,2019,2020
• EACL	2017
• AAAI	2017-2020
• CoNLL	2017-2018, 2020

## Reviewer: Workshops

• NAACL Student Research Workshop	2016
• Workshop for Women in Machine Learning (WiML) at NeurIPS	2016
• Machine Reading for Question Answering (MRQA) Workshop	2017

## Reviewer: Journals

• Computational Linguistics	2019 - PRESENT
• Transactions of ACL	2020 - PRESENT
• Journal of AI Research	2019

## Co-organizer

• EMNLP 2019: Workshop on Deep Learning in Low Resource NLP (DeepLo)	2019
• West Coast NLP Workshop	2018
• Women's Research Day	2017
• UW-NLP Retreat	2015-2016

## Mentorship at Conferences

- EMNLP 2020, ACL 2020, EMNLP 2018

## COMMUNITY SERVICE

### Allen Institute for AI

• Diversity, Equity and Inclusion Committee	2019-PRESENT
• AllenAI Outstanding Engineer Scholarship Committee	2019

### Paul G. Allen School of CSE, University of Washington

• Faculty Recruiting Liaison	2019
• Graduate Student Advisory Council	2018
• Outreach: Highline Public High School	2017
• CS: Gradswomen	2016-2019

## References

---

<b>Yejin Choi</b> , yejin@cs.washington.edu	<b>UW; AI2</b>
<b>Noah A. Smith</b> , noah@cs.washington.edu	<b>UW; AI2</b>
<b>Chris Dyer</b> , cdyer@google.com	<b>Google DeepMind</b>
<b>Luke Zettlemoyer</b> , lsz@cs.washington.edu	<b>UW; FAIR</b>