

Swabha Swayamdipta

POSTDOCTORAL INVESTIGATOR, ALLEN INSTITUTE FOR AI

✉ swabhas@allenai.org | 🏠 <https://swabhs.com> | 🌐 swabhs | 🐦 @swabhz

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

GUEST LECTURES

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

TEACHING ASSISTANTSHIP

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

Mentorship

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

Invited Talks

Nov 02, 2020	Responsible AI: Addressing Biases in Datasets and Models , E+D Product Leads	<i>Microsoft</i>
Oct 30, 2019	Biases and Learning Challenges in Natural Language Processing , Rising Stars in EECS	<i>UIUC</i>
Oct 18, 2019	Sprucing up a Dataset: Adversarially Filtering Dataset Artifacts , Linguistics Colloquium	<i>UW Linguistics</i>
Apr 08, 2019	Learning Challenges in Natural Language Processing , AI2 Seminar	<i>Allen Institute for AI</i>
Apr 04, 2019	Learning Challenges in Natural Language Processing , NLP Seminar	<i>Stanford University</i>
Mar 13, 2019	Learning Challenges in Natural Language Processing , CLSP Seminar	<i>Johns Hopkins University</i>
Mar 11, 2019	Learning Challenges in Natural Language Processing , CILVR Seminar	<i>New York University</i>
Jan 22, 2019	Learning Challenges in Natural Language Processing , MSR Seminar	<i>Microsoft Research, New York</i>
Sep 12, 2018	Representation Learning with Linguistic Structure , MSR AI Breakthroughs Workshop	<i>Microsoft Research, Redmond</i>
Apr 27, 2018	Syntactic Scaffolds for Semantic Structures , NorthWest NLP	<i>Microsoft Research, Redmond</i>

Awards and Honors

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

Publications

CONFERENCE PAPERS

Dataset Cartography: Mapping and Diagnosing Datasets with Training Dynamics	<i>EMNLP 2020</i>
• S. Swayamdipta , 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, S. Swayamdipta , R. LeBras, J. Wang, C. Bhagavatula, Y. Choi, and D. Downey	
Adversarial Filters of Dataset Biases	<i>ICML 2020</i>
• R. LeBras, S. Swayamdipta , C. Bhagavatula, R. Zellers, M. E. Peters, A. Sabharwal, and Y. Choi	

The Right Tool for the Job: Matching Model and Instance Complexities	ACL 2020
• R. Schwartz, G. Stanovsky, S. Swayamdipta , J. Dodge, and N. A. Smith	
Don't Stop Pretraining: Adapt Language Models to Domains and Tasks	ACL 2020
• S. Gururangan, A. Marasović, S. Swayamdipta , K. Lo, I. Beltagy, D. Downey, and N. A. Smith	
• Best Paper Honorable Mention	
Syntactic Scaffolds for Semantic Structures	EMNLP 2018
• S. Swayamdipta , S. Thomson, K. Lee, L. Zettlemoyer, C. Dyer, N. A. Smith.	
Learning Joint Semantic Parsers from Disjoint Data	NAACL 2018
• H. Peng, S. Thomson, S. Swayamdipta , N. A. Smith	
Annotation Artifacts in Natural Language Inference Data	NAACL 2018
• S. Gururangan*, S. Swayamdipta *, O. Levy, R. Schwartz, S. Bowman, and N. A. Smith	
• * equal contribution	
Polyglot Semantic Role Labeling	ACL 2018
• P. Mulcaire, S. Swayamdipta , and N. A. Smith	
Multi-Mention Learning for Reading Comprehension with Neural Cascades	ICLR 2018
• S. Swayamdipta , A. Parikh, T. Kwaikowski	
Greedy, Joint Syntactic and Semantic Parsing with Stack LSTMs	CoNLL 2016
• S. Swayamdipta , M. Ballesteros, C. Dyer, N. A. Smith	
A Dependency Parser for Tweets	EMNLP 2014
• L. Kong, N. Schneider, S. Swayamdipta , A. Bhatia, C. Dyer, N. A. Smith	
The Pursuit of Power and its Manifestation in Written Dialog	ICSC 2012
• S. Swayamdipta , O. Rambow	

WORKSHOP PAPERS

Multi-Task Learning for Incremental Parsing using Stack LSTMs	WiML @ NeurIPS 2016
• S. Swayamdipta , M. Ballesteros, C. Dyer, N. A. Smith	
CMU: Arc-Factored, Discriminative Semantic Dependency Parsing	SemEval 2014
• S. Thomson, D. Bamman, J. Dodge, S. Swayamdipta , N. Schneider, C. Dyer, N. A. Smith	
The CMU Machine Translation Systems	WMT 2014
• A. Matthews, C. Dyer, A. Lavie, G. Hanneman, W. Ammar, A. Bhatia, S. Swayamdipta , E. Schlinger, Y. Tsvetkov	

WORKING PAPERS

Shallow Syntax in Deep Water	arXiv:1908.11047
• S. Swayamdipta , M. Peters, B. Roof, C. Dyer and N. A. Smith	
Frame-Semantic Parsing with Softmax-Margin Segmental RNNs and a Syntactic Scaffold	arXiv:1706.09528
• S. Swayamdipta , S. Thomson, C. Dyer and N. A. Smith	
DyNet: The Dynamic Neural Network Toolkit	arXiv:1701.03980
• G. Neubig, C. Dyer, Y. Goldberg, A. Matthews, W. Ammar, A. Anastasopoulos, M. Ballesteros, D. Chiang, D. Clothiaux, T. Cohn, K. Duh, M. Faruqui, C. Gan, D. Garrette, Y. Ji, L. Kong, A. Kuncoro, G. Kumar, C. Malaviya, P. Michel, Y. Oda, M. Richardson, N. Saphra, S. Swayamdipta , P. Yin	

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

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