

# Katherine Atwell

*PhD Candidate, Northeastern University*

✉ atwell.ka@northeastern.edu ⬤ Boston, MA, USA ↲ <https://katherine-atwell.github.io/>

⌚ [github.com/katherine-atwell](https://github.com/katherine-atwell)>ID <https://orcid.org/0000-0002-5066-5604>

## EDUCATION

---

### Ph.D., Computer Science

*Northeastern University*

GPA: 4.0

08/2020 – 05/2026

Boston, United States

### B.S., Computer Science

*University of Maryland, Baltimore County*

08/2016 – 05/2020

Baltimore, United States

### B.S., Mathematics

*University of Maryland, Baltimore County*

08/2016 – 05/2020

Baltimore, United States

## PROFESSIONAL EXPERIENCE

---

### Team Member

*Alexa Prize TaskBot Challenge*

Deployed task-oriented dialogue system to real Alexa users; won third place

01/2023 – 09/2023

Pittsburgh, United States

### Software Engineering Intern

*Google, Inc.*

Developed a visualization tool for the Data Studio team

05/2019 – 08/2019

Mountain View, United States

### Engineering Practicum Intern

*Google, Inc.*

Built an internal dashboarding tool for YouTube

05/2018 – 08/2018

Mountain View, United States

### Research Intern (SPIRE-EIT REU)

*Iowa State University*

Designed the UX for new software for Traffic Incident Managers

06/2017 – 08/2017

Ames, United States

## PUBLICATIONS

---

### BASIL: Bayesian Assessment of Sycophancy in LLMs

2026

*ArXiv preprint*

**Katherine Atwell**, Pedram Heydari, Anthony Sicilia, Malihe Alikhani

### Perspectives of Global Deaf and Hard-of-hearing Community Members on Sign Language Technologies

2026

*ArXiv preprint*

**Katherine Atwell**, Saki Imai, Danielle Bragg, Malihe Alikhani

### MixDPO: Modeling Preference Strength for Pluralistic Alignment

2026

*ArXiv preprint*

Saki Imai, Pedram Heydari, Anthony Sicilia, Asteria Kaeberlein, **Katherine Atwell**, Malihe Alikhani

### Measuring Bias and Agreement in Large Language Model Presupposition Judgments

2025

*Findings of ACL 2025*

Katherine Atwell, Mandy Simons, Malihe Alikhani

**Contextual ASR Error Handling with LLMs Augmentation for Goal-Oriented Conversational AI**  
*COLING 2025 Industry Track*  
Yuya Asano, Sabit Hassan, Paras Sharma, Anthony Sicilia, **Katherine Atwell**, Diane Litman, Malihe Alikhani

2025

**Generating Signed Language Instructions in Large-Scale Dialogue Systems**  
*NAACL 2024 Industry Track*  
Mert Inan, **Katherine Atwell**, Anthony Sicilia, Lorna Quandt, Malihe Alikhani

2024

**Studying and Mitigating Biases in Sign Language Understanding Models**  
*EMNLP 2024*  
**Katherine Atwell**, Danielle Bragg, Malihe Alikhani

2024

**The Importance of Including Signed Languages in Natural Language Processing**  
*Sign Language Machine Translation, Springer Nature Switzerland*  
Kayo Yin, **Katherine Atwell**, Julie A Hochgesang, Malihe Alikhani

2024

**Combining Discourse Coherence with Large Language Models for More Inclusive, Equitable, and Robust Task-Oriented Dialogue**  
*COLING 2024*  
**Katherine Atwell**, Mert Inan, Anthony B Sicilia, Malihe Alikhani

2024

**Dialogue with Robots: Proposals for Broadening Participation and Research in the SLIVAR Community**  
*arXiv preprint arXiv:2404.01158*  
Casey Kennington, Malihe Alikhani, Heather Pon-Barry, **Katherine Atwell**, Yonatan Bisk, Daniel Fried, Felix Gervits, Zhao Han, Mert Inan, Michael Johnston, Raj Korpan, Diane Litman, Matthew Marge, Cynthia Matuszek, Ross Mead, Shiwali Mohan, Raymond Mooney, Natalie Parde, Jivko Sinapov, Angela Stewart, Matthew Stone, Stefanie Tellex, Tom Williams

2024

**ISABEL: An inclusive and collaborative task-oriented dialogue system**  
*Alexa Prize TaskBot Challenge 2 Proceedings*  
Anthony Sicilia, Yuya Asano, **Katherine Atwell**, Qi Cheng, Dipunj Gupta, Sabit Hassan, Mert Inan, Jennifer Nwogu, Paras Sharma, Malihe Alikhani.

2023

**Multilingual Content Moderation: A Case Study on Reddit.**  
*EACL 2023*  
Meng Ye, Karan Sikka, **Katherine Atwell**, Sabit Hassan, Ajay Divakaran, Malihe Alikhani

**How people talk about each other: Modeling Generalized Intergroup Bias and Emotion**  
*EACL 2023*  
Venkata Govindarajan, **Katherine Atwell**, Barea Sinno, Malihe Alikhani, David Beaver, Junyi Jessy Li.

2023

**APPDIA: A Discourse-aware Transformer-based Style Transfer Model for Offensive Social Media Conversations**  
*COLING 2022*  
**Katherine Atwell\***, Sabit Hassan\*, Malihe Alikhani

2022

**The Role of Context and Uncertainty in Shallow Discourse Parsing**  
*COLING 2022*  
**Katherine Atwell**, Remi Choi, Junyi Jessy Li, Malihe Alikhani

2022

**PAC-Bayesian Domain Adaptation Bounds for Multiclass Learners**  
*UAI 2022*  
Anthony Sicilia, **Katherine Atwell**, Malihe Alikhani, Seong Jae Hwang

**Political Ideology and Polarization: A Multi-dimensional Approach**  
*NAACL 2022*  
Barea Sinno\*, Bernardo Oviedo\*, **Katherine Atwell\***, Malihe Alikhani, Junyi Jessy Li

2022

**The Change that Matters in Discourse Parsing: Estimating the Impact of Domain Shift on Parser Error**

2022

*Findings of ACL 2022*

**Katherine Atwell\***, Anthony Sicilia\*, Seong Jae Hwang, Malihe Alikhani

**Where Are We in Discourse Relation Recognition?**

2021

*SIGDIAL 2021*

**Katherine Atwell**, Junyi Jessy Li, Malihe Alikhani

**Developing an Optimized UI for Traffic Incident Managers**

2018

*HFES 2018*

Andrina Helgerson, Jamiahus Walton, Celia Loya, Christopher Kawell, **Katherine Atwell**, Quinn Monaghan, Lakshay Ahuja, Hesham Hassan, Stephen Gilbert, & Anuj Sharma

## TEACHING EXPERIENCE

---

**Instructor of Record**

01/2026 – Present  
Boston, US

*Northeastern University*

Teaching CS4120, an introductory computer science course for undergraduate students at Northeastern

**Teaching Assistant**

09/2025 – Present  
Boston, United States

*Northeastern University*

Taught an NLP course and started an LLM track connecting principles learned in class to modern NLP systems

**Teaching Assistant**

2024

*Northeastern University*

Taught an unsupervised learning and data mining course

**Teaching Assistant**

2021

*University of Pittsburgh*

Taught an AI ethics course

**Teaching Assistant**

Pittsburgh, United States

*University of Maryland, Baltimore County*

Taught an object-oriented programming course

2019 – 2020

Baltimore, United States

**Mathematics Tutor**

2017 – 2019

*University of Maryland, Baltimore County*

Tutored for variety of math classes for the Learning Resource Center, including calculus, statistics, and real analysis

Baltimore, United States

**Teaching Assistant**

2018

*University of Maryland, Baltimore County*

Taught Calculus I

Baltimore, United States

## INVITED TALKS

---

**Prompt Design and Post-Training for LLMs**

2024

*Northeastern University*

Boston, United States

**Responsible Artificial Intelligence and Marginalized Communities**

2024

*Bilkent University*

Ankara, Türkiye

**Conversational AI Systems**

2024

*Northeastern University*

Boston, United States

**Coreference**

2023

*University of Pittsburgh*

Pittsburgh, United States

**APPDIA: A Discourse-aware Transformer-based Style Transfer Model for Offensive Social Media Conversations**

2022

*The University of Texas at Austin*

Austin, United States

## AWARDS

---

<b>Third Prize</b>	2023
<i>Alexa Prize TaskBot Challenge</i>	
<b>Best Paper</b>	2022
<i>Association for Uncertainty in Artificial Intelligence</i>	
<b>Travel Scholarship Recipient</b>	2022
<i>CRA-W Scholarship for Early Career Workshop and the Graduate Cohort</i>	
<b>Travel Scholarship Recipient</b>	2018
<i>Grace Hopper Celebration</i>	
<b>Meyerhoff Scholar</b>	2016
<i>University of Maryland, Baltimore County</i>	

## ACADEMIC SERVICE

---

<b>Organizer</b>	2025
<i>NLP for Positive Impact</i>	Vienna, Austria
Co-organized the fourth workshop for NLP for Positive Impact, which was co-located with ACL 2025	
<b>Organizer</b>	2023
<i>PittInclusion</i>	Pittsburgh, United States
Organized the PittInclusion program to provide underrepresented students with research experience with faculty mentorship	
<b>Area Chair</b>	2024 – 2025
<i>ACL Rolling Review, CoNLL</i>	
<b>Ethics Reviewer</b>	2024 – present
<i>ACL Rolling Review</i>	
<b>Reviewer</b>	
<i>ACL Rolling Review, EMNLP, ACL, NAACL, UAI, ACM FAccT, Journal of Educational Data Mining, INLG, AAAI</i>	
<b>Program Committee</b>	
<i>ACM FAccT, SIGDIAL, Workshop on Computational Approaches to Discourse</i>	
<b>Graduate Application Reviewer</b>	2024
<i>Northeastern University</i>	
<b>Hiring Interview Scribe</b>	2024
<i>Northeastern University</i>	