Katherine Atwell

✉ atwell.ka@northeastern.edu 𝕏 @katwellll
🌐 katherine-atwell.github.io 🦋 Katherine Atwell
in Katherine Atwell

Employment History

2026 – · · · ·	🔖	**Instructor of Record,** Northeastern University
2020 – 2026	🔖	**Research and Teaching Assistant,** Northeastern University
2023	🔖	**Student leader**, Alexa Prize TaskBot Challenge
2017 – 2020	🔖	**Teaching Assistant and Mathematics Tutor**, UMBC
2019	🔖	**Software Engineering Intern**, Google, Inc.
2018	🔖	**Engineering Practicum Intern**, Google, Inc.
2017	🔖	**Research Intern**, SPIRE-EIT REU, Iowa State University

Education

2020 – 2026	🔖	**Ph.D., Computer Science** Northeastern University
2016 – 2020	🔖	**B.S., Computer Science** University of Maryland, Baltimore County
2016-2020	🔖	**B.S., Mathematics** University of Maryland, Baltimore County

Research Publications

Preprints

1. S. Imai, P. Heydari, A. Sicilia, A. Kaeberlein, **K. Atwell**, and M. Alikhani, "Mixdpo: Modeling preference strength for pluralistic alignment," *arXiv preprint arXiv:2601.06180*, 2026.

2. **K. Atwell**, P. Heydari, A. Sicilia, and M. Alikhani, "Basil: Bayesian assessment of sycophancy in llms," *arXiv preprint arXiv:2508.16846*, 2025.

3. **K. Atwell**, S. Imai, D. Bragg, and M. Alikhani, "" nothing about us without us": Perspectives of global deaf and hard-of-hearing community members on sign language technologies," *arXiv preprint arXiv:2512.08839*, 2025.

4. C. Kennington et al., "Dialogue with robots: Proposals for broadening participation and research in the slivar community," *arXiv preprint arXiv:2404.01158*, 2024.

Conference Proceedings

1. Y. Asano et al., "Contextual asr error handling with llms augmentation for goal-oriented conversational ai," in *Proceedings of the 31st International Conference on Computational Linguistics: Industry Track*, 2025, pp. 374–386.

2. **K. Atwell**, M. Simons, and M. Alikhani, "Measuring bias and agreement in large language model presupposition judgments," in *Findings of the Association for Computational Linguistics: ACL 2025*, 2025, pp. 2096–2107.

3. **K. Atwell** et al., "Proceedings of the fourth workshop on nlp for positive impact (nlp4pi)," in *Proceedings of the Fourth Workshop on NLP for Positive Impact (NLP4PI)*, 2025.

4. **K. Atwell**, M. Inan, A. B. Sicilia, and M. Alikhani, "Combining discourse coherence with large language models for more inclusive, equitable, and robust task-oriented dialogue," in *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, 2024, pp. 3538–3552.

5. V. S. Govindarajan, **K. Atwell**, B. Sinno, M. Alikhani, D. I. Beaver, and J. J. Li, "How people talk about each other: Modeling generalized intergroup bias and emotion," in *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, 2023, pp. 2496–2506.

6. M. Ye, K. Sikka, **K. Atwell**, S. Hassan, A. Divakaran, and M. Alikhani, "Multilingual content moderation: A case study on reddit," in *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, 2023, pp. 3828–3844.

7. **K. Atwell**, R. Choi, J. J. Li, and M. Alikhani, "The role of context and uncertainty in shallow discourse parsing," in *Proceedings of the 29th International Conference on Computational Linguistics*, 2022, pp. 797–811.

8. **K. Atwell**, S. Hassan, and M. Alikhani, "Appdia: A discourse-aware transformer-based style transfer model for offensive social media conversations," in *Proceedings of the 29th International Conference on Computational Linguistics*, 2022.

9. **K. Atwell**, A. Sicilia, S. J. Hwang, and M. Alikhani, "The change that matters in discourse parsing: Estimating the impact of domain shift on parser error," in *Findings of the Association for Computational Linguistics: ACL 2022*, 2022, pp. 824–845.

10. S. Hassan, K. J. Atwell, and M. Alikhani, "Studying the effect of moderator biases on the diversity of online discussions: A computational cross-linguistic study," in *Proceedings of the Annual Meeting of the Cognitive Science Society*, vol. 44, 2022.

11. A. Sicilia, **K. Atwell**, M. Alikhani, and S. J. Hwang, "Pac-bayesian domain adaptation bounds for multiclass learners," in *Uncertainty in artificial intelligence*, PMLR, 2022, pp. 1824–1834.

12. B. Sinno, B. Oviedo, **K. Atwell**, M. Alikhani, and J. J. Li, "Political ideology and polarization: A multi-dimensional approach," in *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2022, pp. 231–243.

13. **K. Atwell**, J. J. Li, and M. Alikhani, "Where are we in discourse relation recognition?" In *Proceedings of the 22nd Annual Meeting of the Special Interest Group on Discourse and Dialogue*, 2021, pp. 314–325.

14. A. Helgerson et al., "Developing an optimized ui for traffic incident managers," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, SAGE Publications Sage CA: Los Angeles, CA, vol. 62, 2018, pp. 292–296.

Books and Chapters

1. K. Yin, **K. Atwell**, J. A. Hochgesang, and M. Alikhani, "The importance of including signed languages in natural language processing," in *Sign Language Machine Translation*, Springer Nature Switzerland Cham, 2024, pp. 73–87.

Awards and Achievements

2023	**Third place**, Alexa Prize TaskBot competition
2022	**Best Paper Award**, UAI 2022
	Travel Scholarship Recipient, CRA-W Scholarship for Early Career Workshop and the Graduate Cohort
2016 – 2020	**Meyerhoff Scholar**, University of Maryland, Baltimore County
2018	**Travel Scholarship Recipient**, Grace Hopper Celebration

Invited Talks

2026	**Are Models Telling us What we Want to Hear? Exploring sycophancy in LLMs**, Responsible AI Practice, Northeastern University
2025	**Conversational AI systems**, Northeastern University
2024	**Responsible Artificial Intelligence and Marginalized Communities**, Bilkent University
2022	**Coreference**, University of Pittsburgh
	APPDIA: A Discourse-aware Transformer-based Style Transfer Model for Offensive Social Media Conversations, The University of Texas at Austin

Skills

Languages	Native English speaker with beginner to intermediate Spanish proficiency
Coding	Python, C++, C, HTML, CSS, JavaScript, · · · ·
Libraries	HuggingFace, PyTorch, Pandas, scikit-Learn, · · · ·
Misc.	Mathematics, academic research, teaching, writing, data analysis

Teaching Experience

2026	**Instructor of Record: CS 4120, Natural Language Processing** *Northeastern University* Developed syllabus and course materials, delivered weekly lectures to 40+ students, facilitated semester-long research projects in NLP
2025	**Teaching Assistant: CS 4120, Natural Language Processing** *Northeastern University* Prepared course materials, graded homeworks, and gave a weekly series on modern applications of LLMs
2024	**Teaching Assistant: DS 5230, Unsupervised Learning and Data Mining** *Northeastern University* Gave in-class labs, graded assignments, assisted students with course projects
2021	**Teaching Assistant: CS 1684, Bias and Ethical Implications in Artificial Intelligence** *University of Pittsburgh* Held office hours, graded assignments, assisted students with course projects
2019-2020	**Teaching Assistant: CMSC 202, Introduction to Computer Science II** *UMBC* Assisted students with their assignments, held office hours, graded projects and exams
2018	**Teaching Assistant: Calculus I** *UMBC* Held discussion sections, proctored quizzes, graded quizzes and exams, held office hours
2017-2018	**Tutor: Math and Science Center** *UMBC* Tutored students for a variety of mathematics, statistics, and computer science courses at the university tutoring center

Academic Service

2025	**Organizer**, NLP for Positive Impact Workshop at ACL 2025
2024 – · · · ·	**Area Chair**, ACL Rolling Review, CoNLL

Academic Service (continued)

	Ethics Reviewer, ACL Rolling Review
2021 – · · · ·	**Reviewer**, ACL Rolling Review, EMNLP, ACL, NAACL, UAI, ACM FAccT, Journal of Educational Data Mining, INLG, AAAI
2022 – · · · ·	**Program Committee,** ACM FAccT, SIGDIAL, Workshop on Computational Approaches to Discourse
2021 – · · · ·	**Reviewer**, ACL Rolling Review, EMNLP, ACL, NAACL, UAI, ACM FAccT, Journal of Educational Data Mining, INLG, AAAI
2024	**Graduate Application Reviewer**, Northeastern University
	Hiring Interview Scribe, Northeastern University
2023	**Organizer**, PittInclusion