

Andrew Lee

Contact Information	E-mail: ajyl@umich.edu Website: https://ajyl.github.io	
Research Interests	Natural Language Processing: Dialogue, Computational Social Science	
Education	University of Michigan Ph.D. Candidate in Computer Science Advisor: Rada Mihalcea Ann Arbor, MI 2020 - Present	
	University of Michigan Master's in Computer Science GPA: 3.85/4.0 -- Summa Cum Laude Ann Arbor, MI 2015	
	Northwestern University Bachelor of Science in Computer Science GPA: 3.61/4.0 Evanston, IL 2013	
Publications	A. Lee , J. K. Kummerfeld, L. An, R. Mihalcea. 2021. Micromodels for Efficient, Explainable, and Reusable Systems: A Case Study on Mental Health. <i>Findings of Empirical Methods in Natural Language Processing (Findings of EMNLP)</i> .	
	S. Larson, A. Mahendran, J. J. Peper, C. Clarke, A. Lee , P. Hill, J. K. Kummerfeld, K. Leach, M. A. Laurenzano, L. Tang and J. Mars. 2019. An Evaluation for Intent Classification and Out-of-Scope Prediction. <i>Empirical Methods in Natural Language Processing (EMNLP)</i> .	
	S. Larson, A. Mahendran, A. Lee , J. K. Kummerfeld, P. Hill, M. Laurenzano, J. Hauswald, L. Tang, J. Mars. 2019. Outlier Detection for Improved Data Quality and Diversity in Dialog Systems. <i>North American Chapter of the Association for Computational Linguistics (NAACL)</i> .	
Experience	Microsoft Research Research Intern Advisor: Silviu-Petru Cucerzan Redmond, WA May 2021 - August 2021	
	Clinic, Inc. Core AI R&D - Senior Software Engineer, Team Lead Core AI R&D - Software Engineer Ann Arbor, MI June 2019 - August 2020 June 2017 - June 2019	
	Ford Motor Company Software Engineer Dearborn, MI March 2016 - June 2017	
Patents	Systems and methods for constructing an artificially diverse corpus of training data samples for training a contextually-biased model for a machine learning-based dialogue system. A. Lee , S. Larson, C. Clarke, K. Leach, J. Kummerfeld, P. Hill, J. Hauswald, M. Laurenzano, L. Tang, J. Mars. <i>US Patent 10,796,104. 2020.</i>	

Systems and methods for automatically congruing training data for training machine learning models of a machine learning-based dialogue system including seeding training samples or curating a corpus of training data based on instances of training data identified as anomalous. S. Larson, A. Mahendran, **A. Lee**, J. Kummerfeld, P. Hill, M. Laurenzano, J. Hauswald, L. Tang, J. Mars.
US Patent 10,679,150. 2020.

Professional Services

ACL - Association for Computational Linguistics
2021 - ACL Rolling Review
CLPsych - Workshop on Computational Linguistics and Clinical Psychology
2022 - Organizing Committee
2021 - Volunteer (@ NAACL)
CBO - International Symposium on Code Generation and Optimization
2019 - Program Committee

References

Available upon request.