

# Andrew Lee

**Contact**      **E-mail:** [ajyl@umich.edu](mailto:ajyl@umich.edu)  
**Information**    **Website:** <https://ajyl.github.io>  
                    **Google Scholar:** [Link](#)

**Research**      Interpretability, Language Models, Dialogue, Grounding, Natural Language Processing,  
**Interests**      Machine Learning

**Education**    **University of Michigan**      Ann Arbor, MI  
                    Ph.D. Candidate in Computer Science      2020 - Present  
                    Advisor: Rada Mihalcea  
                    **University of Michigan**      Ann Arbor, MI  
                    Master's in Computer Science      2015  
  
                    **Northwestern University**      Evanston, IL  
                    Bachelor of Science in Computer Science      2013

**Publications**    J. Xu, **A. Lee**, S. Sukhbaatar, J. Weston. 2023. Some things are more CRINGE than others: Preference Optimization with the Pairwise Cringe Loss. *Pre-print*

N. Nanda\*, **A. Lee**\*, M. Wattenberg. 2023. Emergent Linear Representations in World Models of Self-Supervised Sequence Models. In *BlackboxNLP (EMNLP Workshop)*.  
**Honorable Mention for Best-Paper.**

O. Ignat, Z. Jin, A. Abzaliev, L. Biester, S. Castro, N. Deng, X. Gao, A. Gunal, J. He, A. Kazemi, M. Khalifa, N. Koh, **A. Lee**, S. Liu, D.J. Min, S. Mori, J. Nwatu, V. Perez-Rosas, S. Shen, Z. Wang, W. Wu, R. Mihalcea. 2023. A PhD Student's Perspective on Research in NLP in the Era of Very Large Language Models. *Pre-print*.

**A. Lee**, J. K. Kummerfeld, L. An, R. Mihalcea. 2023. Empathy Identification Systems are not Accurately Accounting for Context. *European Chapter of the Association for Computational Linguistics (EACL)*.

**A. Lee**, D. Wu, E. Dinan, M. Lewis. 2022. Improving Chess Commentaries by Combining Language Models with Symbolic Reasoning Engines. *Pre-print*

**A. Lee**, Z. Chen, K. Leach, J. Kummerfeld. 2022. Augmenting Task-Oriented Dialogue Systems with Relation Extraction. In *Proceedings of the 10th Dialog System Technology Challenges (AAAI Workshop)*.

**A. Lee**, J. K. Kummerfeld, L. An, R. Mihalcea. 2021. Micromodels for Efficient, Explainable, and Reusable Systems: A Case Study on Mental Health. *Findings of Empirical Methods in Natural Language Processing (Findings of EMNLP)*.

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 Dataset for Intent Classification and Out-of-Scope Prediction. *Empirical Methods in Natural Language Processing (EMNLP)*.

\* Equal Contribution

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. *North American Chapter of the Association for Computational Linguistics (NAACL)*.

## Experience

### Meta AI Research

New York, NY

Research Intern - RAM (Reasoning, Attention, Memory) Team

May 2023 - August 2023

Advisors: Jing Xu, Jason Weston

### Meta AI Research

New York, NY

Research Intern - Diplomacy Team

May 2022 - August 2022

Advisors: Mike Lewis, Emily Dinan

### Microsoft Research

Redmond, WA

Research Intern - KTX Team

May 2021 - August 2021

Advisor: Silviu-Petru Cucerzan

### Clinic, Inc.

Ann Arbor, MI

Core AI R&D - Senior Software Engineer, Team Lead

June 2019 - August 2020

Core AI R&D - Software Engineer

June 2017 - June 2019

### Ford Motor Company

Dearborn, MI

Software Engineer

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.

*US Patent 10,796,104. 2020.*

Systems and methods for automatically configuring 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.*

## Awards

### Excellence in Climate, Diversity, Equity, and Inclusion

2023 - Award given to 5 students in the Computer Science and Engineering department for one's commitment and action in increasing diversity, equity, and inclusion in the division.

### University of Michigan CSE Research Honors AI Lab Nominee

## Professional Services

**ACL** - Association for Computational Linguistics

2021, 2022 - ACL Rolling Review

**CLPsych** - Workshop on Computational Linguistics and Clinical Psychology (@NAACL)

2022 - Organizing Committee

2021 - Volunteer

**CBO** - International Symposium on Code Generation and Optimization

2019 - Program Committee