

AKHILA YERUKOLA

PhD in Computer Science — Carnegie Mellon University (CMU)
ayerukol@andrew.cmu.edu ◇ <http://akhila-yerukola.github.io/>

RESEARCH INTERESTS

My research focuses on enhancing the reasoning and generation capabilities of language models across various contexts, with a particular emphasis on (a) **pragmatics and interactional contexts** to enable cooperative communication, and (b) the **cross-cultural safety and adaptation** of models.

EDUCATION

Carnegie Mellon University, School of Computer Science <i>Doctor of Philosophy in Language Technologies Institute</i> <i>Advisor:</i> Prof. Maarten Sap	2022 - Present
Stanford University <i>Master of Science in Computer Science</i> <i>Advisor:</i> Prof. Christopher Manning	2017 - 2019
National Institute of Technology Trichy (NITT), Tamil Nadu, India <i>Bachelor of Technology in Computer Science & Engineering</i>	2012 - 2016

RESEARCH EXPERIENCE

Allen Institute for Artificial Intelligence (Ai2) PhD Research Intern <i>Advisors:</i> Maarten Sap, Valentina Pyatkin, Jena D. Hwang <u>Project:</u> Improving Pragmatic Communication in LLMs to Better Satisfy Users' Intended Goals	May 2025 - Current
UCLA, Department of Computer Science Visiting Researcher <i>Advisors:</i> Prof. Nanyun (Violet) Peng, Saadia Gabriel <u>Project:</u> Enhancing Multimodal Model Sensitivity to Non-Verbal Communication Cues across Cultures (Gestures, Actions, etc)	June 2024 - Dec 2024
UW, Paul G. Allen School of Computer Science & Engineering Research Collaborator <i>Advisors:</i> Prof. Yejin Choi <u>Project:</u> Evaluating Commonsense Reasoning in Story Generation Models	March 2021 - March 2022
Stanford University, Department of Computer Science Graduate Research Assistant <i>Advisors:</i> Prof. Christopher Manning, Abigail See <u>Project:</u> Story Telling Capabilities of Large Pretrained Models	Sept 2018 - July 2019

SELECTED PUBLICATIONS

- Mind the Gesture: Evaluating AI Sensitivity to Culturally Offensive Non-Verbal Gestures**
Akhila Yerukola, Saadia Gabriel, Nanyun Peng, Maarten Sap
Association for Computational Linguistics (ACL), 2025.
- Words Like Knives: Backstory-Personalized Modeling and Detection of Violent Communication**
Jocelyn Shen, Akhila Yerukola, Xuhui Zhou, Cynthia Breazeal, Maarten Sap, Hae Won Park
Empirical Methods in Natural Language Processing (EMNLP), 2025.

3. **PolyGuard: A Multilingual Safety Moderation Tool for 17 Languages**
 Priyanshu Kumar, Devansh Jain, Akhila Yerukola, Liwei Jiang, Himanshu Beniwal, Thomas Hartvigsen, Maarten Sap
Conference on Language Modeling (COLM), 2025.
4. **NormAd: A Framework for Measuring the Cultural Adaptability of Large Language Models**
Akhila Yerukola*, Abhinav Rao*, Vishwa Shah, Katharina Reinecke, Maarten Sap
Nations of the Americas Chapter of the Association for Computational Linguistics (NAACL), 2025.
5. **Is the Pope Catholic? Yes, the Pope is Catholic. Generative Evaluation of Intent Resolution in LLMs**
Akhila Yerukola, Saujas Vaduguru, Daniel Fried, Maarten Sap
Association for Computational Linguistics (ACL), 2024.
6. **Don't Take This Out of Context! On the Need for Contextual Models and Evaluations for Stylistic Rewriting**
Akhila Yerukola, Xuhui Zhou, Elizabeth Clark, Maarten Sap
Empirical Methods in Natural Language Processing (EMNLP), 2023.
7. **Beyond Denouncing Hate: Strategies for Countering Implied Biases and Stereotypes in Language**
 Jimin Mun, Emily Allaway, Akhila Yerukola, Laura Vianna, Sarah-Jane Leslie, Maarten Sap
Findings of Empirical Methods in Natural Language Processing (EMNLP), 2023.
8. **COBRA Frames: Contextual Reasoning about Effects and Harms of Offensive Statements**
 Xuhui Zhou, Hao Zhu, Akhila Yerukola, Thomas Davidson, Jena D. Hwang, Swabha Swayamdipta, Maarten Sap
Findings of Association for Computational Linguistics (ACL), 2022.
9. **Explainable Slot Type Attentions to Improve Joint Intent Detection and Slot Filling**
 Kalpa Gunaratna, Vijay Srinivasan, Akhila Yerukola, Hongxia Jin
Findings of Empirical Methods in Natural Language Processing (EMNLP), 2022.
10. **Data Augmentation for Voice-Assistant NLU using BERT-based Interchangeable Rephrase**
Akhila Yerukola*, Mason Bretan*, Hongxia Jin
European Chapter of the Association for Computational Linguistics (EACL), 2021.
11. **Do Massively Pretrained Language Models Make Better Storytellers?**
 Abigail See, Akhila Yerukola*, Aneesh Pappu*, Rohun Saxena*, Christopher D. Manning
Computational Natural Language Learning (CoNLL), 2019.

(* equal contribution)

INDUSTRY EXPERIENCE

Samsung Research America (SRA) , Mountain View, USA	Aug 2019 - June 2022
Senior AI Research Engineer Advisors: <i>Hongxia Jin, Vijay Srinivasan</i>	
— Worked on fine-grained natural language understanding for Bixby [EACL 2021]	
— Built an internal AI Chatbot to handle HR, IT & Legal user queries [SRA Q3 2021 President's Award]	
— Delivered a low-resource domain classification system for Bixby with a 50% error rate reduction, across Samsung's mobile phones and smart home products [SRA Q4 2020 President's Award]	

Microsoft R&D, Hyderabad, India

June 2016 - Aug 2017

Software Engineer | *Advisors: Sudhakar Sadashivuni*

— Built Microsoft Social Engagement's sentiment analysis model to improve sales through social insights

— Developed an orchestration engine library for dense Azure cloud deployments

TEACHING EXPERIENCE

Ethics, Safety, and Social Impact in NLP and LLMs 11-430/830

Spring 2025

Teaching Assistant

Instructor: Prof. Maarten Sap

— Worked with a team of 3 TAs for 25 students to refine and grade course assignments, and mentored 2 course project student teams

Advanced NLP 11-711

Spring 2024

Teaching Assistant

Instructor: Prof. Graham Neubig

— Worked with a team of 7 TAs for 100 students to refine and grade course assignments, and mentored 5 course project student teams

Natural Language Understanding CS224U

Spring 2019

Graduate Teaching Assistant

Instructors: Prof. Christopher Potts, Prof. Bill MacCartney

— Worked with a team of 10 TAs for 250+ students to refine and grade course assignments. Mentored 11 course project student teams

— Taught a lecture on “Probing black box models”

Machine Learning CS229

Spring 2019

Graduate Teaching Assistant

Instructor: Adjunct Prof. Andrew Ng

— Worked with a team of 30+ TAs for 850+ students to develop new assignments, refine and grade course assignments. Mentored 30 course project student teams

Delta, a web development club of NITT

June 2013 - May 2014

Undergraduate Volunteer Instructor

— Taught basics of C++, Java and Android application development

AWARDS

- Recipient of K&L Gates Presidential Fellowship 2025-2026
- Q3 2021 SRA President's Award for excellence in research and innovation
- Q4 2020 SRA President's Award for excellence in research and innovation
- 2nd Best Project Award, CS231n (Convolutional Neural Networks for Visual Recognition) Stanford 2018. Picked out of 200+ projects.
- 2nd Best Project Award, CS224n (NLP with Deep Learning) Stanford 2018. Picked out of 145+ projects.
- Full-ride undergraduate scholarship by Ministry of Human Resource Development (MHRD), India 2012-2016. *The scholarship is awarded to the top 0.01 % of students in the All India Engineering Exam, 2012.*
- Third Highest GPA Award, 2015. Awarded by the Department of Computer Science, NITT.

PROFESSIONAL EXPERIENCE AND VOLUNTEERING

Organizer: Socially Responsible Language Modelling Research (SoLaR) Workshop at COLM 2025, Generation, Evaluation, Metrics (GEM) Benchmark and Workshop at ACL 2021

Reviewer: ARR July 2025, ARR May 2025, NLP4PosImpact 2025, ARR Feb 2025, C3NLP 2025, NAACL 2025, ACL 2024, EMNLP 2024 (**Outstanding Reviewer**), NAACL 2024, EMNLP 2023, EMNLP 2020

Secondary Reviewer: NAACL-HLT 2021, COLING 2020, CIKM 2020

Volunteer: EMNLP 2021, EMNLP 2020

STUDENT MENTORSHIP

- Jenna Godsey (Undergraduate at CMU)
- Tianyu Cao (MIIS at CMU)
- Neel Bhandari (MIIS at CMU)
- Sanidhya Vijayvargiya (MSCS at CMU)
- Vishwa Shah (MIIS at CMU)
- Abhinav Rao (MIIS at CMU)
- Hyuk Joon Kwon (Intern at Samsung Research America)

DEPARTMENTAL SERVICE

- Member of Belonging and Engagement at LTI Committee, 2025
- LTI Student Mentoring Program, 2024
- LTI Student Mentoring Program, 2023
- Member of the Stanford MSCS Admissions Committee, 2019. Reviewed student applications with faculty to select the incoming MS students for 2020.
- Editor of Bits and Bytes, the official newsletter of the Department of Computer Science, NITT, 2016. Managed a team of 10 people to release a newsletter twice a month.
- Head of Public Relations at Vortex, Annual Symposium of Department of Computer Science, NITT, 2016.