

# Atharva Kulkarni

Ph.D. in Computer Science | University of Southern California

⌚ athrvkk.github.io ⌚ atharva.kulkarni@usc.edu ⌚ github.com/athrvkk ⌚ Google Scholar ⌚ linkedin.com/in/athrvkk  
🐦 x.com/athrvkk ⌚ athrvkk.bsky.social

Research Interests: Theoretical & Empirical Understanding of LLMs | Learning Dynamics | Trustworthy NLP

## Education

May 2029	University of Southern California   Thomas Lord Dept. of Computer Science	Los Angeles, CA
Aug 2024	Doctor of Philosophy (Ph.D.) in Computer Science > Advisor: <a href="#">Swabha Swayamdipta</a> > Honors: Viterbi School of Engineering Fellowship (AY 2024-25)	
May 2025	University of California, Berkeley   Simons Institute for the Theory of Computing	Berkeley, CA
Jan 2025	Visiting Ph.D. Student   Program: Special Year on LLMs and Transformers (Part 2)	
May 2024	Carnegie Mellon University   Language Technologies Institute	Pittsburgh, PA
Aug 2022	Master of Science (M.S.) in Language Technologies > Advisors: <a href="#">Graham Neubig</a> & <a href="#">Barnabás Póczos</a> > Collaborators: <a href="#">Aditi Raghunathan</a> , <a href="#">Ameet Talwalkar</a> , <a href="#">David Mortensen</a> > Honors: CMU LTI Research Fellowship (AY 2023-24)	
May 2020	Savitribai Phule Pune University   PVG's College of Engineering and Technology	Pune, India
Aug 2016	Bachelor of Engineering (B.E.) in Computer Science	

## Experience

Aug 2024	Apple Inc.   Siri & Information Intelligence (SII)	Seattle, WA
May 2024	Machine Learning Research Intern   Host: <a href="#">Yuan Zhang</a> & <a href="#">Hong Yu</a> Collaborators: <a href="#">Joel Moniz</a> , <a href="#">Xiou Ge</a> , <a href="#">Bo-Hsiang Tseng</a> , & <a href="#">Dhivya Piraviperumal</a> > Meta evaluation of hallucination detection metrics for LLM-generated texts ( <a href="#">Findings of EMNLP'25</a> ).	
Aug 2023	Apple Inc.   Siri & Information Intelligence (SII)	Seattle, WA
May 2023	Machine Learning Research Intern   Host: <a href="#">Shruti Bhargava</a> & <a href="#">Hong Yu</a> Collaborators: <a href="#">Bo-Hsiang Tseng</a> , <a href="#">Joel Moniz</a> , & <a href="#">Dhivya Piraviperumal</a> > Synthetic data generation using LLMs for dialog state tracking. ( <a href="#">EACL'24</a> )	
July 2022	IIIT Delhi   Laboratory for Computational Social Systems (LCS2) [⌚]	Delhi, India
March 2021	Predoctoral Researcher / Advisor: <a href="#">Tanmoy Chakraborty</a> > Hate speech detection under subpopulation and covariate shifts. ( <a href="#">KDD'23</a> ) > Semantic role labeling of multimodal memes using optimal transport. ( <a href="#">EACL'23</a> ) > Claim span identification using structural constraints and domain-specific rationales. ( <a href="#">EMNLP'22</a> ) > Sarcasm explanation generation in multi-modal multi-party dialogues. ( <a href="#">ACL'22</a> )	

## Selected Publications (\* denotes equal contribution)

- [12] **Economies of Open Intelligence: Tracing Power & Participation in the Model Ecosystem** [[PDF](#)]  
Shayne Longpre, Christopher Akiki, Campbell Lund, [Atharva Kulkarni](#),..., Lucie-Aimée Kaffee  
*Under Review*
- [11] **Evaluating Evaluation Metrics – The Mirage of Hallucination Detection** [[PDF](#)]  
[Atharva Kulkarni](#), Yuan Zhang, Joel Moniz, Xiou Ge, ..., Swabha Swayamdipta, Hong Yu  
*Findings of the Association for Computational Linguistics* [[Findings of EMNLP'25](#)]
- [10] **Still Not Quite There! Evaluating Large Language Models for Comorbid Mental Health Diagnosis** [[PDF](#) | [code](#)]  
Amey Hengle\*, [Atharva Kulkarni](#)\*, Shantanu Patankar, ..., Rashmi Gupta  
*Conference on Empirical Methods in Natural Language Processing* [[EMNLP'24](#)]
- [9] **Multitask Learning Can Improve Worst-Group Outcomes** [[PDF](#) | [code](#)]  
[Atharva Kulkarni](#)\*, Lucio Dery\*, Amritth Setlur, Aditi Raghunathan, Ameet Talwalkar, Graham Neubig  
*Transactions of Machine Learning Research* [[TMLR'24](#)]
- [8] **SynthDST: Synthetic Data is All You Need for Few-Shot Dialog State Tracking** [[PDF](#) | [code](#)] [[Oral Presentation](#)]  
[Atharva Kulkarni](#), Bo-Hsiang Tseng, Joel Moniz, Dhivya Piraviperumal, Hong Yu, Shruti Bhargava  
*Conference of the European Chapter of the Association for Computational Linguistics* [[EACL'24](#)]

- [7] **The Student Becomes the Master: Matching GPT3 on Scientific Factual Error Correction** [PDF | code]  
Dhananjay Ashok, Atharva Kulkarni, Hai Pham, Barnabás Póczos  
*Findings of the Association for Computational Linguistics* [Findings of EMNLP'23]
- [6] **Counting the Bugs in ChatGPT's Wugs: A Multilingual Investigation into the Morphological Capabilities of a Large Language Model** [PDF | code]  
Leonie Weisseler, ..., Atharva Kulkarni, ..., Hinrich Schuetze, Kemal Oflazer, David Mortensen  
*Conference on Empirical Methods in Natural Language Processing* [EMNLP'23]
- [5] **Learning & Reasoning Multifaceted Longitudinal Data for Poverty Estimates and Livelihood Capabilities of Lagged Regions in Rural India** [PDF] [AI for Social Good Award]  
Atharva Kulkarni, Raya Das, Ravi Srivastava, Tanmoy Chakraborty  
*International Joint Conference on Artificial Intelligence* [IJCAI'23]
- [4] **Revisiting Hate Speech Benchmarks: From Data Curation to System Deployment** [PDF | code]  
Atharva Kulkarni\*, Sarah Masud\*, Vikram Goyal, Tanmoy Chakraborty  
*ACM SIGKDD Conference on Knowledge Discovery and Data Mining* [KDD'23]
- [3] **Characterizing Entities in Harmful Memes: Who is Hero, Villain, Victim?** [PDF | code]  
Shivam Sharma\*, Atharva Kulkarni\*, ..., Preslav Nakov, Md. Shad Akhtar, Tanmoy Chakraborty  
*Conference of the European Chapter of the Association for Computational Linguistics* [EACL'23]
- [2] **Empowering the Fact-checkers! Automatic Identification of Claim Spans on Twitter** [PDF | code]  
Megha Sundriyal\*, Atharva Kulkarni\*, Vaibhav Pulastya, Md. Shad Akhtar, Tanmoy Chakraborty  
*Conference on Empirical Methods in Natural Language Processing* [EMNLP'22]
- [1] **When did you become so smart, oh wise one?! Sarcasm Explanation in Multi-modal Multi-party Dialogues**  
[PDF | code] [Oral Presentation]  
Shivani Kumar\*, Atharva Kulkarni\*, Md. Shad Akhtar, Tanmoy Chakraborty  
*Annual Meeting of the Association for Computational Linguistics* [ACL'22]

## Academic Service

---

**Conference Reviewer** COLM'25, ICLR'24, ACL ('23-'25), EMNLP ('22-'25)

**Workshop Reviewer** WiNLP 2024, NeurIPS SSLTheoryPractice 2023, NeurIPS RO-FoMo 2023

**Conference Volunteer** EACL 2024, EACL 2023, EACL 2021

## Teaching & Mentorship Roles

---

### Mentor – CS Undergraduate Mentorship Program | USC

Fall'24

› Mentoring USC undergraduate students for starting out in research (Mentee – Tshephang Makubate & Grace Li).

### Teaching Assistant – Advanced NLP | CMU

Fall'23

› TA for 11-711 Advanced NLP taught by Prof. Daniel Fried. Designed assignments and guided teams for course projects.

### Teaching Assistant – Introduction to NLP | CMU

Spring'23

› TA for 11-611 Intro to NLP taught by Prof. David Mortensen. Designed assignments and guided teams for course projects.

### Mentor – AI Undergraduate Mentorship Program | CMU

Fall'22

› Mentored CMU undergraduate students for starting out in research (Mentee – Yi (Nicole) Xiang).

## Talks

---

### SynthDST: Synthetic Data is All You Need for Few-Shot Dialog State Tracking

› NiklasOPF AI podcast [video link]

Nov 2024 – YouTube

› EACL'24 Oral Presentation

March 2024 – St. Julian's, Malta

› Siri & Information Intelligence (SII), Apple. Inc

Aug 2024 – Seattle, WA

### When did you become so smart, oh wise one?! Sarcasm Explanation in Multi-modal Multi-party Dialogues

› Student Research Symposium, CMU LTI

Aug 2022 – Pittsburgh, PA

## Honours & Awards

---

1. **USC Viterbi Fellowship (2024-25)** – Awarded to selected incoming PhD student at USC, covering full tuition & stipend.
2. **CMU LTI Research Fellowship (2023-24)** – Awarded to selected MLT students at CMU, covering full tuition & stipend.
3. **Best Paper – AI for Social Good | IJCAI'23 [🔗]** – Poverty estimation using neural networks for rural Indian setting.
4. **Shared Task Winner | WASSA @ EACL'21 [🔗]** – Empathy & distress detection in reaction to news stories.
5. **Shared Task Winner | TechDoFication @ ICON'20 [🔗]** – Coarse-grained technical domain identification of documents for indigenous Indian languages.
6. **Bachelor's Thesis Award | ZS ASPIRE 2020 [🔗]** – 2<sup>nd</sup> place amongst 33,000 teams & cash prize of 2,00,000 INR for Bachelor's Thesis project.

## References

---

<b>Prof. Swabha Swayamdipta</b>	<i>Assistant Professor, University of Southern California, USA</i> [🔗, 📧]
<b>Prof. Barnabás Póczos</b>	<i>Associate Professor, Carnegie Mellon University, USA</i> [🔗, 📧]
<b>Prof. Graham Neubig</b>	<i>Associate Professor, Carnegie Mellon University, USA</i> [🔗, 📧]