



# Shafiuddin Rehan Ahmed


shah7567 [at] colorado [dot] edu

 ahmeshaf.github.io

 ahmeshaf

 ahmeshaf

 t3IALyoAAAAJ

 ahmeshaf

## SUMMARY

---

My research interests lie in the understanding of events in natural language, detecting important textual cues, and finding common threads in linguistic variations. My dissertation: ***Efficient Cross-document Event Coreference Resolution*** broadly covers but is not limited to the following areas: *Efficient NLP* [8, 7, 5], *Neural and Symbolic Representation Learning* [8, 5, 3], *Semantics* [6, 5, 3], *Structured Prediction and Synthetic Data Generation with LLMs* [5, 3, 1].

My most recent project deals with synergetic learning of related NLP tasks.

## EDUCATION

---

*MS and Ph.D.*, Computer Science and Engineering *Aug 2017 - May 2024*  
**University of Colorado**, Boulder, USA

*Bachelor of Technology*, Computer Science and Engineering *Aug 2008 - May 2012*  
**Indian Institute of Technology**, Hyderabad, India

## RESEARCH EXPERIENCE (5+ YEARS)

---

**Graduate Research Assistant** (Python) *Jan 2024 - May 2024*  
University of Colorado, Boulder, USA

- Our paper on *Synthetic Data Generation with LLMs* got accepted at ACL 2024.
- Papers in-review about Synergetic task training and Text2text Semantic Role Labeling.

**Machine Learning Engineer Intern - Remote** (Python) *May 2023 - Aug 2023*  
ExplosionAI GmbH (makers of spaCy), Berlin, Germany

- Our papers got accepted at EACL-System Demo 2024 & LREC-COLING 2024b

**Machine Learning Engineer Intern - Remote** (Python) *May 2022 - Aug 2022*  
ExplosionAI GmbH (makers of spaCy), Berlin, Germany

- Our paper on model-in-the-loop annotations got accepted at LAW 2023

**Graduate Research Assistant** (Java, Python) *Jan 2018 - May 2022*  
University of Colorado, Boulder, USA

- Technical lead of our team for the cross-university project on *Knowledge Graph Construction - AIDA*. Papers accepted at TAC 2018, TAC 2019, ACL 2023 (Findings), EMNLP 2023 (System Demo).

**Software Engineering Intern II** (Python) *June 2018 - Aug 2018*  
Sopris Health, USA

- Irrelevant utterance detection in clinical conversations recorded by Sopris Health app.

**Senior R&D Engineer** (Python) *Aug 2016 - Aug 2017*  
HP Inc. R&D, Bangalore, India

- NLP solutions for printer customer care service. Filed patent: US20210049489A1

**Software Developer** (C#) *Aug 2012 - Aug 2016*  
HP Inc. R&D, Bangalore, India

- Front-end development for printer testing tools.

## PATENTS

---

1. Shameed Sait M A, Shafiuddin Rehan Ahmed, Niranjana Damera Venkata. *Providing Solutions Using Stochastic Modelling*. en. 2018. URL: <https://patents.google.com/patent/US20210049489A1/>

## AWARDS

---

1. Best Frame Recall for Cross Document Event Coreference Resolution in Text Analysis Conference, 2019
2. 3rd Place Outstanding Poster - In-Progress Research, Graduate Students' Research expo., 2018-2019
3. Merit-cum-Means Scholarship for Undergraduate Studies, 2009-2012

## PAPERS

---

1. **Shafiuddin Rehan Ahmed**, Zhiyong Wang, George Baker, Kevin Stowe, and James H. Martin. *Generating Harder Cross-document Event Coreference Resolution Datasets using Metaphoric Paraphrasing*. ACL 2024
2. Micah Zhang, **Shafiuddin Rehan Ahmed**, and James H. Martin. *FtG-CoT at SemEval-2024 Task 9: Solving Sentence Puzzles Using Fine-Tuned Language Models and Zero-Shot CoT Prompting*
3. **Shafiuddin Rehan Ahmed**, George Arthur Baker, Evi Judge, Michael Reagan, Kristin Wright-Bettner, Martha Palmer, and James H. Martin. *Linear Cross-document Event Coreference Resolution with X-AMR*. LREC-COLING 2024
4. Abhijnan Nath, Huma Jamil, **Shafiuddin Rehan Ahmed**, George Baker, Rahul Ghosh, James H. Martin, Nathaniel Blanchard, and Nikhil Krishnaswamy. *Multimodal Cross-Document Event Coreference Resolution Using Linear Semantic Transfer and Mixed-Modality Ensembles*. LREC-COLING 2024
5. **Shafiuddin Rehan Ahmed**, Jon Cai, Martha Palmer, and James H. Martin. *X-AMR Annotation Tool*. EACL-System Demo 2024
6. Jon Cai, **Shafiuddin Rehan Ahmed**, Julia Bonn, Kristin Wright-Bettner, Martha Palmer, and James H. Martin. *CAMRA: Copilot for AMR Annotation*. EMNLP-System Demo 2023
7. **Shafiuddin Rehan Ahmed**, Abhijnan Nath, Michael Regan, Adam Pollins, Nikhil Krishnaswamy, and James H. Martin. *How Good Is the Model in Model-in-the-loop Event Coreference Resolution Annotation?* LAW-XVII 2024
8. **Shafiuddin Rehan Ahmed**, Abhijnan Nath, James H. Martin, and Nikhil Krishnaswamy.  *$2n$  is better than  $n^2$ : Decomposing Event Coreference Resolution into Two Tractable Problems*. ACL-Findings 2023
9. Cecilia Mauceri, **Shafiuddin Rehan Ahmed**, Timothy O'Gorman, Chris Koski, Peter G. Anick, David White, and Martha Palmer. *RAMFIS: Integrating Diverse TA1's*. TAC 2019
10. Cecilia Mauceri, **Shafiuddin Rehan Ahmed**, and Timothy O'Gorman. *RAMFIS Sys Report*. TAC 2018

## REVIEWER DUTIES

---

ACL 2024 · EMNLP 2023, 24 · StarSEM 2023, 24

## COURSES

---

Deep Learning · Natural Language Processing · Algorithms · Machine Learning · Programming Languages · Probabilistic Models · Numerical Linear Algebra · Transformer Models · Probabilistic Programming Languages

## IMPORTANT LINKS

---

- Portfolio/Personal: <https://ahmeshaf.github.io>
- GitHub: <https://github.com/ahmeshaf>
- Google Scholar: <https://scholar.google.com/citations?user=t3IALyoAAAAJ>
- Semantic Scholar: <https://www.semanticscholar.org/author/Shafiuddin-Rehan-Ahmed/1418353489>
- LinkedIn: <https://linkedin.com/in/ahmeshaf>
- Hugging Face: <https://huggingface.co/ahmeshaf/>
- E-mail: [shah7567@colorado.edu](mailto:shah7567@colorado.edu)
- Phone: +1 720 643-9283