

Adithya Pratapa

Ph.D. student
Language Technologies Institute
School of Computer Science
Carnegie Mellon University

apratapa.xyz
vpratapa@cs.cmu.edu

Research Interests

Long-context language modeling, retrieval-augmented generation, efficient NLP

Education

Carnegie Mellon University, Pittsburgh, US

Ph.D., Language and Information Technologies, 2025 (expected)

M.S., Language Technologies, 2021

Indian Institute of Technology, Mumbai, India

B.Tech with Honors, Computer Science and Engineering, 2017

Industry Experience

Research Intern, Bloomberg AI – Search Team, New York

Summer 2024

Long-context and compression methods for large-scale summarization.

Applied Scientist Intern, Amazon Alexa – Web Information Team, Seattle

Summer 2022

Timeline summarization of events.

Research Fellow, Microsoft Research, Bangalore, India

2017–2019

Code-mixed language modeling and linguistically-grounded data augmentation.

Awards and Recognition

LTI Ph.D. Fellowship, 2024

Outstanding Paper, EMNLP 2023

Best Paper, Multilingual Information Access Workshop, NAACL 2022

Selected Publications

Adithya Pratapa, Teruko Mitamura. [Scaling Multi-Document Event Summarization: Evaluating Compression vs. Full-Text Approaches.](#)

NAACL 2025 Main

Sang Keun Choe, Hwijeen Ahn, Juhan Bae, Kewen Zhao, Minsoo Kang, Youngseog Chung, **Adithya Pratapa**, Willie Neiswanger, Emma Strubell, Teruko Mitamura, Jeff Schneider, Eduard Hovy, Roger Grosse, Eric Xing. [What is Your Data Worth to GPT? LLM-Scale Data Valuation with Influence Functions.](#)

arXiv preprint 2024

Adithya Pratapa, Kevin Small, Markus Dreyer. [Background Summarization of Event Timelines.](#)
EMNLP 2023 Main

Outstanding Paper

Yanlin Feng, **Adithya Pratapa**, David R Mortensen. [Calibrating Seq2seq Models for Ultra-fine Entity Typing.](#)

EMNLP 2023 Findings

Jiefu Ou, **Adithya Pratapa**, Rishubh Gupta, Teruko Mitamura. [Hierarchical Event Grounding.](#)
AAAI 2023

Adithya Pratapa, Rishubh Gupta, Teruko Mitamura. [Multilingual Event Linking to Wikidata.](#)
Multilingual Information Access Workshop @ NAACL 2022

Best Paper

Adithya Pratapa, Zhengzhong Liu, Kimihiro Hasegawa, Linwei Li, Yukari Yamakawa, Shikun Zhang, Teruko Mitamura. [Cross-document Event Identity via Dense Annotation.](#)

CoNLL 2021

Professional Service

Mentoring

Research advisor to Masters students (CMU, 2021–22, 2023–24)

Graduate student mentor (CMU, 2021–25)

Conference Reviewing

ARR (2022–25), EMNLP (2022–23), ACL 2023, AAAI (2023–24), NAACL 2022

Teaching Assistant

Advanced NLP (CMU, Spring 2024)

Question Answering (CMU, Spring 2022)

updated March 11, 2025