

# TANMAY PAREKH

@ tparekh@cs.ucla.edu

@ tparekh97@gmail.com

⌚ <https://tanmayparekh.github.io/>

in [linkedin.com/in/tanmayparekh/](https://linkedin.com/in/tanmayparekh/)

## EDUCATION

University of California Los Angeles (UCLA)

**Doctor of Philosophy in Computer Science**

📅 2021-2026

☑ 4.0/4.0

Carnegie Mellon University (CMU)

**Master of Science in Language Technologies**

📅 2019-2021

☑ 4.02/4.0

Indian Institute of Technology Bombay (IITB)

**B. Tech in Computer Science & Engineering**

📅 2014-2018

☑ 9.37/10

## ACHIEVEMENTS & ROLES

- Recipient of three PhD fellowships: **Amazon Science Ph.D. Fellowship**, **Bloomberg Ph.D. Data Science Fellowship**, and **UCLA Computer Science Fellowship**
- Served as **Program Chair** for SoCal NLP Symposium '23
- Co-leading the organization of **UCLA NLP Seminar Series**
- Worked as a Teaching Assistant for 10+ courses

## RESEARCH PROJECTS

### EVENT-CENTRIC NEWS NARRATIVE EVALUATION AND SAFE PERSONALIZED GENERATION

- Utilizing fine-grained event analysis to study and improve the factual faithfulness and persuasiveness of LLMs' generations for safely personalizing news narratives

### DIVERGENT-CONVERGENT LLM REASONING FOR ZERO-SHOT EVENT DETECTION

- Proposed *divergent reasoning* to explore without constraints to boost recall, and *convergent reasoning* to ground the divergent thoughts to boost precision, in LLM-guided Event Detection
- Achieved best zero-shot results on 6 datasets across 5 domains

### DOMAIN-AWARE DATA GENERATION FOR LOW RESOURCE INFORMATION EXTRACTION

- Synthesized data using LLMs by unsupervised extraction and conditioning on domain-specific cues for information extraction. Showed the efficacy across three domains and three languages.

### CONTEXTUAL LABEL PROJECTION FOR CROSS-LINGUAL STRUCTURE PREDICTION [Best Paper Nomination]

- Utilized instruction-tuned language models to perform contextual machine translation for cross-lingual label projection
- Showed improvements for Event Argument Extraction (EAE) and Named Entity Recognition (NER) for 50 languages

### MULTILINGUAL EVENT EXTRACTION FROM SOCIAL MEDIA FOR EPIDEMIC PREPAREDNESS

- Created the first Multilingual Event Detection framework SPEED for epidemic events in social media for 65 languages

### BENCHMARKING GENERALIZABILITY FOR EVENT ARGUMENT EXTRACTION

- Created GENEVA, a diverse and comprehensive dataset with 100+ event types and argument roles using expert human annotations from a semantic role labeling dataset FrameNet

## RESEARCH INTERESTS

Planning, Code Generation, Reasoning, Efficiency in LLMs, Personalized Generation, Data Generation

## INDUSTRY EXPERIENCE

Data Science Intern

📅 Jun '25 – Sep '25

**Bloomberg AI**

- Developing a **software-engineering-inspired agentic workflow** for low-latency parallel exploration and complex code generation
- Achieved SOTA on the text-to-SQL Spider 2.0 dataset

Research Scientist Intern

📅 Jun '24 – Sep '24

**Meta (GenAI)**

- Fine-tuned LLMs to encourage **dynamic planning** to optimize performance and cost for different prompt/tools
- Showed the efficacy of our technique on four prompting/tool strategies on three question-answering datasets

Applied Scientist Intern

📅 Jun '22 – Sep '22

**Amazon (Alexa AI)**

- Explored the alignment of higher-order semantics across languages in multilingual models; showed the effectiveness of our alignment on sentiment analysis

Applied Scientist

📅 Jul '18 – Jun '19

**Amazon (Machine Learning)**

- Proposed semi-supervised learning and regularization techniques for product attribute extraction from product titles without using partially labeled data

## SELECTED PUBLICATIONS

- T Parekh, et. al., "PExA: Parallel Exploration Agent for Complex Text-to-SQL", under submission
- T Parekh, et. al., "DiCoRe: Enhancing Zero-shot Event Detection via Divergent-Convergent LLM Reasoning", In Proceedings of EMNLP 2025
- T Parekh, et. al., "SNaRe: Domain-aware Data Generation for Low-Resource Event Detection", In Proceedings of EMNLP 2025
- T Parekh, et. al., "Dynamic Strategy Selection for Efficient Question Answering with Large Language Models", in Proceedings of NAACL Findings 2025
- T Parekh, et. al., "SPEED++: A Multilingual Event Extraction Framework for Epidemic Prediction and Preparedness", in Proceedings of EMNLP 2024
- T Parekh, et. al., "Contextual Label Projection for Cross-Lingual Structure Extraction", in Proceedings of NAACL 2024 (**BEST PAPER NOMINATION**)
- T Parekh, et. al., "GENEVA: Pushing the Limit of Generalizability for Event Argument Extraction with 100+ Event Types", in Proceedings of ACL 2023
- A Madaan\*, A Setlur\*, T Parekh\*, et. al., "Politeness Transfer: Tag and Generate Approach", in Proceedings of ACL 2020