

Akshay Parekh, Ph.D.
Consultant, NLP Applied Research Scientist
email: sendtoakshayparekh@gmail.com
mobile: +91 8462000540

I am an NLP Data Scientist in the Advanced Intelligence (AI) team at Eli Lilly & Co., Bangalore, where I have been working full-time since June 2022 after joining as an intern in February 2022. I am currently exploring reasoning models, reinforcement learning, and efficient small models for domain-specific applications, and applying these through the development of agentic pipelines for drug discovery computation models and automated first-draft document generation.

CURRENT ROLE

- **Advanced Intelligence (AI), Eli Lilly & Co.** Bangalore, India
NLP Data Scientist *Feb 2022 - Present*
 - **Advisor Data Scientist:** Squad Lead for AI Authoring, AI for Drug Discovery
 - **Consultant Data Scientist:** Generative AI, Information Extraction & Retrieval, Knowledge Graphs, and Single cell Transformer Models
 - **NLP Research Intern:** Biomedical Information Retrieval, Unsupervised Key-Phrase Extraction, Transformers, BioClinicalBERT, Biomedical Named Entity Recognition

RECENT EDUCATION

- **Indian Institute of Technology Guwahati** Guwahati, India
Ph.D., Thesis Topic: Large Scale Relation Extraction, Advisors: Dr. Ashish Anand & Dr. Amit Awekar *Jul 2016 – 2023*

KEY PUBLICATIONS

Patent Application

- **Neural-Based Ontology Generation and Refinement:** Balaji Ganesan, Riddhiman Dasgupta, **Akshay Parekh**, Hima Patel, Berthold Reinwald - US-11520986-B

Conference

- **Taxonomical hierarchy of canonicalized relations from multiple Knowledge Bases.:** **Akshay Parekh**, Ashish Anand, Amit Awekar - CoDS-COMAD 2020
- **Improving Relation Classification using Relation Hierarchy.:** **Akshay Parekh**, Ashish Anand, Amit Awekar - NLDB 2022
- **Document Structure aware Relational Graph Convolutional Networks for Ontology Population.:** Abhay M Shalghar, Ayush Kumar, Shobha G, Balaji Ganesan, Aswin Kannan, **Akshay Parekh** - DLG4NLP Workshop at ICLR 2022
- **Relation Classification Model Performance as a Function of Dataset Label Noise.:** **Akshay Parekh**, Ashish Anand, Amit Awekar - CoDS-COMAD 2023

TECHNICAL SKILLS

- **Languages:** Python, C
- **Data Analysis and Visualization:** Numpy, Pandas, Matplotlib, Seaborn
- **ML-DL:** Scikit-learn, PyTorch
- **NLP & Knowledge Graphs:** NLTK, spacy, gensim, transformers, langchain, neo4j, & cypher

Full CV available upon request.