

Sandeep Kumar



Department of Computer Science & Engineering
Indian Institute of Technology Patna, India

Email: sandeep_2121cs29@iitp.ac.in
[Google Scholar](#)
Links: [Github](#), [LinkedIn](#)
Phone: +918294544941

Education

Indian Institute of Technology, Patna <i>Ph.D. Computer Science (Prime Minister's Research Fellow)</i> Advisor: Dr. Asif Ekbal GPA: 8.5/10.0	2021 - 2025
Indian Institute of Technology, Patna <i>Masters in Mathematics and Computing</i> Advisor: Dr. Asif Ekbal GPA: 8.86/10.0	2019 - 2021
Bangalore Institute of Technology, Bangalore <i>Bachelors in Computer Science</i> Aggregate: 60.6%	2013 - 2017

Research and Work Experience

Samsung Research (SRIB), Bangalore PhD. Internship Topic: On Human AI Alignment	April 2024- Oct 2024
Oak Ridge National Laboratory, USA Research Collaboration Topic: LLM in Science	Jan 2024- present
Lexitas Legal, USA Machine Learning Research Intern Topic: NLP and ML techniques for Transcript Analysis	June 2021- March 2022

Publications

Sandeep Kumar, Abhijit A Nargund, Vivek Sridhar, “ CourtEval: A Courtroom-Based Multi-Agent Evaluation Framework,” accepted at ACL 2025 (Findings)

Sandeep Kumar, Mohit Sahu, Gunnet Singh Kohli, Tirthankar Ghosal, Asif Ekbal, “ Toward Watermarking Peer Reviews Generated by Large Language Models,” Submitted in ARR 2025

Sandeep Kumar, Samarth Garg, Sagnik Sengupta, Tirthankar Ghosal, Asif Ekbal, “MixRevDetect: Towards Detecting AI-Generated Content in Hybrid Peer Reviews. ,” NAACL 2025 [Paper](#)

Sandeep Kumar, Tirthankar Ghosal, Vinayak Goyal and A. Ekbal, “Can Large Language Models Unlock Novel Scientific Research Ideas? ,” Submitted in ARR 2025 [Paper](#)

Sandeep Kumar, Mohit Sahu, Vardhan Gacche, Tirthankar Ghosal, Asif Ekbal, “ 'Quis custodiet ipsos custodes?' Who will watch the watchmen? On Detecting AI-generated peer-reviews” in The 2024 Conference on *Empirical Methods in Natural Language Processing* Main , to be presented in Nov. 2024. EMNLP 2024 [Paper](#)

Sandeep Kumar, Gunnet Singh Kohli, Tirthankar Ghosal and A. Ekbal, “ Longform Multimodal Lay Summarization of Scientific Papers: Towards Automatically Generating Science Blogs from Research Articles,” LREC-COLING 2024 [Paper](#)

Sandeep Kumar, T. Ghosal, and A. Ekbal, “When reviewers lock horns: Finding disagreements in scientific peer reviews,” in The 2023 Conference on *Empirical Methods in Natural Language Processing* Main , Dec. 2023. EMNLP 2023 [Paper](#)

Sandeep Kumar, G. Singh, T. Ghosal, and A. Ekbal, “Mup-scidocsum: Leveraging multi-perspective peer review summaries for scientific document summarization,” in Accepted at The 25th *International Conference on Asia-Pacific Digital Libraries* , Online, Dec. 2023. ICADL 2023 [Paper](#)

Sandeep Kumar, T. Ghosal, and A. Ekbal, “APCS: Towards argument based pros and cons summarization of peer reviews,” in The *2nd Workshop on Information Extraction from Scientific Publications, IJCNLP-AACL*, Nov. 2023 WIESP@IJCNL-AACL 2023 [Paper](#)

Sandeep Kumar, T. Ghosal, and A. Ekbal, “DeepMetaGen: an unsupervised deep neural approach to generate template-based meta-reviews leveraging on aspect category and sentiment analysis from peer reviews,” Mar. 2023. IJDL Journal 2023 [Paper](#)

Sandeep Kumar, G. S. Kohli, K. Shinde, and A. Ekbal, "Team AINLPML@ MuP in SDP 2022: Scientific Document Summarization by End-to-End Extractive and Abstractive Approach" *Proceedings of the Third Workshop on Scholarly Document Processing*.
SDP 2023@ACL 2022 [Paper](#)

Sandeep Kumar, Hardik Arora, Tirthankar Ghosal, Asif Ekbal "DeepASPeer: towards an aspect-level sentiment controllable framework for decision prediction from academic peer reviews" *22nd Joint Conference on Digital Libraries*
ACM/IEEE JCDL 2022 (🏆 Best Student Paper Award 🏆) [Paper](#)

Tirthankar Ghosal, **Sandeep Kumar**, Prabhat Kumar Bharti, Asif Ekbal "Peer review analyze: A novel benchmark resource for computational analysis of peer reviews" *Public Library of Science PLOSOne Journal* 2022 [Paper](#)

Sandeep Kumar, Tirthankar Ghosal, Prabhat Kumar Bharti, Asif Ekbal "Sharing is Caring! Joint Multitask Learning Helps Aspect-Category Extraction and Sentiment Detection in Scientific Peer Reviews" *Joint Conference on Digital Libraries*
ACM/IEEE JCDL 2021 [Paper](#)

Hardik Arora, Tirthankar Ghosal, **Sandeep Kumar**, Suraj Patwal, Phil Gooch "INNOVATORS at SemEval-2021 Task-11: A Dependency Parsing and BERT-based model for Extracting Contribution Knowledge from Scientific Papers" *Proceedings of the 15th International Workshop on Semantic Evaluation*
SemEval-2021 [Paper](#)

Sandeep Kumar, Tirthankar Ghosal, Asif Ekbal "DataQuest: An Approach to Automatically Extract Dataset Mentions from Scientific Papers" Towards Open and Trustworthy Digital Societies: *23rd International Conference on Asia-Pacific Digital Libraries*, 23
ICADL 2021 [Paper](#)

Research Interests

Research Goal: My long-term goal is to develop AI tools that can assist the scientific community and prevent the misuse of AI.

Broad: Natural Language Processing, Deep Learning, Summarization, Recommender Systems

Specific: NLP for Scientific Literature

Awards and Achievements

- Awarded Prime Minister's Research Fellowship (PMRF) Issued by MoE (Received by top 1% researchers in India) 2021
- Won Best Student Paper Award, for DeepASPeer at JCDL 2022, Cologne, Germany
- Rank 1 SDP@COLING, Our team AINLPML achieved Rank 1 in the MuP shared task at the Scholarly Document Processing workshop at COLING 2022
- World Rank 5 in the Task 11 of SemEval competition 2021 among 115 participating teams

- Kaggle Rank 53 out of 1608 teams in kaggle competition held by Coleridge Initiative, New York University, United States
- Secured a Gate Score of 637 and a rank of 1563 in Gate 2019 among 1.30lac candidates
- **Grants:** Microsoft Travel Grant 2023, Student Travel Grant and SIGIR 2022 Student Travel Grant; D&I grant - ACL

Position of Responsibility

Teaching

Foundation of Machine Learning/Information Extraction in Mother's Institute of Technology (AICTE approved Technical Institute at Bihta)

Teaching Assistance

Artificial Intelligence (Spring 2022, Spring 2021), Natural Language Processing (Autumn 2023, Autumn 2020), C++ Programming (Spring 2023), Computer Architecture (Autumn 2021)

Academic Service

- Volunteer of ICONIP 2023
- PC member of SIGDial 2021 Special Session on Summarizing of Dialogues and Multi-Party Meetings
- Reviewer (Journals): Scientometrics Springer 2022, Computer Speech & Language
- Reviewers (Conference): ICONIP 2023, AAAI 2023, IJCNN 2023, SDP 2022 IJCNN 2022, CIKM 2020 Resource Track
- Mentor: Mentoring 2 undergraduate students as a part of AI-NLP-ML Lab, IIT Patna

Technical Skills

Programming

Python, C, C++

Technologies

MySQL, Flask, Docker, Git, CUDA, OpenMP, StanfordCoreNLP, Spacy, Wordnet, NLTK, OpenCV, Gensim, Sklearn, Pytorch.

Personal Information

Languages Known: Hindi, English