Pawan Goyal

Contact Information	Room 308, Department of Computer Science and Enginnering IIT Kharagpur, Kharagpur, West Bengal – 721302

E-mail:
pawang@cse.iitkgp.ac.in
pawang.iitk@gmail.com
DOB: - August 10, 1986

Research Interests

Computational Linguistics, Text Mining, Information Retrieval

Education

Ph.D. July, 2007 - September, 2011
Intelligent Systems Research Center, Faculty of Computing and Engineering, University of Ulster, NI, UK, September 2011
- Thesis Topic: *Analytic Knowledge Discovery Techniques for Ad-Hoc Information Retrieval and Text Summarization*
- Supervisors: Professor L. Behera and Prof. T. M. McGinnity

B. Tech. July, 2003 - May, 2007
Electrical Engineering, IIT Kanpur, Kanpur, U.P., India, May 2007
- CPI: 9.1/10 (overall), 9.8/10 (EE major)

Academic Appointments

Associate Professor March, 2020 - Present
Department of Computer Science and Engineering,
IIT Kharagpur

Assistant Professor July, 2013 - March, 2020
Department of Computer Science and Engineering,
IIT Kharagpur

Postdoctoral Researcher October 2011 to July, 2013
Department of Computer Science and Applied Mathematics,
INRIA Paris Rocquencourt
- Certification of Sanskrit Morphology
- Advisor : Professor Gérard Huet

Awards

Faculty Excellence Award (Associate Professor) from IIT Kharagpur for the year 2022.

ACM HyperText Ted Nelson Newcomer Award, 2021.

INAE Young Engineer Award 2020.

Google India AI/ML Research Awards 2020.

Facebook AI and Ethics Research Award India, 2019.

WebSci Best Paper Award (honorable mention), 2019.

GYTI Appreciation from SRISTI, 2017.

Academic Excellence Award from IIT Kanpur for the year 2005-2006.

Scientific Responsibilities

Senior PC / Meta-reviewer: AAAI 2021, AAAI 2023

Organization: Organized 6th International Sanskrit Computational Linguistics from October 23-25, 2019 at IIT Kharagpur.

PC Member for Top Conferences (Recent): ACL 2023, EMNLP 2022, SIGIR 2022, EMNLP 2021, SIGIR 2021, ACL 2021, NAACL 2021, SIGIR 2020, ACL 2020, IJCAI 2020

Sponsored Projects as Principal Investigator	2022-2025

2022-2025

An Interpretable Unified Framework for Text-to-Text Translation among Indian Languages using Sanskrit-based Interlingua Representation
- Sponsoring Agency: MeitY
- Project Cost: INR 3,43,76,700 (Consortium project with 6 institutes)

2022-2025

Voice Based Natural Interaction for Goal Oriented Tasks in Healthcare
- Sponsoring Agency: I-Hub Foundation for Cobotics
- Project Cost: INR 75,00,000

2022-2023

Microsoft Academic Partnership Grant (MAPG 2022)
- Sponsoring Agency: Microsoft India
- Project Cost: INR 15,00,000

2021-2022

Microsoft Academic Partnership Grant (MAPG 2021)
- Sponsoring Agency: Microsoft India
- Project Cost: INR 10,00,000

2019-2021

Unrestricted Research Grant for Research in Dialog Response Generation
- Sponsoring Agency: Google India
- Project Cost: INR 15,00,000

2021-2023

Fintalk: Automated text highlight generation from financial data
- Sponsoring Agency: Goldman Sachs
- Project Cost: INR 70,80,000

2019-2021

Conversation Agents for Urban Navigation
- Sponsoring Agency: IMPRINT 2, SERB.
- Project Cost: INR 68,20,000

2016-2021

Teaching Learning Centre for Internet of Things
- Sponsoring Agency: MHRD, Department of Higher Education, New Delhi.
- Project Cost: INR 101,40,000

2016-2019

Aspect Based Opinion Extraction for Automated Product Comparison
- Sponsoring Agency: DST - SERB (Early Career)
- Project Cost: INR 18,40,000

2020-2021

Advanced Natural Language Understanding
- Sponsoring Agency: Samsung India Pvt. Ltd.
- Project Cost: INR 36,40,000

2018-2019

Self Evolving and Language Independent Knowledge Graphs
- Sponsoring Agency: Samsung India Pvt. Ltd.
- Project Cost: INR 35,40,000

2018-2023

AI/ML Techniques for Online Shopping
- Sponsoring Agency: Flipkart India Pvt. Ltd.
- Project Cost: INR 22,60,000

2019-2020

Natural Language Processing for the Data and texts for Material Science Literature
- Sponsoring Agency: IKST, Bangalore
- Project Cost: INR 16,49,000

2015-2020

Unrestricted Grant for Research in Review Mining
- Flipkart India Pvt. Ltd.
- Project Cost: INR 12,57,000

2014-2019

IndicView: Because Language is no more a barrier
- Google India Pvt. Ltd.
- Project Cost: INR 5,00,000

2018-2019

Multiuser DeInterleaving in SmartHomes
- Samsung R&D, Bangalore
- Project Cost: INR 6,00,000

TEACHING

- CS21004: Formal Language and Automata Theory (4 times)
- CS21003: Algorithms-1 theory and Lab (5 times)
- CS10001: Programming and Data Structures (2 times)
- CS19002: Programming and Data Structures Laboratory (5 times)
- CS60057: Speech and Natural Language Processing (5 times)
- CS60017: Social Computing (3 times)
- CS60092: Information Retrieval (4 times)
- CS60010: Deep Learning (once)
- **Top Teaching Feedback**
    - 2nd in theory courses (100 or more): FLAT (2016)
    - 2nd in first year undergraduate lab courses: PDS Lab (2019), PDS Lab (2015)
    - 4th in lab courses (100 or more): Algorithms-1 Lab (2017)

REFEREED
JOURNAL
PUBLICATIONS

Roy, Aniruddha, Sharma, Isha, Sarkar, Sudeshna and Goyal, Pawan. Meta-ED: Cross-lingual Event Detection using Meta-learning for Indian Languages. *ACM Transactions on Asian and Low-Resource Language Information Processing*, Vol. 22, No. 2, Feb 2023.

Das, Kishalay, Samanta, Bidisha, Goyal, Pawan, Lee, Seung-Cheol, Bhattacharjee, Satadeep and Ganguly, Niloy. CrysXPP: An Explainable Property Predictor for Crystalline Materials. *npj Computational Materials*, Vol. 8, No. 1, 2022.

Jana, Abhik, Haldar, Siddhant and Goyal, Pawan. Network Embeddings from Distributional Thesauri for Improving Static Word Representations. *Expert Systems with Applications*, Vol. 187, January 2022.

Nayak, Tapas, Majumder, Navonil, Goyal, Pawan and Poria, Soujanya. Deep Neural Approaches to Relation Triplets Extraction: A Comprehensive Survey. *Cognitive Computation*, Vol. 13, pp. 1215-1232, 2021.

Ghosh, Krishnendu, Nangi, Sharmila Reddy, Kanchugantla, Yashasvi, Rayapati, Pavan Gopal, Bhowmick, Plaban Kumar and Goyal, Pawan. Augmenting Video Lectures: Predicting Off-topic Concepts and Linking to Relevant Video Lecture Segments. *International Journal of Artificial Intelligence in Education*, Springer, 2021.

Guha, Souradip, Mullick, Ankan, Agrawal, Jatin, Ram, Swetarekha, Ghui, Samir, Lee, Seung-Cheol, Bhattacharjee, Satadeep and Goyal, Pawan. MatScIE: An automated tool for the generation of databases of methods and parameters used in the computational materials science literature. *Computational Materials Science, Elsevier*, 2021.

Das, Mithun, Mathew, Binny, Saha, Puyajoy, Goyal, Pawan and Mukherjee, Animesh. Hate speech in online social media. *ACM SigWeb Newsletter*, Nov. 2020.

Krishna, Amrith, Santra, Bishal, Gupta, Ashim, Satuluri, Pavankumar and Goyal, Pawan. A Graph Based Framework for Structured Prediction Tasks in Sanskrit. *Computational Linguistics, MIT Press*, Vol. 46, No. 4, Dec 2020.

Jana, Abhik, Mukherjee, Animesh and Goyal, Pawan. Network Measures: A New Paradigm Towards Reliable Novel Word Sense Detection. *Information Processing and Management, Elsevier*, Vol. 57, No. 6, 2020.

Pandey, Pradumn Kumar, Singh, Mayank, Goyal, Pawan, Mukherjee, Animesh and Chakrabarti, Soumen. Analysis of Reference and Citation Copying in Evolving Bibliographic Networks. *Journal of Informetrics, Elsevier*, vo. 14, no. 1, 2020.

Rudra, Koustav, Goyal, Pawan, Ganguly, Niloy, Imran, Mohammad and Mitra, Prasenjit. Summarizing Situational Tweets in Crisis Scenarios: An Extractive-Abstractive Approach. *IEEE Transactions on Computational Social Systems*, Vol. 6, No. 5, 2019, pp. 981-993.

Paheli Bhattacharya, Pawan Goyal and Sudeshna Sarkar. Using Communities of Words Derived from Multilingual Word Vectors for Cross-Language Information Retrieval in Indian Languages. *ACM Transactions on Asian and Low-Resource Language Information Processing*. Vol. 18, No. 1, 2019.

Maity, Suman Kalyan, Chakraborty, Aishik, Goyal, Pawan and Mukherjee, Animesh. *Opinion Conflicts: An Effective Route to Detect Incivility in Twitter*, Proceedings of the ACM on Human-Computer Interaction, 2018, Vol. 2, No. CSCW, pp. 1–27.

Ankan Mullick, Pawan Goyal, Niloy Ganguly and Manish Gupta. Harnessing Twitter for Answering Opinion List Queries. *IEEE Transactions on Computational Social Sciences*, Vol. 5, No. 4, 2018.

Koustav Rudra, Niloy Ganguly, Pawan Goyal and Saptarshi Ghosh. Extracting and Summarizing Situational Information from the Twitter Social Media during Disasters. *ACM Transactions on the Web*. Vol. 12, No. 3, 2018.

Mayank Singh, Tanmoy Chakraborty, Animesh Mukherjee and Pawan Goyal. Is this conference a top-tier? ConfAssist: An assistive conflict resolution framework for conference categorization. *Journal of Informetrics*. Vol. 10, no. 4, pp. 1005 – 1022, November, 2016.

Pawan Goyal and Gérard Huet. Design and analysis of a lean interface for Sanskrit corpus annotation. *Journal of Language Modeling*, Vol. 4, No 2, pp . 145 – 182, 2016.

Bhattacharya, Paheli, Pawan Goyal, and Sudeshna Sarkar. Using word embeddings for query translation for hindi to english cross language information retrieval. *Computation & Systems*, Vol. 20. No. 3 (2016): pp. 435 – 447.

Tanmoy Chakraborty, Suhansanu Kumar, Pawan Goyal, Niloy Ganguly and Animesh Mukherjee. On the categorization of scientific citation profiles in computer sciences. *Communications of the ACM*. Vol 58, no. 9, pp. 82–90, September, 2015.

Siddhartha R Jonnalagadda, Pawan Goyal and Mark D Huffman. Automating Data Extraction in Systematic Reviews: A Systematic Review. *Systematic Reviews*. 4:78, June, 2015.

Sunny Mitra, Ritwik Mitra, Suman Kalyan Maity, Martin Riedl, Chris Biemann, Pawan Goyal and Animesh Mukherjee. An automatic approach to identify word sense changes in text media across timescales. *JNLE special issue on Graph Methods for NLP*. April 2015

Pawan Goyal, Laxmidhar Behera and T. M. McGinnity. Query Representation through Lexical Association for Information Retrieval. *IEEE Transactions on Knowledge and Data Engineering.* Vol. 24, no. 12, pp. 2260 - 2273, Dec. 2012.

Pawan Goyal, Laxmidhar Behera and T. M. McGinnity. A Novel Neighborhood Based Document Smoothing Model for Information Retrieval. *Information Retrieval, Springer.* Vol. 16, Issue 3, pp. 391- 425, May 2013.

Pawan Goyal, Laxmidhar Behera and T. M. McGinnity. A Context based Word Indexing Model for Document Summarization. *IEEE Transactions on Knowledge and Data Engineering*, Vol. 25, no. 8, pp. 1693 - 1705, Aug. 2013.

Pawan Goyal, Laxmidhar Behera and Martin McGinnity. Application Of Bayesian Framework In Natural Language Understanding. *IETE Technical Review*, 25(5):251–269, 2008.

SELECTED CONFERENCE PUBLICATIONS

Sandhan, Jivnesh, Behera, Laxmidhar and Goyal, Pawan.. Systematic Investigation of Strategies Tailored for Low-Resource Settings for Low-Resource Dependency Parsing. *EACL (short)*, 2023.

Mullick, Ankan, Mondal, Ishani, Ray, Sourjyadip, R, Raghav, Chaitanya, G Sai and Goyal, Pawan. Intent Identification and Entity Extraction for Healthcare Queries in Indic Languages. *EACL findings*, 2023.

Das, Kishalay, Samanta, Bidisha, Goyal, Pawan, Lee, Seung-Cheol, Bhattacharjee, Satadeep and Ganguly, Niloy. CrysGNN : Distilling pre-trained knowledge to enhance property prediction for crystalline materials. *AAAI*, 2023.

Mukherjee, Rajdeep, Bohra, Abhinav, Banerjee, Akash, Sharma, Soumya, Hegde, Manjunath, Shaikh, Afreen, Shrivastava, Shivani, Dasgupta, Koustuv, Ganguly, Niloy, Ghosh, Saptarshi and Goyal, Pawan. ECTSum: A New Benchmark Dataset For Bullet Point Summarization of Long Earnings Call Transcripts. *EMNLP*, 2022.

Sandhan, Jivnesh, Singla, Rathin, Rao, Narein, Samanta, Suvendu, Behera, Laxmidhar and Goyal, Pawan. TransLIST: A Transformer-Based Linguistically Informed Sanskrit Tokenizer. *EMNLP findings*, 2022.

Shukla, Abhay, Bhattacharya, Paheli, Poddar, Soham, Mukherjee, Rajdeep, Ghosh, Kripabandhu, Goyal, Pawan and Ghosh, Saptarshi. Legal Case Document Summarization: Extractive and Abstractive Methods and their Evaluation. *AACL*, 2022.

Kar, Debanjana, Sarkar, Sudeshna and Goyal, Pawan. ArgGen: Prompting Text Generation Models for Document-Level Event-Argument Aggregation. *AACL findings (short)*, 2022.

Chakraborty, Souvic, Goyal, Pawan and Mukherjee, Animesh. Fast Few shot Self-attentive Semi-supervised Political Inclination Prediction. *ICADL*, 2022.

Sandhan, Jivnesh, Gupta, Ashish, Terdalkar, Hrishikesh, Sandhan, Tushar, Samanta, Suvendu, Behera, Laxmidhar and Goyal, Pawan. A Novel Multi-Task Learning Approach for Context-Sensitive Compound Type Identification in Sanskrit. *Coling*, 2022.

Roy, Aniruddha, Thankur, Rupak Kumar, Sharma, Isha, Gupta, Ashim, Krishna, Amrith, Sarkar, Sudeshna and Goyal, Pawan (2022). Does Meta-learning Help mBERT for Few-shot Question Generation in a Cross-lingual Transfer Setting for Indic Languages? *Coling (short)*, 2022.

Kumar, Rishabh, Adiga, Devaraja, Ranjan, Rishav, Krishna, Amrith, Ramakrishnan, Ganesh, Goyal, Pawan and Jyothi, Preethi. Linguistically Informed Post-processing for ASR Error correction in Sanskrit. *Interspeech*, 2022.

Santra, Bishal, Roychowdhury, Sumegh, Mandal, Aishik, Gurram, Vasu, Naik, Atharva, Gupta, Manish and Goyal, Pawan. Representation Learning for Conversational Data using Discourse Mutual Information Maximization. *NAACL*, 2022.

Mullick, Ankan, Purkayastha, Sukannya, Goyal, Pawan and Ganguly, Niloy. A Framework to Generate High-quality Datapoints for Multiple Novel Intent Detection. *Findings of NAACL*, 2022.

Roy, Kalyani, Goel, Avani and Goyal, Pawan. Using Data Augmentation to Identify Relevant Reviews for Product Question Answering. *The ACM Web Conference (poster)*, 2022.

Paul, Shounak, Goyal, Pawan and Ghosh, Saptarshi. LeSICiN: A Heterogeneous Graph-based Approach for Automatic Legal Statute Identification from Indian Legal Documents. *AAAI*, 2022.

Mukherjee, Rajdeep, Vishnu, Uppada, Peruri, Hari Chandana, Ganguly, Niloy, Bhattacharya, Sourangshu, Goyal, Pawan and Rudra, Koustav. MTLVS: A Multi-Task Framework to Verify and Summarize Crisis-Related Microblogs. *WSDM*, 2022.

Mukherjee, Rajdeep, Nayak, Tapas, Butala, Yash, Bhattacharya, Sourangshu and Goyal, Pawan. PASTE: A Tagging-free Decoding Framework using Pointer Networks for Aspect Sentiment Triplet Extraction. *EMNLP*, 2021.

Nandy, Abhilash, Sharma, Soumya, Maddhashiya, Shubham, Sachdeva, Kapil, Goyal, Pawan and Ganguly, Niloy. Question Answering over Electronic Devices: A New Benchmark Dataset and a Multi-Task Learning based QA Framework. *EMNLP Findings*, 2021.

Das, Mithun, Saha, Puyajoy, Dutt, Ritam, Goyal, Pawan, Mukherjee, Animesh and Mathew, Binny. You too Brutus! Trapping Hateful Users in Social Media: Challenges, Solutions & Insights. *ACM HyperText*, 2021.

Singh, Shruti, Singh, Mayank and Goyal, Pawan. COMPARE: A Taxonomy and Dataset of Comparison Discussions in Peer Reviews. *JCDL (short)*, 2021.

Adiga, Devaraja, Kumar, Rishabh, Krishna, Amrith, Jyothi, Preethi, Ramakrishnan, Ganesh and Goyal, Pawan. Automatic Speech Recognition in Sanskrit: A New Speech Corpus and Modelling Insights. *Findings of ACL*, 2021.

Santra, Bishal, Potnuru, Anusha and Goyal, Pawan. Hierarchical Transformer for Task Oriented Dialog Systems. *NAACL-HLT*, 2021 (online).

Binny Mathew, Punyajoy Saha, Seid Muhie Yimam, Chris Biemann, Pawan Goyal and Animesh Mukherjee. HateXplain: A Benchmark Dataset for Explainable Hate Speech Detection. *AAAI*, Feb 2-9, 2021 (Online)

Rima Hazra, Hardik Aggarwal, Pawan Goyal, Animesh Mukherjee and Soumen Chakrabarti. Joint Autoregressive and Graph Models for Software and Developer Social Networks. *ECIR*, March 28th - April 1st, 2021 (Online).

Rajdeep Mukherjee, Shreyas Shetty, Subrata Chattopadhyay, Subhadeep Maji, Samik Datta and Pawan Goyal. Reproducibility, Replicability and Beyond: Assessing Production Readiness of Aspect Based Sentiment Analysis in the Wild. *ECIR: Reproducibility Track*, March 28th - April 1st, 2021 (Online).

Paul, Shounak, Goyal, Pawan and Ghosh, Saptarshi. Automatic Crime Identification from Facts: A Few Sentence-Level Crime Annotations is All You Need. *Coling*, December 8-11, 2020 (online).

Krishna, Amrith, Gupta, Ashim, Garasangi, Deepak, Satuluri, Pavankumar and Goyal, Pawan. Keep it Surprisingly Simple: A Simple First Order Graph Based Parsing Model for Joint Morphosyntactic Parsing in Sanskrit. *EMNLP (short)*, Nov. 16-20, 2020 (online).

Mathew, Binny, Illendula, Anurag, Saha, Punyajoy, Sarkar, Soumya, Goyal, Pawan and Mukherjee, Animesh. Hate begets Hate: A Temporal Study of Hate Speech. *CSCW*, October 17-21, 2020 (online).

Maji, Subhadeep, Kumar, Rohan, Bansal, Manish, Roy, Kalyani and Goyal, Pawan. Logic Constrained Pointer Networks for Interpretable Textual Similarity. *IJCAI-PRICAI*, January, 2021, Yokohama, Japan.

Mukherjee, Rajdeep, Peruri, Hari Chandana, Vishnu, Uppada, Goyal, Pawan, Bhattacharya, Sourangshu and Ganguly, Niloy. Read what you need: Controllable Aspect-based Opinion Summarization of Tourist Reviews. *SIGIR (short)*, July 25-30, 2020 (online).

Chakraborty, Souvic, Goyal, Pawan and Mukherjee, Animesh. Aspect-based Sentiment Analysis of Scientific Reviews. *JCDL*, August 1-5, 2020 (online).

Sharma, Soumya, Santra, Bishal, Jana, Abhik, Tokala, Santosh, Ganguly, Niloy and Goyal, Pawan. Incorporating Domain Knowledge into Medical NLI using Knowledge Graph, *EMNLP-IJCNLP (short)*, November 3-7, 2019, Hong Kong.

Jana, Abhik, Puzyrev, Dima, Panchenko, Alexander, Goyal, Pawan, Biemann, Chris and Mukherjee, Animesh. On the Compositionality Prediction of Noun Phrases using Poincaré embeddings, *ACL*, July 28th - August 3rd, 2019, Florence, Italy.

Krishna, Amrith, Sharma, Vishnu Dutt, Santra, Bishal, Satuluri, Pavan Kumar and Goyal, Pawan. Poetry to Prose Conversion in Sanskrit as a Linearisation Task: A case for Low-Resource Languages, *ACL (short)*, July 28th - August 3rd, 2019, Florence, Italy.

Maji, Subhadeep, Roy, Kalyani, Kumar, Rohan, Bansal, Manish, Kumar, Mohit and Goyal, Pawan. Addressing Vocabulary Gap in E-commerce Search, *SIGIR*, Paris, France, July 20-25, 2019, pp. 1073–1076.

Mathew, Binny, Dutt, Ritam, Goyal, Pawan and Mukherjee, Animesh. Spread of hate speech in online social media, *ACM WebSci*, June 30 - July 3, 2019, Boston, MA, USA (Best paper award honorable mention).

Mathew, Binny, Saha, Punyajoy, Tharad, Hardik, Rajgaria, Subham, Singhania, Prajwal, Maity, Suman Kalyan, Goyal, Pawan and Mukherjee, Thou shalt not hate: Countering online hate speech, *ICWSM*, June 11-14, 2019, Munich, Germany.

Singh, Mayank, Sarkar, Rajdeep, Vyas, Atharva, Goyal, Pawan, Mukherjee, Animesh and Chakrabarti, Soumen. Automated Early Leaderboard Generation From Comparative Tables, *ECIR*, April 14-18, 2019, Cologne, Germany.

Palod, Priyank, Patwari, Ayush, Bahety, Sudhanshu, Bagchi, Saurabh and Goyal, Pawan. Misleading Metadata Detection on YouTube, *ECIR*, April 14-18, 2019, Cologne, Germany (short paper).

Gupta, Ashim, Goyal, Pawan, Sarkar, Sudeshna and Gattu, Nandeshwar. Fully Contextualized Biomedical NER, *ECIR*, April 14-18, 2019, Cologne, Germany (short paper).

Maity, Suman Kalyan, Panigrahi, Abhishek, Ghosh, Sayan, Banerjee, Arundhati, Goyal, Pawan and Mukherjee, Animesh. DeepTagRec: A Content-cum-User based Tag Recommendation Framework for Stack Overflow, *ECIR*, April 14-18, 2019, Cologne, Germany (short paper).

Jana, Abhik, Mukherjee, Animesh and Goyal, Pawan. Detecting Reliable Novel Word Senses: A Network-Centric Approach, *ACM SAC Knowledge and Language Processing Track*, April 8-12, 2019, Limassol, Cyprus.

Amrith Krishna, Bishal Santra, Sasi Prasanth Bandaru, Sahu, Gaurav, Sharma, Vishnu Dutt, Satuluri, Pavan Kumar and Pawan Goyal. Free as in Free Word Order: An Energy Based Model for Word Segmentation and Morphological Tagging in Sanskrit, *EMNLP*, Brussels, Belgium, October 31-November 4, 2018.

Krishna, Amrith, Majumder, Bodhisattwa Prasad, Bhat, Rajesh and Goyal, Pawan. Upcycle Your OCR: Reusing OCRs for Post-OCR Text Correction in Romanised Sanskrit, *CoNLL*, October 31 - November 1, 2018, Brussels, Belgium.

Abhik Jana, Pranjal Kanojia, Pawan Goyal and Animesh Mukherjee. WikiRef: Wikilinks as a route to recommending appropriate references for scientific Wikipedia pages. *Coling*, Santa Fe, New Mexico, USA, August 20-26, 2018.

Rudra, Koustav, Goyal, Pawan, Ganguly, Niloy, Mitra, Prasenjit and Imran, Mohammad. Identifying Sub-events and Summarizing Information during Disasters. *SIGIR*, Ann Arbor, Michigan, July 8-12, 2018.

Singh, Mayank, Dogga, Pradeep, Patro, Sohan, Barnwal, Dhiraj, Dutt, Ritam, Haldar, Rajarshi, Goyal, Pawan and Mukherjee, Animesh. CL Scholar: The ACL Anthology Knowledge Graph Miner. *NAACL-HLT (Demo)*, New Orleans, Louisiana, June 1-6, 2018.

Pani, Sandeep Kumar, R, Naresh, Goyal, Pawan and Bhowmick, Plaban Kumar. Learning to Extract Comparison Points of Entity Pairs from Wikipedia Articles. *JCDL (Poster)*, Fort Worth, Texas, June 3-6, 2018.

Jana, Abhik and Goyal, Pawan. Can Network Embedding of Distributional Thesaurus be Combined with Word Vectors for Better Representation?. *NAACL-HLT*, New Orleans, Louisiana, June 1-6, 2018.

Majumder, Anirban, Pande, Abhay, Vonteru, Kondalarao, Gangwar, Abhishek, Maji, Subhadeep, Bhatia, Pankaj and Goyal, Pawan. Automated Assistance in E-commerce: An Approach based on Category-Sensitive Retrieval. *40th European Conference on Information Retrieval (ECIR)*, Grenoble, France (short paper), March 26-29, 2018.

Mullick, Ankan, Dastider, Surjodoy G, Maheshwari, Shivam, Sahoo, Srotaswini, Maity, Suman Kalyan, C, Soumya and Goyal, Pawan. Identifying Opinion and Fact Subcategories from the Social Web. *ACM International Conference on Supporting Group Work (Group)*, Florida, USA (poster), January 7-10, 2018.

Jatin Arora, Sumi Agrawal, Pawan Goyal and Sayan Pathak (2017). Extracting Entities of Interest from Comparative Product Reviews. *CIKM*, Singapore, November 6-10, 2017.

Mayank Singh, Rajdeep Sarkar, Pawan Goyal, Animesh Mukherjee and Soumen Chakrabarti. Relay-Linking Models for Prominence and Obsolescence in Evolving Networks. *KDD*, Halifax, Canada, August 10-13, 2017.

Ankan Mullick, Pawan Goyal, Niloy Ganguly and Manish Gupta. Extracting Social Lists from Twitter. *ASONAM*, Sydney, Australia (short paper), July 31-August 3, 2017.

Mayank Singh, Ajay Jaiswal, Priya Shree, Arindam Pal, Animesh Mukherjee and Pawan Goyal. Understanding the Impact of Early Citers on Long-Term Scientific Impact. *ACM/IEEE Joint Conference on Digital Libraries (JCDL)*, Toronto, Ontario, Canada, June 19-23, 2017.

Abhik Jana, Sruthi Mooriyath, Animesh Mukherjee and Pawan Goyal (2017). WikiM: Metapaths based Wikification of Scientific Abstracts. *ACM/IEEE Joint Conference on Digital Libraries (JCDL)*, Toronto, Ontario, Canada, June 19-23, 2017.

Mayank Singh, Abhishek Niranjan, Divyansh Gupta, Nikhil Angad Bakshi, Animesh Mukherjee and Pawan Goyal (2017). Citation sentence reuse behavior of scientists: A case study on massive bibliographic text dataset of computer science. *ACM/IEEE Joint Conference on Digital Libraries (JCDL) (short paper)*, Toronto, Ontario, Canada, June 19-23, 2017.

Ankan Mullick, Shivam Maheshwari, Soumya C, Pawan Goyal and Niloy Ganguly. A Generic Opinion-Fact Classifier with Application in Understanding Opinionatedness in Various News Sections. *26th World Wide Web (WWW) (Poster)*, Perth, Australia, April 3rd - 7th, 2017.

Amrith Krishna, Bishal Santra, Pavan Kumar Satuluri, Sasi Prasanth Bandaru, Bhumi Faldu, Yajuvendra Singh and Pawan Goyal. Word Segmentation in Sanskrit Using Path Constrained Random Walk. *26th International Conference on Computational Linguistics (Coling) (Poster)*, Osaka, Japan, December 11-16, 2016.

Mayank Singh, Barnopriyo Barua, Priyank Palod, Manvi Garg, Sidhartha Satapathy, Samuel Bushi, Kumar Ayush, Krishna Sai Rohith, Tulasi Gamidi, Pawan Goyal and Animesh Mukherje. OCR++: A Robust Framework For Information Extraction from Scholarly Articles. *26th International Conference on Computational Linguistics (Coling) (Poster)*, Osaka, Japan, December 11-16, 2016 .

Abhishek Sikchi, Pawan Goyal and Samik Datta. PEQ : An explainable, specification-based, aspect-oriented product comparator for e-commerce. *25th ACM Conference on Information and Knowledge Management (CIKM)*, Indianapolis, October 24-28, 2016.

Koustav Rudra, Siddhartha Banerjee, Niloy Ganguly, Pawan Goyal, Mohammad Imran and Prasenjit Mitra. Summarizing Situational Tweets in Crisis Scenario. *27th ACM Conference on Hypertext and Social Media*, Halifax, Canada, July 10-13, 2017.

Tanmoy Chakraborty, Amrith Krishna, Mayank Singh, Niloy Ganguly, Pawan Goyal and Animesh Mukherjee. FeRoSA: A Faceted Recommendation System for Scientific Articles. *20th Pacific Asia Conference on Knowledge Discovery and Data Mining (PAKDD)*, Auckland, New Zealand, April 19-22, 2016.

Mayank Singh, Vikas Patidar, Suhansanu Kumar, Tanmoy Chakraborty, Animesh Mukherjee and Pawan Goyal. The role of citation context in predicting long-term citation profiles: an experimental study based on a massive bibliographic text dataset. *24th ACM Conference on Information and Knowledge Management (CIKM)*, Melbourne, Australia, October 19-23, 2015

Koustav Rudra, Shubham Ghosh, Niloy Ganguly, Pawan Goyal and Saptarshi Ghosh. Extracting Situational Information from Microblogs during Disaster Events: A Classification-Summarization Approach. *24th ACM Conference on Information and Knowledge Management (CIKM)*, Melbourne, Australia, October 19-23, 2015.

Tanmoy Chakraborty, Sikhar Patranabis, Pawan Goyal and Animesh Mukherjee. On the formation of circles in co-authorship networks. *21st ACM SIGKDD*, Sydney, Australia, August 10-13, 2015.

Mayank Singh, Tanmoy Chakraborty, Animesh Mukherjee and Pawan Goyal. ConfAssist: A Conflict resolution framework for assisting the categorization of Computer Science conferences. *JCDL*, Knoxville, Tennessee, USA, June 21-25, 2015 (Poster).

Suman Kalyan Maity, Abhishek Gupta, Pawan Goyal and Animesh Mukherjee. A stratified learning approach for predicting the popularity of Twitter Idioms. *9th International AAAI Conference on Web and Social Media*, Oxford, UK, May 26-29, 2015 (Poster).

Tanmoy Chakraborty, Suhansanu Kumar, Pawan Goyal, Niloy Ganguly and Animesh Mukherjee. Towards a Stratified Learning Approach to Predict Future Citation Counts. *In ACM/IEEE Joint Conference on Digital Libraries (JCDL)*, London, UK, September 8-12, 2014, pp. 351–360.

Pawan Goyal and Amba Kulkarni. Converting Phrase Structures to Dependency Structures in Sanskrit. *In proceedings of 25th International Conference on Computational Linguistics (COLING)*, Dublin, Ireland, August 25-29, 2014, pp. 1834–1843.

Sunny Mitra, Ritwik Mitra, Martin Riedl, Chris Biemann, Animesh Mukherjee and Pawan Goyal. That's sick dude!: Automatic identification of word sense change across different timescales. *In the proceedings of ACL 2014*, Baltimore, USA, June 22-27, 2014.

Pawan Goyal, Gérard Huet, Amba Kulkarni, Peter Scharf and Ralph Bunker. A Distributed Platform for Sanskrit Processing. *In proceedings of Coling 2012*, IIT Bombay, INDIA, December 2012, pp 1011-1028.

OTHER
CONFERENCE,
WORKSHOP
PUBLICATIONS
AND BOOK
CHAPTERS

Krishna, Amrith, Vidhyut, Shiv, Chawla, Dilpreet, Sambhavi, Sruti and Goyal, Pawan. SHR++: An Interface for Morpho-syntactic annotation of Sanskrit Corpora, *LREC*, May 11-16, 2020, Marseille, France.

Jana, Abhik, Varimala, Nikhil and Goyal, Pawan. Using Distributional Thesaurus Embedding for Co-hyponymy Detection, *LREC*, May 11-16, 2020, Marseille, France.

Mathew, Binny, Kumar, Navish, Goyal, Pawan and Mukherjee, Animesh. Interaction dynamics between hate and counter users on Twitter, *CODS-COMAD*, January 5-7, 2020, Hyderabad, India.

Mathew, Binny, Maity, Suman Kalyan, Goyal, Pawan and Mukherjee, Animesh. Competing Topic Naming Conventions in Quora: Predicting Appropriate Topic Merges and Winning Topics from Millions of Topic Pairs, *CODS-COMAD*, January 5-7, 2020, Hyderabad, India.

Kayal, Pratik, Singh, Mayank and Goyal, Pawan. Weakly-Supervised Deep Learning for Domain Invariant Sentiment Classification, *CODS-COMAD (short)*, January 5-7, 2020, Hyderabad, India.

Hazra, Rima, Singh, Mayank, Goyal, Pawan, Adhikari, Bibhas and Mukherjee, Animesh. The rise and rise of interdisciplinary research: Understanding the interaction dynamics of three major fields – Physics, Mathematics & Computer Science. ICADL (short), November 4-7, 2019, Kualalumpur.

Sandhan, Jivnesh, Krishna, Amrith, Goyal, Pawan and Behera, Laxmidhar. Revisiting the Role of Feature Engineering for Compound Type Identification in Sanskrit, *6th ISCLS*, October 23-25, 2019, WB, India.

Mathew, Binny, Dutt, Ritam, Maity, Suman Kalyan, Goyal, Pawan and Mukherjee, Animesh. Deep Dive into Anonymity: Large Scale Analysis of Quora Questions, *SocInfo*, November 18-21, 2019, Doha, Qatar .

Patro, Jasabanta, Baruah, Sabyasachee, Gupta, Vivek, Choudhury, Monojit, Goyal, Pawan and Mukherjee, Animesh. Characterizing the spread of exaggerated health news content over social media, *ACM HyperText (poster)*, September 17 - 20, 2019, Hof, Germany.

Jana, Abhik and Goyal, Pawan. Network Features Based Co-hyponymy Detection. *11th International Conference on Language Resources and Evaluation (LREC)*, Miyazaki, Japan, May 7-12, 2018.

Reddy, Vikas, Krishna, Amrith, Sharma, Vishnu Dutt, Gupta, Prateek, R, Vineeth M and Goyal, Pawan. Building a Word Segmenter for Sanskrit Overnight. *11th International Conference on Language Resources and Evaluation (LREC)*, Miyazaki, Japan, May 7-12, 2018.

Krishna, Amrith, Majumder, Bodhisattwa Prasad and Goyal, Pawan. An 'Ekalavya' Approach to Learning Context Free Grammar Rules for Sanskrit Using Adaptor Grammar. *17th World Sanskrit Conference (WSC), section on Computational Sanskrit & Digital Humanities*, Vancouver, Canada, July 9-13, 2018.

Amrith Krishna, Pavankumar Satuluri, Harshavardhan Ponnada, Muneeb Ahmed, Gulab Arora, Kaustubh Hiware and Pawan Goyal (2017). A Graph Based Semi-Supervised Approach for Analysis of Derivational Nouns in Sanskrit. TextGraphs, workshop at ACL 2017, Vancouver, Canada, July 30-August 4, 2017.

Binny Mathew, Suman Kalyan Maity, Pratip Sarkar, Animesh Mukherjee and Goyal, Pawan (2017). Adapting predominant and novel sense discovery algorithms for identifying corpus-specific sense differences. TextGraphs, workshop at ACL 2017, Vancouver, Canada, July 30-August 4, 2017.

Amrith Krishna, Pavankumar Satuluri and Pawan Goyal (2017). A Dataset for Sanskrit Word Segmentation. LaTeCH-CLfL, workshop at ACL 2017, Vancouver, Canada, July 30-August 4, 2017.

Mayank Singh, Soham Dan, Sanyam Agarwal, Pawan Goyal and Animesh Mukherjee (2017). AppTechMiner: Mining Applications and Techniques from Scientific Articles. eC6th International Workshop on Mining Scientific Publications, workshop at JCDL 2017, Toronto, Ontario, Canada, June 19-23, 2017.

Krishnendu Ghosh, Plaban Bhowmick and Pawan Goyal (2017). Using Re-ranking to Boost Deep Learning based Community Question Retrieval. Web Intelligence (WI), Leipzig, Germany, August 23-26, 2017.

Kalpana Raja, Naman Dasot, Pawan Goyal and Siddhartha R Jonnalagadda. Towards Evidence-Based Precision Medicine: Extracting Population Information from Biomedical Text using Binary Classifiers and Syntactic Patterns. *AMIA 2016 Joint Summits on Translational Science*, San Francisco, USA, March 21-24, 2016.

Peter Scharf, Pawan Goyal, Anuja Ajotikar and Tanuja Ajotikar. Voice, preverb, and transitivity restrictions in Sanskrit verb use. *Sanskrit Syntax*, 2015, pp. 157–202.

Peter Scharf, Anuja Ajotikar, Sampada Savardekar and Pawan Goyal. Distinctive features of poetic syntax: preliminary results. *Sanskrit Syntax*, 2015, pp. 305–324.

Keshav Melnad, Pawan Goyal and Peter Scharf. Meter identification of Sanskrit verse. *Sanskrit Syntax*, 2015, pp. 325–346.

Amrith Krishna and Pawan Goyal. Towards automating the generation of derivative nouns in Sanskrit by simulating Panini. *16th World Sanskrit conference, Sanskrit and the IT world*, Bangkok, Thailand, June 28-July 02, 2015.

Pujari Rajkumar, Swara Desai, Niloy Ganguly and Pawan Goyal. A Novel Two-stage Framework for Extracting Opinionated Sentences from News Articles. *Textgraphs-9, workshop at EMNLP 2014*. Doha, Qatar, September 25-29, 2014, pp. 25–33.

Uwe Quasthoff, Ritwik Mitra, Sunny Mitra, Thomas Eckart, Dirk Goldhahn, Pawan Goyal and Animesh Mukherjee. Large Web Corpora of High Quality for Indian Languages. *In the proceedings of LREC WILDRE 2014*, Reykjavik, Iceland, May 26-31, 2014 (poster).

Gérard Huet and Pawan Goyal. Design of a Lean Interface for Sanskrit Corpus Annotation. *In the proceedings of ICON 2013*, CDAC Noida, India, December 18-20, 2013.

Pawan Goyal and Gérard Huet. Completeness Analysis of a Sanskrit Reader. *In Proceedings of Fifth Sanskrit Computational Linugistics Symposium*, IIT Bombay, INDIA, 2013.

Pawan Goyal, Laxmidhar Behera and T. M. McGinnity. An Information Retrieval Approach Based on Semantically Adapted Vector Space Model. *Proceedings of the 2009 International Conference on Natural Language Processing*, Hyderabad, INDIA, December 14-17, 2009.

Pawan Goyal, Laxmidhar Behera and T. M. McGinnity. Entailment of Causal Queries in Narratives Using Action Language. *Proceedings of the International Conference on Knowledge Discovery and Information Retrieval*, October 04-06, Funchal, Portugal. pp 112–118, 2009.

Pawan Goyal, Laxmidhar Behera and T. M. McGinnity. An Information Retrieval Model Based On Automatically Learnt Concept Hierarchies. *Proceedings of the IEEE International Conference on Semantic Computing*, Berkeley, CA, USA. pp 458-465, 2009.

Pawan Goyal, Vipul Arora and Laxmidhar Behera. Analysis of Sanskrit Text: Parsing and Semantic Relations. *Sanskrit Computational Linguistics: Revised Selected and Invited Papers*, Springer-Verlag, pp 200 - 218, 2009.

Pawan Goyal, Amba Kulkarni and Laxmidhar Behera. Computer Simulation of Ashtadhyayi: Some insights. *Sanskrit Computational Linguistics: Revised Selected and Invited Papers*, Springer-Verlag, pp 139 - 160, 2009.

Pawan Goyal and R. Mahesh K. Sinha. A Study towards Design of an English to Sanskrit Machine Translation System. *Sanskrit Computational Linguistics: Revised Selected and Invited Papers*, Springer-Verlag, pp 287 - 305, 2009.

Pawan Goyal and R. Mahesh K. Sinha. Translation Divergence Between English-Sanskrit-Hindi Language Pairs. *Proceedings of Third Sanskrit Computational Linguistic Symposium*, Springer-Verlag, pp 134-143, 2009.

Invited Talks

What's in a dialog? Inductive-bias motivated design principles for better dialog modeling, invited talk at Microsoft India, October 19th, 2021.

HIER: Hierarchical Transformer for Task-oriented Dialog Systems, invited talk at Goldman Sachs, September 23rd, 2021.

Sequence Learning using Recurrent Neural Networks and LSTMs, invited talk at Jadavpur University at Workshop on Machine Learning and Data Analytics, March 27th, 2019.

Sanskrit Computational Linguistics: Learning linguistic structures for Free word order Languages, invited talk at Manipal University, Jaipur, January 18th, 2019.

Attention Mechanism and Memory Networks, Invited talk at ISI Kolkata, September 26th, 2018.

Activity De-interleaving in a Multi-user Smarthome, Invited talk at Industry Research Network event, Samsung Bangalore, September 7th, 2018.

Introduction to Machine Learning, Invited talk at Calcutta University, March 16th, 2018.

Semantic Processing in Sanskrit, Invited talk at IIT Bombay, August 19th, 2016.

FeRoSA: A Faceted Recommendation System for Scientific Articles, Invited talk at TU Darmstadt, June 17, 2014.

A Context Based Word Indexing Model for Document Summarization, Invited talk at the Department of Electrical Engineering, IIT Kanpur, March 2013.

Orienting methods used in compositional and distributional semantics to languages rich in morphological and agglutinative structures, Invited talk in the Chair Blaise Pascal lecture series, CNRS Paris 7, France, May 2012.

Analytic Knowledge Discovery Models for Information Retrieval and Text Summarization, Invited talk at the Junior Research Seminar, INRIA Paris Rocquencourt, March 2012.