

Report on the 3rd Symposium on NLP for Social Good (NSG 2025)

Procheta Sen

University of Liverpool
UK

`procheta.sen@liverpool.ac.uk`

Tulika Saha

International Institute of Information Technology Bangalore
India

`tulika.saha@iiitb.ac.in`

Danushka Bollegala

University of Liverpool
UK

`danushka.bollegala@liverpool.ac.uk`

Abstract

Artificial intelligence (AI), specifically, Natural Language Processing (NLP) is being hailed as a new breeding ground for immense innovation potential. Researchers believe that NLP-based technologies could help to solve societal issues such as equality and inclusion, education, health, and hunger, climate action etc., and many more. Tackling these questions requires a concerted, collaborative effort across all sectors of society. The 3rd Symposium on NLP for Social Good (NSG) was a novel effort that aimed to enable NLP researchers and scholars from inter-disciplinary field who want to think about the societal implications of their work for solving humanitarian and environmental challenges. The objective of the symposium was to support fundamental research and engineering efforts and empower the social sector with tools and resources, while collaborating with partners from all sectors to maximise effect in solving problems within public health, nature & society, accessibility, crisis response etc. In its 3rd year, we invited speakers from academia and industry to provide an overview of some areas from NLP applications such as healthcare, climate and legal domains in order to provide a platform to stimulate discussion regarding the current state of NLP in these varied fields.

Date: 25–26 June, 2025.

Website: <https://nlp4social.github.io/NSG/>.

1 Introduction and Motivation for NSG

Artificial intelligence (AI), specifically, Natural Language Processing (NLP) is being hailed as a new breeding ground for immense innovation potential. While scholars believe that NLP has enormous potential for excessive growth, one question remains: how can it be used for the better welfare of the society? Researchers believe that NLP-based technologies could help to solve societal issues such as equality and inclusion, education, health, and hunger, and climate action etc.

and many more. The field is focused on delivering positive social impact in accordance with the priorities outlined in the United Nations’ 17 Sustainable Development Goals (SDGs). Tackling these questions requires a concerted, collaborative effort across all sectors of society. The Symposium on NLP for Social Good is a novel effort that aims to enable NLP researchers and scholars from interdisciplinary field who want to think about the societal implications of their work for solving humanitarian and environmental challenges. The symposium aims to support fundamental research and engineering efforts and empower the social sector with tools and resources, while collaborating with partners from all sectors to maximise effect in solving problems within public health, nature & society, climate & energy, accessibility, crisis response etc.

2 NSG 2025

NSG-2025 (website link ¹) was held as a virtual event from 25th to 26th June, 2025. The symposium was organized by NLP academicians from University of Liverpool, United Kingdom & International Institute of Information Technology Bangalore, India. The event hosted three keynote speakers and four invited talks from academia to share their research works related to NSG and their insights in regards to the potential of NSG. There were 130 registered participants for NSG 2025 from across the globe out of which 61.50% were “males”, 38.48% were “females” and 0.02% were from “others”. In addition to the participation from United Kingdom & India, there were participants from various other countries. The next section describes in detail the keynote and invited talks at NSG 2025.

2.1 Invited Talks at NSG 2025

Here, we enlist the three keynote and four invited talks delivered in NSG 2025. We are grateful to all the speakers for their insightful talks and the attendees for the meaningful discussion in the Q&A sessions. The video recordings of the invited talks are available on the YouTube channel².

2.1.1 Keynote Talk #1 by Prof. Tanmoy Chakraborty

Speaker Bio: The first keynote speaker, Prof. Tanmoy Chakraborty is a Rajiv Khemani Young Faculty Chair Professor in AI and an Associate Professor in the Dept. of Electrical Engineering and the School of AI at IIT Delhi. He leads the Laboratory for Computational Social Systems (LCS2), a research group that primarily focuses on building economical, adaptable and interpretable language models and applying them specifically to two areas – mental health counselling and cyber-informatics. He served as the DAAD visiting professor at MPI Saarbrücken, PECFAR visiting professor at TU Munich and ELISE visiting professor at TU Darmstadt. Tanmoy has received numerous recognitions, including the Humboldt Fellowship, Ramanujan Fellowship, ACL ’23 Outstanding Paper Award, IJCAI’23 AI for Good Award, and several faculty awards from industries like Google, LinkedIn, JP Morgan, and Adobe. He has authored two textbooks – ”Social Network Analysis” and ”Introduction to Large Language Models”. Tanmoy earned his PhD from IIT Kharagpur in 2015 as a Google PhD Scholar. He is currently serving as the PC Chair

¹<https://nlp4social.github.io/NSG/>

²<https://www.youtube.com/@NLPforSocialGood>

of EMNLP'25 and local chair of ACL'25. More details may be found at tanmoychak.com.

Title: Towards Enhanced Conversational Dynamics for Effective Virtual Therapist-Assistive Counseling

Abstract: The increasing demand for digital healthcare, coupled with current infrastructure limitations, requires digital therapeutic interventions. My talk will focus on the design and implementation of Virtual Mental Health Assistants modules that serve as therapist-assistive mechanisms to automate their complex work cycle. We work on building novel LLM-based methods for dialogue understanding, summarization, and generation, and our research captures the intricacies of therapeutic communication while incorporating signs into human behavior analysis. In support of this, we also develop datasets and resources, many of which are first-of-its-kind, including HOPE, MEMO, MENTAL-TRUST, MentalCLOUDS, and BeCOPE, all of which are available for research purposes.

2.1.2 Keynote Talk #2 by Dr. Tommaso Caselli

Speaker Bio: The second keynote speaker, Dr. Tommaso Caselli is Assistant Professor in Computational Semantics at the Center for Language and Cognition (CLCG) of the Faculty of Arts of the University of Groningen. His main research interests are in event extraction and framing, hate speech and misinformation detection and countering. He is one of the founders of the "Event and Stories in the News" workshop series and the co-editor of the volume "Computational Analysis of Storylines" (CUP, 2021). He has been involved in the organization of several semantic evaluation campaigns for NLP for English and Italian. He has covered (senior) area chair positions (COLING, ACL, EMNLP, EACL) and his work is featured in major *CL conferences and journals. He has been awarded two "Outstanding Paper Award" (COLING 2022; ACL 2023) and one "Best Paper Award" (ACL 2022). Since November 2024, he is the coordinator of the theme "AI and Language" of the Jantina Tammes School of Digital Society, Technology, and AI of the University of Groningen.

Title: Framing Perspectives on Environmental Sustainability

Abstract: Communication is at the core of every human activity. The way we speak, or narrate something, activates (consciously or unconsciously) perspectives on things that happen in the world. These perspectives are not simple points of view but they encode and influence our perception of events and phenomena. A ubiquitous device to encode and convey such perspectives is framing. The difference between "climate change" and "climate crisis" is primarily a difference in frames that these words activate in the minds of receivers: a "change" is more neutral and less urgent than a "crisis". In this talk, I will present and discuss ongoing research on frame activation and generation at the lexical level concerning the food transition and parliamentary debates on climate change in the European Union.

2.1.3 Keynote Talk #3 by Dr. Aparna Taneja

Speaker Bio: The third speaker, Dr. Aparna Taneja is a researcher at the Multi Agents Systems for Social Impact team in Google Deepmind India. She collaborates with several NGO's and

academic partners in the fields of public health and conservation, and her primary focus is the collaboration with ARMMAN, an NGO focused on improving maternal and child health outcomes in India. She was named on the TIME 100 most influential people in Health globally in 2025, for her work with ARMMAN. Aparna joined Google in 2015 and worked with the Google Maps team in Zurich to improve search quality for Maps. She received her Phd in Computer Science at ETH Zurich under the supervision of Prof. Marc Pollefeys. She then pursued a postdoc at Disney Research Zurich. She received her Bachelors and Masters in Computer Science from IIT Delhi in 2006.

Title: Using AI to assist in improving maternal and child health outcomes in underserved communities in India

Abstract: The widespread availability of cell phones has enabled non-profits to deliver critical health information to their beneficiaries in a timely manner. However, significant fraction of beneficiaries drop out of the program and non-profits often have limited health-worker resources to place crucial service calls for live interaction with beneficiaries to prevent such engagement drops. To assist non-profits in optimizing this limited resource, we developed a Restless Multi-Armed Bandits (RMABs) system. The RMAB system was evaluated in collaboration with an NGO via a real-world service quality improvement study and showed a 30% reduction in engagement drops. This has inspired a lot of research from the team in the broad area of limited resource allocation using RMABs. More recently, we have presented efforts towards a foundation model for RMABs, additionally empowered by LLMs to offer more flexibility and adaptability to changing goals.

2.1.4 Invited Talk #1 by Dr. M R Hasan

Speaker Bio: Dr. M R Hasan is an assistant professor at the University of Nebraska-Lincoln, USA.

Title: Bridging Modalities, Improving Lives: How Multimodal AI Systems Can Enhance Educational Equity and Outcomes

Abstract: This talk explores the transformative potential of multimodal AI systems, integrating natural language processing and vision capabilities, to advance educational interventions and improve learning outcomes. At the Human-First Artificial Intelligence Lab (HAL 2.0), our research on modeling complex longitudinal experiential (LE) data, capturing students' cognitive, emotional, and behavioral dynamics over time, has highlighted significant challenges in achieving generalizable insights. Drawing from our NSF-supported research on the "Messages From A Future You" AI system, which initially explored methods like large language models for analyzing noisy, sparse, and heterogeneous student data collected throughout an academic semester, we encountered limitations in generalizing predictive models across student cohorts and contexts. To overcome these fundamental challenges inherent in LE data modeling, we developed a novel multimodal framework, leveraging vision-language models. By transforming LE data into complementary textual narratives and visual representations, our approach is specifically designed to capture intricate structural dynamics and overcome data limitations, enabling forecasting of learning outcomes and behavioral attributes with greater precision and robust generalizability. This

multimodal AI framework shows promising potential for delivering personalized interventions informed by the nuanced variations in students' learning experiences, thereby enhancing educational equity and outcomes, while establishing a foundational paradigm that can extend beyond education to healthcare, mental wellness, and other domains where understanding complex human experiences is essential for positive social impact.

2.1.5 Invited Talk #2 by Dr. Shivani Kumar

Speaker Bio: Dr. Shivani Kumar is a postdoctoral Fellow at the University of Michigan, USA

Title: Are Rules Meant to be Broken? Understanding Multilingual Moral Reasoning as a Computational Pipeline with UniMoral

Abstract: Moral reasoning is fundamental to human decision-making, influencing social interactions, policy-making, and ethical AI development. However, its computational study remains fragmented, with existing NLP research relying on disparate datasets and isolated tasks. To advance NLP for social good, we introduce UniMoral, a multilingual dataset designed to facilitate the development of AI systems that understand and navigate ethical dilemmas in diverse cultural settings. UniMoral integrates psychologically grounded and real-world moral dilemmas from social media, annotated with action choices, ethical principles, contributing factors, and consequences, alongside annotators' moral and cultural profiles. Recognizing the cultural relativity of moral reasoning, UniMoral spans six languages—Arabic, Chinese, English, Hindi, Russian, and Spanish—enabling cross-cultural analysis. We assess its impact through benchmark evaluations of three large language models (LLMs) across four tasks: action prediction, moral typology classification, factor attribution analysis, and consequence generation. Our findings highlight that while LLMs can leverage implicit moral contexts, significant challenges remain in ensuring these models reason ethically across diverse sociocultural landscapes. UniMoral lays the foundation for more equitable, context-aware AI systems, fostering NLP applications that promote fairness, inclusivity, and ethical awareness in automated decision-making.

2.1.6 Invited Talk #3 by Shubham Kumar Nigam

Speaker Bio: Shubham Kumar Nigam is a PhD student at the Indian Institute of Technology, Kanpur.

Title: NyayaSutra: Enabling Reliable and Interpretable Legal Judgment through Structured Thinking.

Abstract: In high-stakes domains like law, opaque AI models pose a significant barrier to real-world adoption. Legal professionals demand not just accurate predictions but interpretable reasoning paths that align with judicial logic. While explainability techniques have emerged to address this, they often provide post-hoc justifications rather than surfacing the actual reasoning that led to a decision, leading to a growing gap between model outputs and human trust. NyayaSutra introduces an interpretable and reliable AI framework for legal judgment prediction and reasoning, tailored to the Indian judiciary. It leverages a structured thinking paradigm, breaking down

judgments into rhetorical segments, Facts, Issues, Arguments, Reasoning, and Decision, to ensure transparency and traceability. The system employs hybrid legal retrieval, instruction-tuned LLMs trained on annotated Indian judgments, and GRPO-based optimization using structured “thinking tokens.” By making legal reasoning interpretable from the ground up, NyayaSutra empowers legal professionals, researchers, and policymakers with factual, explainable, and trustworthy AI outputs, contributing meaningfully to the larger vision of NLP for Social Good.

2.1.7 Invited Talk #4 by Dr. Vivek Kumar

Speaker Bio: Dr. Vivek Kumar is a senior researcher at the Chair of Open Source Intelligence in INF 7 Institute for Data Security at Germany.

Title: Towards Empathetic AI: Safe AI Practice in Behavioral Therapy.

Abstract: While Large Language Models (LLMs) have demonstrated remarkable capabilities in generating human-like text, their reliable application in low-resource domains such as mental health remains limited. This challenge primarily arises from the domain’s inherent complexity and the scarcity of high-quality, labeled data, which can contribute to bias, hallucinations, and a lack of emotional nuance. This work presents a mixed-methods approach to evaluate the efficacy of LLMs in psychotherapy, specifically in generating accurate summaries of Motivational Interviewing (MI) dialogues. The experimental design explores two summarization levels: (i) full-session and (ii) utterance-level MI summaries. The work evaluates 13 state-of-the-art (SOTA) large language models (LLMs) alongside classical natural language processing (NLP) methods for plausible in-context data generation. The work introduces a novel two-stage evaluation scheme grounded in the Motivational Interviewing Treatment Integrity (MITI) framework, assessing key components such as evocation, collaboration, autonomy, direction, empathy, and non-judgmental attitude. These criteria are used to evaluate semantic drift, hallucinations, MI adherence, fluency, and contextual coherence. Additionally, the study benchmarks LLMs on NLP downstream tasks, including automated annotation and multi-class, multi-label classification. The findings offer insights into the capacity of LLMs to model complex psychological constructs and provide best practices for mitigating semantic drift in therapeutic contexts. This study also contributes a publicly available dataset comprising 1,764 MI dialogues—both low- and high-quality—encompassing approximately 81,000 therapist-client talk turns.

2.2 Organizers of NSG-2025

- **Dr. Procheta Sen**³ is a Lecturer (Assistant Professor) in the Department of Computer Science at University of Liverpool (UoL), United Kingdom. She is a part of the NLP group at UoL. Before joining UoL, she worked with Emine Yilmaz as a postdoctoral researcher in the Web Intelligence Group at University College London (UCL). She finished her PhD. in 2021 from Dublin City University, Ireland. Her recent research interest lies in building intelligent systems fair and explainable. Her work on explaining information retrieval systems from an IR practitioner’s perspective and measuring trust of an information retrieval system

³<https://procheta.github.io/sprocheta/>

got published in SIGIR 2020 and CIKM 2022 respectively. She has given a tutorial on Explainability for IR and NLP models at FIRE 2022, SIGIR 2023. She was one of the organizers of the first symposium on NLP for Social Good at the University of Liverpool. Some of her publications are as follows :

1. Automated Argument Generation from Legal Facts [Tuvey and Sen, 2023]. arXiv preprint arXiv:2310.05680.
 2. Towards socially responsible ai: Cognitive bias-aware multi-objective learning [Sen and Ganguly, 2020]. In Proceedings of the AAAI Conference on Artificial Intelligence.
- **Dr. Tulika Saha**⁴ (corresponding author) is an Assistant Professor at the Dept. of DSAI & CSE in the International Institute of Information Technology Bangalore (IIIT-B), India. She also serves as an Honorary Lecturer at the School of CS & Informatics in the University of Liverpool (UoL), United Kingdom (UK) where she worked earlier as a full-time Lecturer prior to joining IIIT-B. Her current research interests include NLP typically Dialogue Systems, Social Good, Social Media Analysis using Deep Learning and Reinforcement Learning etc. She was a postdoctoral research fellow at the National Centre for Text Mining, University of Manchester, UK. Previously she earned her Ph.D. from Indian Institute of Technology Patna, India. Her research articles are published in top-tier conferences such as ACL, ACM SIGIR, EMNLP, NAACL, CoLing etc. and several peer-reviewed journals. She has organized several workshops/symposium namely NLP for Social Good 2023, 2024, AICAI 2023⁵, LLMIT 2023 [Saha et al., 2023a] and presented tutorials in ECIR [Saha et al., 2023b], InterSpeech⁶ 2023 etc. Some of her recent publications in the Social Good theme are as follows :
1. A shoulder to cry on: towards a motivational virtual assistant for assuaging mental agony [Saha et al., 2022b]. In Proceedings of the 2022 conference of the North American chapter of the association for computational linguistics: Human language technologies.
 2. Towards motivational and empathetic response generation in online mental health support [Saha et al., 2022a]. In Proceedings of the 45th international ACM SIGIR conference on research and development in information retrieval.
 3. Symptoms are known by their companies: towards association guided disease diagnosis assistant [Tiwari et al., 2022]. BMC bioinformatics, 23(1).
 4. Domain infused conversational response generation for tutoring based virtual agent [Jain et al., 2022]. In 2022 International Joint Conference on Neural Networks (IJCNN), IEEE.
 5. T-VAKS: A Tutoring-based Multimodal Dialog System via Knowledge Selection [Jain et al., 2023b]. In Proceedings of the 26th European Conference on Artificial Intelligence, 2023.
 6. Can you Summarize my learnings? Towards Multi-modal Perspective-based Educational Dialogue Summarization [Jain et al., 2023a]. In Findings of EMNLP, 2023.
 7. "Reasoning before Responding": Towards Legal Long-form Question Answering with Interpretability [Ujwal et al., 2024]. In Proceedings of the 33rd ACM International Conference on Information and Knowledge Management, 2024.

⁴<https://sahatulika15.github.io/>

⁵<https://aicai2023.site/>

⁶<https://interspeech2023.org/tutorials/>

-
8. HiLEx: Image-Based Hierarchical Layout Extraction from Question Papers [Aich et al., 2026]. In Proceedings of ICDAR 2025.
 9. Do Sentiment and Emotion Affect Mental Health? A Multi-task Classification Framework for Comprehensive Understanding of Mental Health, Emotion, and Sentiment from Motivational Conversations [Pathak et al., 2025]. In ACM Transactions on Computing for Healthcare, 2025.

• **Prof. Danushka Bollegala**⁷ is a Professor in the Department of Computer Science at the University of Liverpool and a Amazon Scholar with Amazon Search. He has published over 190 peer-reviewed papers related to NLP/ML and has worked as a Program Chair, Senior Area chair, Senior PCs at NLP/AI conferences. His research interests are lexical and compositional semantics, information extraction, domain adaptation and text summarization. He has received various awards such as the conference best paper awards (GECCO, PRICAI) and IEEE Young Author Award in the past for his research excellence. Some of his publications relevant to the theme are as follows :

1. Unmasking the Mask – Evaluating Social Biases in Masked Language Models [Kaneko and Bollegala, 2022], Proc. of the 36th AAAI Conference on Artificial Intelligence, 2022.
2. Sense Embeddings are also Biased – Evaluating Social Biases in Static and Contextualised Sense Embeddings [Zhou et al., 2022], Proc. of the 60th Annual Meeting of the Association for Computational Linguistics, 2022.
3. Debiasing isn't enough! – On the Effectiveness of Debiasing MLMs and their Social Biases in Downstream Tasks [Kaneko et al., 2022a], Proc. of the 29th International Conference on Computational Linguistics (COLING 2022).
4. Gender Bias in Meta-Embedding [Kaneko et al., 2022b], Proc. of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP 2022).
5. The Impact of Debiasing on the Performance of Language Models in Downstream Tasks is Underestimated [Kaneko et al., 2023], Proceedings of the 13th IJCNLP and the 3rd Conference of the ACL, 2023.

3 Concluding Remarks

In NSG 2025, we discussed about the emerging topics of NLP which has relevance to social good such as Healthcare, Legal-NLP and Education-NLP. The discussions were broadly categorised into two areas. One was focused on the challenges and accessibility of the datasets to pursue research in these areas. The other one was about the application of the recent and advanced NLP and Machine Learning techniques to solve the different challenges in these areas. The discussions with the invited speakers were useful for both students and senior researchers attending NSG.

One of the core contributions of this Symposium is its ability to foster interdisciplinary engagement. NSG transcends traditional academic boundaries, attracting researchers from various

⁷<https://danushka.net/>

fields such as computer science, linguistics, social sciences, public policy, and more. This diversity of perspectives enriches the academic discourse and promotes innovative approaches to solving real-world problems and creates a thriving ecosystem of knowledge exchange and research partnerships. Through this Symposium on NSG, we emphasize the importance of addressing global challenges, such as healthcare, education, environmental sustainability, and social justice etc. This not only widens the scope of NLP research but also encourages researchers to tackle critical issues, resulting in research that has a meaningful impact on society.

The video recordings of NSG-2025 is available on the Symposium on NLP for Social Good (NSG) YouTube channel⁸. We plan to organise future editions of NSG, incorporating all the feedback and suggestions obtained from NSG-2024.

Acknowledgements

The organizers express their gratitude to all the keynote/invited speakers and the attendees for their invaluable contributions to the symposium. The organizers would also like to extend heartfelt thanks to the University of Liverpool, UK & IIIT Bangalore, India for their valuable support to organize NSG-2025.

References

- Utathya Aich, Shinjini Chakraborty, Deepan Sadhukhan, Swarnendu Ghosh, and Tulika Saha. Hilex: Image-based hierarchical layout extraction from question papers. In Xu-Cheng Yin, Dimosthenis Karatzas, and Daniel Lopresti, editors, *Document Analysis and Recognition – ICDAR 2025*, pages 485–505, Cham, 2026. Springer Nature Switzerland. ISBN 978-3-032-04627-7.
- Raghav Jain, Tulika Saha, Souhitya Chakraborty, and Sriparna Saha. Domain infused conversational response generation for tutoring based virtual agent. In *2022 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 2022.
- Raghav Jain, Tulika Saha, Jhagrut Lalwani, and Sriparna Saha. Can you summarize my learnings? towards perspective-based educational dialogue summarization. In Houda Bouamor, Juan Pino, and Kalika Bali, editors, *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 3158–3173, Singapore, December 2023a. Association for Computational Linguistics. doi: 10.18653/v1/2023.findings-emnlp.208. URL <https://aclanthology.org/2023.findings-emnlp.208/>.
- Raghav Jain, Tulika Saha, and Sriparna Saha. T-vaks: A tutoring-based multimodal dialog system via knowledge selection. 2023b.
- Masahiro Kaneko and Danushka Bollegala. Unmasking the mask—evaluating social biases in masked language models. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 11954–11962, 2022.
- Masahiro Kaneko, Danushka Bollegala, and Naoaki Okazaki. Debiasing isn’t enough!—on the effectiveness of debiasing mlms and their social biases in downstream tasks. *arXiv preprint arXiv:2210.02938*, 2022a.

⁸<https://www.youtube.com/@NLPforSocialGood>

-
- Masahiro Kaneko, Danushka Bollegala, and Naoaki Okazaki. Gender bias in meta-embeddings. *arXiv preprint arXiv:2205.09867*, 2022b.
- Masahiro Kaneko, Danushka Bollegala, and Naoaki Okazaki. The impact of debiasing on the performance of language models in downstream tasks is underestimated. *arXiv preprint arXiv:2309.09092*, 2023.
- Agnibh Pathak, Soham Bhattacharjee, Tulika Saha, and Sriparna Saha. Do sentiment and emotion affect mental health? a multi-task classification framework for comprehensive understanding of mental health, emotion, and sentiment from motivational conversations. *ACM Trans. Comput. Healthcare*, 6(2), April 2025. doi: 10.1145/3704740. URL <https://doi.org/10.1145/3704740>.
- Tulika Saha, Vaibhav Gakhreja, Anindya Sundar Das, Souhitya Chakraborty, and Sriparna Saha. Towards motivational and empathetic response generation in online mental health support. In *Proceedings of the 45th international ACM SIGIR conference on research and development in information retrieval*, pages 2650–2656, 2022a.
- Tulika Saha, Saichethan Reddy, Anindya Das, Sriparna Saha, and Pushpak Bhattacharyya. A shoulder to cry on: towards a motivational virtual assistant for assuaging mental agony. In *Proceedings of the 2022 conference of the North American chapter of the association for computational linguistics: Human language technologies*, pages 2436–2449, 2022b.
- Tulika Saha, Debasis Ganguly, Sriparna Saha, and Prasenjit Mitra. Large language models’ interpretability and trustworthiness (llmit). 2023a.
- Tulika Saha, Abhishek Tiwari, and Sriparna Saha. Trends and overview: The potential of conversational agents in digital health. In *European Conference on Information Retrieval*, pages 349–356. Springer, 2023b.
- Procheta Sen and Debasis Ganguly. Towards socially responsible ai: Cognitive bias-aware multi-objective learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(03):2685–2692, Apr. 2020.
- Abhishek Tiwari, Tulika Saha, Sriparna Saha, Pushpak Bhattacharyya, Shemim Begum, Minakshi Dhar, and Sarbajeet Tiwari. Symptoms are known by their companies: towards association guided disease diagnosis assistant. *BMC bioinformatics*, 23(1):1–23, 2022.
- Oscar Tuvey and Procheta Sen. Automated argument generation from legal facts, 2023.
- Utkarsh Ujwal, Sai Sri Harsha Surampudi, Sayantan Mitra, and Tulika Saha. ”reasoning before responding”: Towards legal long-form question answering with interpretability. In *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management, CIKM ’24*, page 4922–4930, New York, NY, USA, 2024. Association for Computing Machinery. ISBN 9798400704369. doi: 10.1145/3627673.3680082. URL <https://doi.org/10.1145/3627673.3680082>.
- Yi Zhou, Masahiro Kaneko, and Danushka Bollegala. Sense embeddings are also biased—evaluating social biases in static and contextualised sense embeddings. *arXiv preprint arXiv:2203.07523*, 2022.