

Report on the 2nd Symposium on NLP for Social Good (NSG 2024)

Procheta Sen¹, Tulika Saha¹ and Danushka Bollegala¹

¹Department of Computer Science, University of Liverpool, United Kingdom

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 first 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 inception, we invited speakers from academia and industry to provide an overview of some areas from NLP applications such as education, healthcare and legal domains in order to provide a platform to stimulate discussion regarding the current state of NLP in these varied fields.

Keywords

AI for Social Good, Natural Language Processing, Climate Change, Legal NLP

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.

2nd Symposium on NLP for Social Good 2024, April 25–26, 2024, Liverpool, UK



© 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).



CEUR Workshop Proceedings (CEUR-WS.org)

2. NSG 2024

NSG-2024 (website link ¹) was held as a hybrid event from 25th to 26th June, 2024 in the University of Liverpool (UoL), United Kingdom. The symposium was organized by NLP academicians from the Department of Computer Science, UoL. The event hosted three keynote speakers and one invited talks from academia to share their research works related to NSG and their insights in regards to the potential of NSG. There were 177 registered participants for NSG 2024 from across the globe. The next section describes in detail the keynote and invited talks at NSG 2023.

3. Invited Talks at NSG 2024

Here, we enlist the three keynote/invited talks delivered in NSG 2024. We are grateful to all the speakers for their insightful talks and the attendees for the meaningful discussion in the Q&A sessions.

3.1. Keynote Talk #1 by Prof. Manfred Stede

Speaker Bio: Manfred Stede is a Professor of applied computational linguistics at Potsdam University, Germany. His research and teaching activities revolve around issues in discourse structure and automatic discourse parsing, including applications in sentiment analysis and argument mining. For several years now, he actively collaborates with social scientists from different disciplines (political science, education science, communication science) on research questions involving political argumentation, social media analysis, and a focus on discourses about climate change. Stede is a (co-) author of four books, 30 journal papers, and 150 conference or workshop papers and book chapters.

Title: NLP on Climate Change Discourse: Two Case Studies

Abstract: The debate around climate change (CC) – its extent, its causes, and the necessary responses – is intense and of global importance. The ongoing discourses are a prominent object of study in several Social Sciences, while in the natural language processing community, this domain has so far received relatively little attention. In my talk, I first give a brief overview of types of approaches and data, and then report on two case studies that we are currently conducting in my research group. The first tackles the notion of "framing" (the perspective taken in viewing an issue) in CC-related editorials of the journals 'Nature' and 'Science': We proceed from a coarse-grained text-level labeling to increasingly detailed clause-level annotation of framing CC, and run experiments on automatic classification. The second involves a corpus of parliamentary speeches, press releases and tweets from the members of the German parliament (2017-2021) and compares their ways of addressing CC, contrasting on the one hand the different communication channels and on the other hand the party affiliations of the speakers.

¹<https://nlp4social.github.io/nlp4socialgood/>

3.2. Keynote Talk #2 by Prof. Sophia Ananiadou

Speaker Bio: Sophia Ananiadou is Professor in the Department of Computer Science, University of Manchester, UK; Director, National Centre for Text Mining; Deputy Director, Institute of Data Science and Artificial Intelligence; Distinguished Research Fellow, Artificial Intelligence Research Centre, AIST, Japan and senior researcher at the Archimedes Research Centre, Greece. She was an Alan Turing Institute Fellow (2018-2023). Currently she is a member of the European Laboratory for Learning and Intelligent Systems (ELLIS). Her research contributions in Natural Language Processing include tasks such as information extraction, summarisation, simplification, emotion detection and misinformation analysis. Her research is deeply interdisciplinary. She has been active bridging the gap between Biomedicine and NLP, via the provision of tools for a variety of translational applications related with personalized medicine, drug discovery, database curation, risk assessment, disease prediction. Her research on recognising affective information (emotions, sentiments) has been applied to mental health applications and misinformation detection. Currently, she is also focusing on the development of LLMs for FinTech applications.

Title: Emotion Detection and LLMs: Transforming Mental Health and Countering Misinformation on Social Media

Abstract: Social media serves as a key resource for analysing mental health through natural language processing (NLP) techniques like emotion detection. While current efforts focus on specific aspects of affective classification, such as sentiment polarity, they overlook regression tasks like sentiment intensity. We recognize the importance of emotional cues in mental health detection and propose MentaLLaMA, an interpretable LLM series for social media mental health analysis. Emotions and sentiment also play vital roles in detecting misinformation and conspiracy theories. However, existing LLM-based approaches often neglect the emotional dimensions of misinformation. By integrating affective cues into automated detection systems, we can improve accuracy. We'll showcase an open-source LLM that leverages emotional cues for enhanced detection of conspiracy theories, utilizing a novel conspiracy dataset.

3.3. Keynote Talk #3 by Prof. Kevin D. Ashley

Speaker Bio: Kevin D. Ashley, Ph.D., is an expert on computer modeling of legal reasoning. He was selected as a Fellow of the American Association of Artificial Intelligence "for significant contributions in computationally modeling case-based and analogical reasoning in law and practical ethics." He has been a principal investigator of a number of National Science Foundation grants and is co-editor in chief of Artificial Intelligence and Law, the journal of record in the field. He wrote Modeling Legal Argument: Reasoning with Cases and Hypotheticals (MIT Press/Bradford Books, 1990) and Artificial Intelligence and Legal Analytics: New Tools for Law Practice in the Digital Age (Cambridge University Press, 2017). He is a full professor at the School of Law, a senior scientist at the Learning Research and Development Center, and a member of the Intelligent Systems Program of the University of Pittsburgh.

Title: Modeling Case-based Legal Argument with Text Analytics

Abstract: Researchers in AI and Law have applied text analytic tools, including Natural Language Processing and Machine Learning, to predict the outcomes of legal cases and to attempt to explain the predictions in terms legal professionals would understand. Such explanations require legal knowledge, but integrating legal knowledge with deep learning can be problematic. Formerly, in modeling how legal professionals argue with cases and analogies, researchers explicitly represented aspects of the legal knowledge that advocates and judges employ in predicting, explaining, and arguing for and against case outcomes, such as legal issues, rules, factors, and values. Representing cases in terms of the applicable knowledge, however, was a manual process. More recently, researchers are employing text analytics to bridge the gap automatically between case texts and their argument models. Advances in large language models and generative AI have expanded the approaches for automatically representing the knowledge but raise new questions about the roles, if any, that argument models and the associated legal knowledge will play in an age of generative AI. This talk surveys a series of recent projects that bear on these questions.

3.4. Keynote Talk #4 by Dr. Swarnendu Ghosh

Speaker Bio: Dr. Swarnendu Ghosh is an academician and researcher specializing in compact and innovative deep learning solutions for computer vision, generative AI, and NLP. With a PhD in Computer Science and Engineering from Jadavpur University, India, and extensive research experience as an Erasmus Mundus Fellow at the University of Evora, Portugal, Dr. Ghosh has honed his expertise in developing cutting-edge methodologies for object recognition, image segmentation, and knowledge graph generation across diverse domains. His academic journey includes a Master's degree from Jadavpur University, where he explored sentiment classification using discourse analysis and graph kernels. Swarnendu has also contributed significantly to multiple Government research projects, such as developing knowledge graphs from images and crafting event-guided natural scene description frameworks for real-time applications. His rapidly growing research profile has already gathered over 800 citations in a short duration and boasts eminent journals such as ACM Computing Surveys, Pattern Recognition and Computer Science Reviews. Dr. Ghosh's expertise extends to teaching and mentoring roles, currently serving as an Associate Professor at IEM Kolkata. He is the founder of the IEM Centre of Excellence for Data Science and he is also the coordinator of the Innovation and Entrepreneurship Development Cell.

Title: Digital Twins in Healthcare: A Forefront for Knowledge Representation Techniques

Abstract: Digital twins have recently gathered significant interest in the healthcare community. This concept promises to unlock various previously unavailable services such as remote monitoring, advanced visualization, simulation of medical procedures, predictive analytics, demographic studies, and so on. At present research in this area is localized and conducted independently. Thus, effective deployment of digital twins in healthcare is still a work in progress due to inconsistent data representation and isolated innovation without effective integration at large

scale. Knowledge representation plays a vital role in structuring, integrating, and reasoning over heterogeneous healthcare data sources such as electronic health records, genomics data, clinical guidelines, reports, medical literature, and more. The process of digitization is relevant not only to patients but also to healthcare professionals, infrastructure facilities, devices, insurance providers, and even historical records. This work proposes to thoroughly highlight this research gap and the current initiatives addressing these issues. It aims to review and consolidate existing efforts in standardizing data structures for healthcare digital twins, with a focus on interoperability, representation and integration across diverse healthcare domains.

References