Preface

The rapid advancement of Natural Language Processing (NLP) and Large Language Models (LLMs) has transformed the landscape of computational linguistics. However, Indo-Aryan and Dravidian Languages (IADL), which represent a significant portion of South Asia's linguistic heritage, remain under-resourced and under-researched in these technological developments. This workshop aims to bridge this gap by bringing together researchers, linguists, and technologists to focus on the unique challenges and opportunities. Participants will explore innovative methods for creating and annotating digital corpora, develop speech and language technologies suited to IADL, and promote interdisciplinary collaborations. By leveraging LLMs, we seek to address the complexities of syntax, morphology, and semantics in these languages to enhance the performance of NLP applications. Furthermore, the workshop will provide a platform for sharing best practices, tools, and resources, enhancing the digital infrastructure necessary for language preservation. Through collaborative efforts, we aim to build a research community to advance NLP for IADL, contributing to linguistic diversity and cultural preservation in the digital age.

In parallel with the workshop, we have also organised a shared task to address key challenges in transliteration for Indian languages. The primary objectives of the shared task are to develop a real-time transliterator, effectively manage linguistic variations, and improve typing accuracy. A significant focus of the task is on enabling the transliterator to handle ad-hoc transliterations, which involve short typing scripts and diverse typing patterns, with or without vowel combinations. This initiative aims to create a robust transliteration system that accommodates the dynamic and complex nature of typing practices in Indian languages.

We received 27 submissions for the workshop and shared task. Following the review process, we accepted 15 papers and 4 shared task submissions to appear in the workshop proceedings.

The success of IndoNLP 2025 would not have been possible without the contributions of several exceptional individuals who supported this initiative. First and foremost, we extend our heartfelt gratitude to the authors who submitted their work to the workshop, driving forward research in low-resource languages across diverse areas of study. We are equally thankful to the program committee members, whose dedicated efforts were instrumental to the success of this workshop. Their timely engagement in the review process and constructive feedback not only enhanced the quality of the submissions but also ensured that the papers met the highest academic standards. Moreover we would like to thank to Prof. Pushpak Bhattacharyya for accepting our invitation to be as the keynote speaker in the workshop. Finally, we would like to express our sincere gratitude to the Informatics Institute of Technology, Colombo, for their generous sponsorship of the workshop. We are truly thankful to everyone who contributed to the success of IndoNLP 2025 through their invaluable support and encouragement.