

Report of Malayalam NLP Workshop (Trivandrum, Kerala), 2010

The regional workshop on NLP was conducted by LDC-IL at the Department of Linguistics, University of Kerala from 2nd - 6th August 2010. The program was as follows:

- 2nd August 2010

The program was inaugurated by the Registrar, University of Kerala. The Head of Linguistic Department, the professors of the various Departments along with all research scholars attended this program. On the first day, there was a presentation by Mr. Thennarasu, S. on the introduction about LDC-IL and its contribution in the area of Language Technology for Indian Languages.

- 3rd August 2010

On the second day of the workshop there was a presentation and discussion base on the POS tagging for Malayalam.

- 4th August 2010

On the third day, there were two presentations viz. one based on Source and Creation of Corpus and the other based on Linguistic Data Analysis and Methods.

- 5th August 2010

On the fourth day, there were four presentations: i) Speech Corpus: An Introduction ii) Making N-grams (A Corpus based study) iii) Morphological Analyzer for Malayalam iv) An Introduction to Speech Annotation using Praat.

- 6th August 2010

On the last day of the workshop, there was a discussion on the evaluation of POS Tagset for Malayalam and demo of LDC-IL NLP (Corpus Analysis) tools. The Certificates were distributed to the participants in the valedictory function.

This was an awareness raising workshop of LDC-IL. It aimed to create awareness and be better informed about the activities of LDC-IL. The participants were asked to comment on the approaches and the outcome of the workshop through feedback forms. About 100 participants participated in the workshop. Participants from both the disciplines humanities and science had participated. All language departments requested LDC-IL to conduct NLP workshop in their respective departments.