Assamese/Bodo NLP Workshop Report
(Department of Computer Science and Information Technology (CSIT), Gauhati University, Guwahati)
31st of January to 10th February 2012

The Linguistic Data Consortium for Indian Languages (LDC-IL), CIIL, Mysore organized a ten day Orientation cum training programme in NLP for Assamese and Bodo from the 31st of January to 10th February 2012 in collaboration with the Department of Computer Science and Information Technology (CSIT), Gauhati University, Guwahati. The first day started off with the inaugural function presided by the head of the department of CSIT. Prof. Shikhar Sharma. Prof. Sharma addressed the audience on the importance of NLP and made people aware of the various NLP projects being undertaken by CSIT. This was followed by an introductory address by Atreyee Sharma who familiarized the participants about the course of the programme and the motto of LDC-IL behind organizing such programmes all over India. After the formal registration and distribution of kits to the participants the lectures of the programme began. There were about 50 formal participants and overall 70 participants including the department staff etc. Participants were from Gauhati University, Dibrugarh University, Tezpur University and Assam University, from departments like Computer Science, Assamese, Bodo, Linguistics and English. Experts were from IIT Delhi (Dr. Priyankoo Sarmah Assistant Professor of linguistics in the department of HSS), Tezpur University (Prof. Madhumita Borbora, from the Department of English and Foreign languages and Dr. Utpal Sarmah, Head, Department of Computer Science and Engineering), Gauhati University (Prof. Swarna Prabha Chainary, Head, Department of Bodo), Prof. Shikhar Sarma, Head, Department of Computer Science and Information Technology (CSIT), Gauhati University and his team members also delivered a few lectures. The 5 member team of LDC-IL who went for the orientation also delivered lectures. The topics of the lectures ranged from basic linguistics to basic concepts of NLP and computational linguistics, corpus and its application, POS annotation, morphological analyzer, MT, OCR, encoding, machine learning, AI, Speech segmentation and annotation, TTS, ASR, word-net, chunking, syntactic parsing and spell checker. Totally there were 28 lectures covering 36 hours of teaching. On the 10th February after lunch, a valedictory function was organized to formally end the ten day orientation programme. Prof. Swarna Prabha Chainary, Head, Department of Bodo presided over the function and Atreyee Sharma and Prof. Shikhar Sarma too spoke on the occasion. Certificates were distributed to the participants after taking a brief feedback from them. The feedback was positive and participants had a unanimous vote of having more programmes of the sort again.