

## **Oriya NLP Workshop Report(School of Computer Science & Engineering, Kalinga Institute of Industrial Training ( KIIT University), Bhubaneswar, Odisha), 8th to 17th February, 2012.**

Linguistic Data Consortium for Indian languages (LDC-IL), CIIL organized a Ten-day Orientation cum Training Programme on Natural Language Processing in collaboration with the School of Computer Science & Engineering, Kalinga Institute of Industrial Training ( KIIT University), Bhubaneswar, Odisha from 8th February to 17th February,2012.

The programme was organized with certain goals such as to spread the knowledge of NLP among Odia scholars and to create awareness of NLP to promote language technology development in Indian Languages

In the inaugural function was on the 8th February at 10 am in the Conference Hall at the School of Computer Science & Engineering, KIIT University, Bhubaneswar. The Vice Chancellor of the University Prof. Ashok Kolasker presided over the inaugural session. In his speech he emphasized the importance of the development of computational approach for linguistics phenomena; he also encouraged the students to study linguistics much more. Prof. A.K. Bisoi the Dean of the KIIT University had given the key note speech it was followed by the welcome address by Dr. Anil Kumar Singh, Assistant Professor, School of Computer Science and Engineering, KIIT University, (also local coordinator). Finally Mr. Pramod Kumar Rout, Senior Research Assistant, LDC-IL had given the vote of thanks.

Nearly 60 participants (both scholars & lecturers) participated in the Orientation Programme. The participants were from universities like Utkal, Ravenshaw, Berhampur, Hi-tech Institute of Education, Sambalpur, KIIT & IIT Kharagpur, Some of the Students were from Jharkhand, Bihar, West Bengal, Manipur (all are KIIT student) who had taken keen interest to participate in all the days. The experts were from University of Hyderabad, IIT Bhubaneswar, KIIT University, Ravenshaw University, Utkal University, Berhampur University and some of the LDC-IL experts had also given lectures on NLP.

Classes started from 8th February at 11.45 am onwards. Every day 4 classes in two sessions were held and duration of each class was 1.15 hours. Lectures covered topics from Core linguistics, Corpus linguistics, language learning, Traditional grammar, Odia Lexicography & NLP module development and challenges.

The detailed schedule of the programme was as the table below.

<b>Day 1 (8th Feb, 2012)</b>		
<b>Expert</b>	<b>Time</b>	<b>Topic</b>
	10.00am	Inaugural Session
	11.30am	Tea Break
Prof. P. Mohanty	11.45am	Odia Morphology
	1.00pm	Lunch Break
Prof. P. Mohanty	2.00pm	Odia Morphology
	3.15pm	Tea Break
Shahnawaz	3.30 pm	Morphological Analysis
<b>Day 2 (9th Feb, 2012)</b>		
Prof. P. Mohanty	10.15am	Odia Syntax
	11.30am	Tea Break
Prof. P. Mohanty	11.45am	Odia Syntax
	1.00pm	Lunch Break
Dr. Anil Kr. Singh	2.00pm	Parsing & Chunking
	3.15pm	Tea Break
Sahid	3.30pm	Basic Concept of AI &NLP/CL
<b>Day 3 (10th Feb, 2012)</b>		
Prof. P. Mohanty	10.15am	Odia Semantics

	11.30am	Tea Break
Prof. P. Mohanty	11.45am	Odia Semantics
	1.00pm	Lunch Break
Sahid	2.00pm	Approaches to NLP Computational Grammar/Deductive Approach Data Driven/Statistical/Inductive Approach
	3.15pm	Tea Break
Sahid	3.30pm	Corpus (Text, Speech & Sign) - Concept, Balancing & Classification. Corpus Annotation-Concept, types &Application
<b>Day 4 (11th Feb, 2012)</b>		
Prof. P. Mohanty	10.15am	Odia Phonetics
	11.30 am	Tea Break
Prof. P. Mohanty	11.45am	Odia Phonetics
	1.00pm	Lunch Break
Dr. Prafulla Tripathy	2.00pm	Identifying Odia POS
	3.15pm	Tea Break
Dr. Prafulla Tripathy	3.30pm	Learning a new Indian Language
<b>Day 5 (13th Feb, 2012)</b>		
Dr. Santosh Tripathy	10.15pm	Traditional Grammar
	11.30am	Tea Break
Dr. Anil Kr. Singh	11.45am	ASR-Concept, Application & Challenges
	1.00pm	Lunch Break
Dr. R. Balabantray	2.00pm	Odia Morphology: Computational Approach
	3.15pm	Tea Break
Dr. R. Balabantray	3.30pm	Odia Spell Checker: Concept & Applications, Development & Challenges
<b>Day 6 (14th Feb, 2012)</b>		
Prof. Girish Ch. Mishra	10.15am	Machine Translation
	11.30 am	Tea Break
Prof. Girish Ch. Mishra	11.45am	TTS-Concept, application and Challenges
	1.00pm	Lunch Break
Dr. Anil Kr. Singh	2.00pm	TTS-Concept, Application, Development & Challenges
	3.15pm	Tea Break
Sriram Chaudhury	3.30pm	MT-Concept, Application and Development
<b>Day 7 (15th Feb, 2012)</b>		
Dr. Anil Ku. Singh	10.15am	Tool Demonstration
	11.30 am	Tea Break
Dr. Anil Ku. Singh	11.45am	OCR-Concept, Application, Development & Challenges
	1.00pm	Lunch Break
Sriram Chaudhury	2.00pm	Encoding –Concept, of Font & Encoding, ASCII, ISCII & Unicode
	3.15pm	Tea Break
Dr. Anil Kr. Singh	3.30pm	Corpus analysis tools (Frequency, N-gram, KWIC-KWOC, Concordances, etc.)
<b>Day 8 (16th Feb, 2012)</b>		
Dr. Anil Kr. Singh	10.15am	POS (annotation/tagging)
	11.30 am	Tea Break
Dr. Anil Kr. Singh	11.45am	Machine Learning (ML)- Basic Statics & Some Statistical Approaches ML (Decision Trees & Decision Lists, HMM, SVM &Neural Network)
	1.00pm	Lunch Break
Dr. Natabar Satapathy	2.00pm	Odia Lexicon Tradition
	3.15pm	Tea Break
Dr. Natabar Satapathy	3.30pm	
<b>Day 9 (17th Feb, 2012)</b>		
Dr. Sadananda Naik	10.15am	Odia Writing Tradition and Reforms
	11.30 am	Tea Break
Pramod	11.45am	Odia Morphological Analyzer & Suffix Stripping Model
Kshirod	12.15pm	Speech Segmentation & Annotation
	1.00pm	Lunch Break
Santosh	2.00pm	Odia Parts of Speech Annotation
	2.45pm	Tea Break
	3.00pm	Practice Session