

**NATIONAL SEMINAR ON INDIAN LANGUAGE TECHNOLOGY:
STATE AND PROSPECTS (Three days)**

21-23 March 2018

Technology and Resource Centre for Malayalam,

**Department of Linguistics
University of Kerala, Kariavattom, Trivandrum**

It is proposed to conduct a national seminar on the topic 'Indian Language Technology: State and Prospects'. Indian languages have made global strides in this era of information and a flurry of LT based systems is easily available even for layman. The benefits of information technology reach him only when software tools and human - machine interface systems are made available in his own language. All the major Indian languages have developed their linguistic resources to a great extent and attained a stage where each has devolved their strong theories and models along with their long tradition. The Three day seminar is proposed to assess the present status of the Indian linguistics and the language technology works related to them.

India is a multilingual country, with 22 official languages and 12 scripts. Indian languages belong to four different families, spoken in India and abroad. Only about 5% of Indians know English and hence the rest is deprived of the benefits of information technology. In order to broaden the scope of ICT in Indian languages tools, products and resources should be made familiar and freely available to the general public. The governments both at the centre and the states have taken initiatives for wider proliferation of Indian language Software Tools and Fonts. They also promote Language Technology standardization through active participation in International and national standardization bodies such as ISO, UNICODE, World-wide-Web consortium (W3C) and BIS (Bureau of Indian Standards) to ensure adequate representation of Indian languages in existing and future language technology standards. Projects on translation systems, information retrieval systems, speech based systems, etc. are being developed.

There have been initiatives for long term research for development of Machine Translation System, Optical Character Recognition, Online Handwriting Recognition System, Cross-lingual Information Access and Speech Processing, corpora development. The consistent and conscious efforts on the part of Government have fuelled the growth of industry in this sector. The spin-offs of these attempts have triggered an increasing interest of MNCs to look at India as a large market for Language Technologies. India is, thus, poised to emerge as Multilingual Computing hub.

Areas of Current Research

The projects like 'Technology Development for Indian Languages (TDIL), LDCIL at CIIL, WORDNET, the Programmes of Department of Electronics and Information Technology, IT Mission and IT@School of the state of Kerala, etc. work on the different research areas with the aim of enabling masses to build a knowledge society and ensuring communication without language barrier and moving up in the knowledge chain. Following are some major areas in the field.

- Development of corpora in machine readable for Machine Translation System
- Development of English to Indian Languages Machine Translation (MT) System

- Development of Indian Language to Indian Language Machine Translation System:
- Development of Cross-lingual Information Access
- Development of Robust Document Analysis & Recognition System for Indian Languages (OCR)
- Development of Online handwriting recognition system (OHWR)
- Development of Text to Speech System for Indian Languages (TTS)
- Development of regional language based Machine Translation System (SRLMT)

The proposed seminar topic has to be presented on the different perspectives of IT and Knowledge era. It is proposed to have a national seminar to think, discuss, explore, interact and interpret on all the above topics. It is also envisioned as a platform for researchers and professionals across the country to share their experiences related to natural language processing and computational linguistics.

Objective, Highlight and Theme of the Seminar, include the following broad areas for research.

1. The Policy and planning in Indian Language technology
2. Philosophy and theoretical aspects of Language engineering
3. History and development of Language technology in India
4. Models and theories in language Engineering
5. Projects on Phonology, Morphology, Syntax, Semantics and their Processing
6. Approaches to syntax/ grammar formalisms
7. Issues on Pragmatics, world knowledge and Discourse
8. Dialogue, speech and Dialogue Systems
9. Language Resources – Corpus, text, discourses and other data bases
10. Corpora in Indian languages and projects like LDCIL,
11. Machine Translation - and ML Language Models
12. POS Tagging and Parsing
13. Information Extraction/ Retrieval
14. Text Mining, Opinion Mining and Summarisation
15. NLP for Educational Applications
16. Recent NLP Advancements
17. University and Institute level Researches in language technology,
18. Public and Private Participations

Academicians, Language experts, Linguists, Literary personals, Sociologists, Language Teachers, Translators, Media persons, Software/ Net experts, Researchers, Policy makers, Computer scientists, and IT Experts, etc. will be the participants in the seminar.

Organizing Secretary
Dr. S. A Shanavas, Hon. Director & Associate Professor
Technology and Resource Centre for Malayalam,
University of Kerala, Kariavattom
Thiruvananthapuram – 695581 , Mob. ++91 9495074811
 Email: langtech2018.uok@gmail.com,

Schedule of the seminar

Notification– 25 January, 2018

Registration and Accepting Abstracts– 20th Feb– 2018
Last date of submission of Final paper– 08 March, 2018
Seminar Date– 21-23, 2018.