IJCNLP Workshop

on

Named Entity Recognition for South and South East Asian Languages

12 January, 2008, IIIT, Hyderabad, India

Submission Deadline: Sept 21, 2007

Motivation

Most of the South and South East Asian (SSEA) languages are scarce in resources and tools and Named Entity Recognition (NER) systems are no exception. It is very important that good systems for NER be available, because many problems in information extraction and machine translation are dependent on accurate NER. However, the issues involved are significantly different for these languages. For example, these languages do not have capitalization, which is a major feature for European languages.

Scope

Papers are invited on substantial, original, and unpublished research on all aspects of NER for SSEA languages. The workshop will have one track for regular research papers on NER for SSEA languages and the second track will be along the lines of a shared task.

Shared Task

In the shared task, the contestants having their own NER systems will be given some annotated test data. The participating systems will be ranked according to their performance on the test data. There may or may not be training data for a particular language. In either case, the contestants will have the freedom to use any technique for NER, e.g. a purely rule based technique or a purely statistical technique.

Important Dates
- Release of Training Data: Aug 2 to Aug 25, 2007 (for different languages)
- Release of Test Data: Sept 13, 2007
- Annotated Data Submission Deadline: Sept 15, 2007
- Paper Submission Deadline: Sept 21, 2007
- Notification of Paper Acceptance: Oct 26, 2007
- Camera Ready Submission Deadline: Nov 16, 2007

Contact Persons
Dipti Misra Sharma, Rajeev Sangal, Anil Kumar Singh
E-mail: dipti@iiit.ac.in, sangal@iiit.ac.in, anil@research.iiit.ac.in
Phone: 91-40-23001412

For more information visit: http://ltrc.iiit.ac.in/ner-ssea-08