

Help Build a Nepali Translator!

Of the approximately 7,000 languages in the world, less than 100 of them have automated instant translation capabilities. Microsoft Research recognizes that there is a need to facilitate the building of an automated Nepali translation system that is developed, improved and maintained by and for the Nepalese.

Kathmandu University, Language Technology Kendra, Nepal and Microsoft Research Connections, U.S.A. will be co-hosting a two day Machine Translation workshop on Jan 24-25, 2012. We invite students and the local community to participate in the creation of a machine translation system for the Nepali language.

We aim to involve at most 35 participants in the workshop - 15 computer science students, 15 students and professional linguists, Languages and Communication students and 5 people from other backgrounds like translation, social sciences and others. So the registration would be basically on a FIRST COME FIRST SERVE basis.

All participants would need to register for the workshop by Jan 2 by sending a confirmation email to mtworkshopnepal2012@hotmail.com and a filled in Workshop Participation Form as a reply to the Call for Participation. Acceptance to participate in the Workshop would be notified via email by Jan 6.

Those accepted for participating in the workshop are expected to develop 50 parallel texts (either English sentences and corresponding Nepali translations or Nepali sentences and corresponding English translations) and 50 test sentences (either English sentences or Nepali sentences) and submit via email by Jan 18, 2012.

Refreshments and certificates of participation will be provided to workshop participants.

Thank you and we look forward to working alongside you at the event!

Workshop schedule

January 24th, 2012:

Morning session

9:00 A.M - Arrival from Kathmandu to Dhulikhel

9:00 A.M - 9:15 A.M: Tea and cookies

9:15 A.M - 10.15 A.M: Orientation session.

MT workshop - An Overview, Orientation on the Microsoft MT Toolkit - Use, Training,

Development, Testing, Deployment

10:15 A.M – 11:40 A.M:

Movie show. [Parallel Event] Workshop participants are invited to attend a public screening of "The Linguists" hosted by David Harrison.

11:50 A.M - 1 P.M: Hands-on session of the Microsoft MT Toolkit (Creating Project owners, Project spaces, uploading samples, feeding sample training data, training the system, testing the system)

1 P.M - 2 P.M: Lunch break

January 24th continued

Afternoon session

2 P.M - 3 P.M [Parallel sessions]

Coding session - Coding, compilation and debugging [For Computer Science students]

Verification of English-Nepali/Nepali-English parallel sentences (texts) and uploading the verified parallel sentences (texts) to the system [For Linguists and Translators]

3 P.M - 3.15 P.M - Tea and snacks

3.15 P.M – 4.15 P.M: Testing of the Microsoft MT Toolkit

4:30 P.M: Leaving for Kathmandu by bus

January 25th:

Morning session

9:00 A.M - Arrival from Kathmandu to Dhulikhel

9:00 A.M - 9:15 A.M: Tea and cookies

9:15 A.M – 11.15 P.M [Parallel sessions]

Coding session - Coding, compilation and debugging [For Computer Science students]

Verification of English-Nepali/Nepali-English parallel sentences (texts) and uploading the verified parallel sentences (texts) to the system [For Linguists and Translators]

11.15 A.M – 12:45 P.M: Testing of the Microsoft MT Toolkit. Comparing BLEU scores.

1 P.M - 2 P.M: Lunch break

Afternoon session

2 P.M - 3 P.M: Closing session, Concluding remarks, Certificates distribution. Note on way forward.

3 P.M - 3:15 P.M: Group photographs, Tea and snacks

3:30 P.M – 4:30 P.M: Sightseeing of KU campus for guests and non-KU participants

5 P.M: Leaving for Kathmandu by bus