

## **A Micro Crowdsourcing Architecture to Localize Web Content for Less-Resourced Languages**

Asanka Wasala

Chris Exton

Ruvan Weerasinghe\*

Reinhard Schäler

Centre for Next Generation Localisation/Localisation Research Centre,  
CSIS Dept, University of Limerick, Limerick, Ireland.

\* Language Technology Research Laboratory, University of Colombo School of  
Computing, 35 Reid Avenue, Colombo 00700, Sri Lanka.

{Asanka.Wasala, Chris.Exton, Reinhard.Schaler}@ul.ie, \*arw@ucsc.cmb.ac.lk

### **Abstract**

We will report on a novel browser extension-based client-server architecture using open standards that allows localization of web content using the power of the crowd. We address issues related to MT-based solutions and propose an alternative approach based on translation memories (TMs). The approach is inspired by Exton et al. (2009) on real-time localization of desktop software using the crowd and Wasala and Weerasinghe (2008) on browser based pop-up dictionary extensions. The architectural approach chosen enables in-context real-time localization of web content supported by the crowd. To best of our knowledge, this is the only practical web content localization methodology currently being proposed that incorporates Translation Memories. The approach also supports the building of resources such as parallel corpora – resources that are still not available for many, but especially for under-served languages.