# Machine translation and language resources for under-resourced European languages challenges and experiences

## Invited Talk

*Mikel L. Forcada,*
Universitat d'Alacant and Apertium, Spain
E-mail : mlf@dlsi.ua.es

## Abstract

Multilingualism is at the core of Europe's identity. However, the official multilingualism of European Union institutions is usually restricted to a set of 24 official languages (bul, ces, dan, deu, ell, eng, spa, est, fin, fra, gle, hrv, hun, ita, lit, lav, mlt, nld, pol, por, ron, slk, slv, swe), with some symbolic support for 5 more semi-official languages (cat, cym, eus, gla, glg). As a result, European funding for machine translation and language resources usually covers these languages, but not many other languages in which Europeans live a great part of their everyday lifes.

In this talk, I will discuss some of the issues and challenges involved in creating machine translation (MT) and MT-related language resources for less-resourced languages of Europe, will show the advantages of a free/open-source, collaborative approach, and will argue in favour of knowledge-based approaches, illustrating them with three cases: two under-resourced languages (Breton and Occitan) and a better resourced language (Catalan).