CoSyne

Multilingual Content Synchronization with Wikis

The overall aim of CoSyne is to automate the dynamic multilingual synchronization process of Wikis.

Main Objectives

The project’s aim is to synchronize user-generated content in different languages through machine translation, cross-lingual entailment checking and (language-independent) structural information. The adaptability of the resulting CoSyne system will be demonstrated using content in four core languages (English, German, Dutch, Italian), and two languages with more limited resources (Bulgarian, Turkish).

Technical Functionalities

The end result of the CoSyne project will be a synchronization tool for Wikis, which automatically merges pages written in different languages by using:

- An adaptive machine translation component, which translates novel information from one page to another, and adapts its translation model by following user edits of the automatically translated content;

- A cross-language textual entailment recognition component, which identifies overlapping and non-overlapping text fragments from different language versions of a page, thus providing information about the text portions to be translated and added;

- A structural analysis module, which builds parallel structural information for different language versions of a page, to facilitate content analysis for entailment and translation.

Wiki Environment

The components of CoSyne will be integrated through web services with the open-source MediaWiki platform, which is the most commonly used Wiki platform.

At a glance: CoSyne

Consortium:
University of Amsterdam (UvA), The Netherlands (co-ordinator), Fondazione Bruno Kessler (FBK), Italy, Dublin City University (DCU), Ireland, Heidelberg Institute for Theoretical Studies (HITS), Germany, Netherlands Institute for Sound and Vision (NISV), The Netherlands, Deutsche Welle (DW), Germany, Vereniging WikiMedia Nederland (WMF), The Netherlands