# CoSyne, a Project on Multilingual Content Synchronization with Wikis

**7ᵗʰ Framework Programme**
**FP7-ICT-Call 4**
**small or medium-scale-focused research project**
**248531**
**http://www.cosyne.eu**

| |
|---|
| University of Amsterdam (UvA), The Netherlands (coordinator) |
| Fondazione Bruno Kessler (FBK), Italy |
| Dublin City University (DCU), Ireland |
| Heidelberg Institute for Technical Studies (HITS), Germany |
| Deutsche Welle (DW), Germany |
| Netherlands Institute for Sound & Vision (NISV), The Netherlands |
| WikiMedia Foundation NL, The Netherlands |

**Project duration: 1 March 2010 — 28 February 2013**

**Summary**

CoSyne aims at automating the dynamic multilingual synchronization process of wikis. It deals with automating the process of analyzing and mutually enriching different wiki pages on the same subject. This includes, but is not limited to, user-generated content. The CoSyne project focuses on robust machine translation and synchronization in six designated languages. Because of the synchronization aspect, the work is done in a wiki environment, and the different services and components use the open-source MediaWiki platform.

The strength of the CoSyne system compared to its competitors is the combination of machine translation with synchronization. The system automatically recognizes which parts of text are not present in the other language version(s) and translates or augments only those parts that are considered different using segment-specific adaptive modeling.

At present the project is entering its third year of activity. In the first year, the focus was on a limited set of language combinations: German-English, Dutch-English and Italian-English (all in both directions). The second year this was expanded to include a combination among all four languages (English, German, Dutch and Italian). The third year will focus on enhanced prototyping, machine learning with identification of factual changes and style changes, and analysis of logged user edits to teach and improve the system. Two additional languages (Turkish and Bulgarian) will be used to test the system for new languages, in the final phase of the project. Contact: c.monz@uva.nl www.cosyne.eu