

## **The TermWise Project: Leveraging Big Data for Terminological Support in Legal Translation**

*Kris Heylen, KU Leuven*

This presentation reports on the TermWise project, a cooperation of terminologists, corpus linguists and computer scientists at KU Leuven. The project responds to a need for more terminological support expressed by the translators at the Belgian Federal Ministry of Justice. Although they use a state-of-the-art CAT tool, their termbase and translation memory only provides limited coverage of relevant legal expressions. Yet, large collections of official bilingual legal documents containing approved translations are available online. Currently, translators search these collections with a general-purpose search engine, but finding relevant examples is often hard and time-consuming. In the TermWise project, we therefore developed dedicated knowledge extraction algorithms that extract bilingual legal terminology automatically from large online document collections and we offer those to the user in a system that identifies the most relevant previous translation example of a domain-specific expression relative to the user's current translation assignment. In the presentation we give an overview of the system, give a demo of the user interface and, more in general, discuss the possibilities of mining big data to support specialised translation.