PLuTO: Why overcoming language barriers fosters a more innovative society
European consortium formed to improve access to multilingual patent digital libraries

Entering a new market is a costly endeavour for individual inventors and SMEs. Searches must be made for any existing patents in the domain and this requires high-level technical and legal expertise. Language barriers represent an additional hurdle, because of the specific language in patents: both technical and legal at the same time. The multi-national distribution of patent sources adds further difficulties due to the need to identify and connect different data sources in multiple languages.

In order to overcome these linguistic barriers and to facilitate the further building of the European Digital Information Space, the European Commission is co-funding a project that will facilitate cross-language retrieval of patent information. PLuTO (Patent Language Translations Online) is funded under the ICT Policy Support Programme (ICT PSP) and has a duration of 3 years (April 2010 – March 2013). The project consortium comprises both academic and industrial partners across 5 European countries: the Centre for Next Generation Localisation (CNGL) of the Dublin City University, Ireland; the Information Retrieval Facility (IRF), Austria; ESTeam, Sweden; Cross Language, Belgium; and the Dutch Patent Information User Group (WON), The Netherlands.

"By funding pilot projects through the ICT-PSP programme, the European Commission aims at showing the public value of research", Susan Fraser, project officer at the European Commission, explains. "If PLuTO manages to give small inventors and SMEs the possibility to exploit their inventions more quickly and at lower cost, it fits perfectly into a sustainable European innovation strategy". The consortium partners aim to develop a rapid solution for online translation services, by integrating existing components and iteratively improving them. The CNGL brings to the target platform a state-of-the-art machine translation engine, MaTrEx, which exploits hybrid statistical, example-based and hierarchical techniques and has demonstrated high quality translation performance in a number of recent evaluation campaigns. ESTeam contributes a comprehensive translation software environment to the project, including server-based, multi-layered, multi-domain translation memory technology. Information retrieval expertise is provided by the IRF which also gives access to its data on patent search use-cases and a large scale, multi-lingual patent repository. PLuTO will also exploit the evaluation expertise of Cross Language who has developed a methodology for evaluating machine translation output. WON will be directly involved in all phases of development, providing valuable user feedback. The consortium also intends to collaborate closely with the European Patent Office in order to profit from their experience in this area.

PLuTO will provide an integrated, online search and translation tool where several human experts (technical, legal, consultants) can take advantage of existing web-content and state-of-the-art machine translation tools to collaboratively search for, select and translate patents. PLuTO will make domain-adaptable machine translation systems more readily available to potential users of such technology. This service will ultimately give rise to a more inclusive innovation society where individual inventors and SMEs have access to a large pool of international knowledge without language barriers.

For more information about the PLuTO project, please contact:

Marie-Pierre Garnier, Information Retrieval Facility, Operngasse 20B, 1040 Vienna
Email : mp.garnier@ir-facility.org, Phone : +43 1 236 94 74-6069