

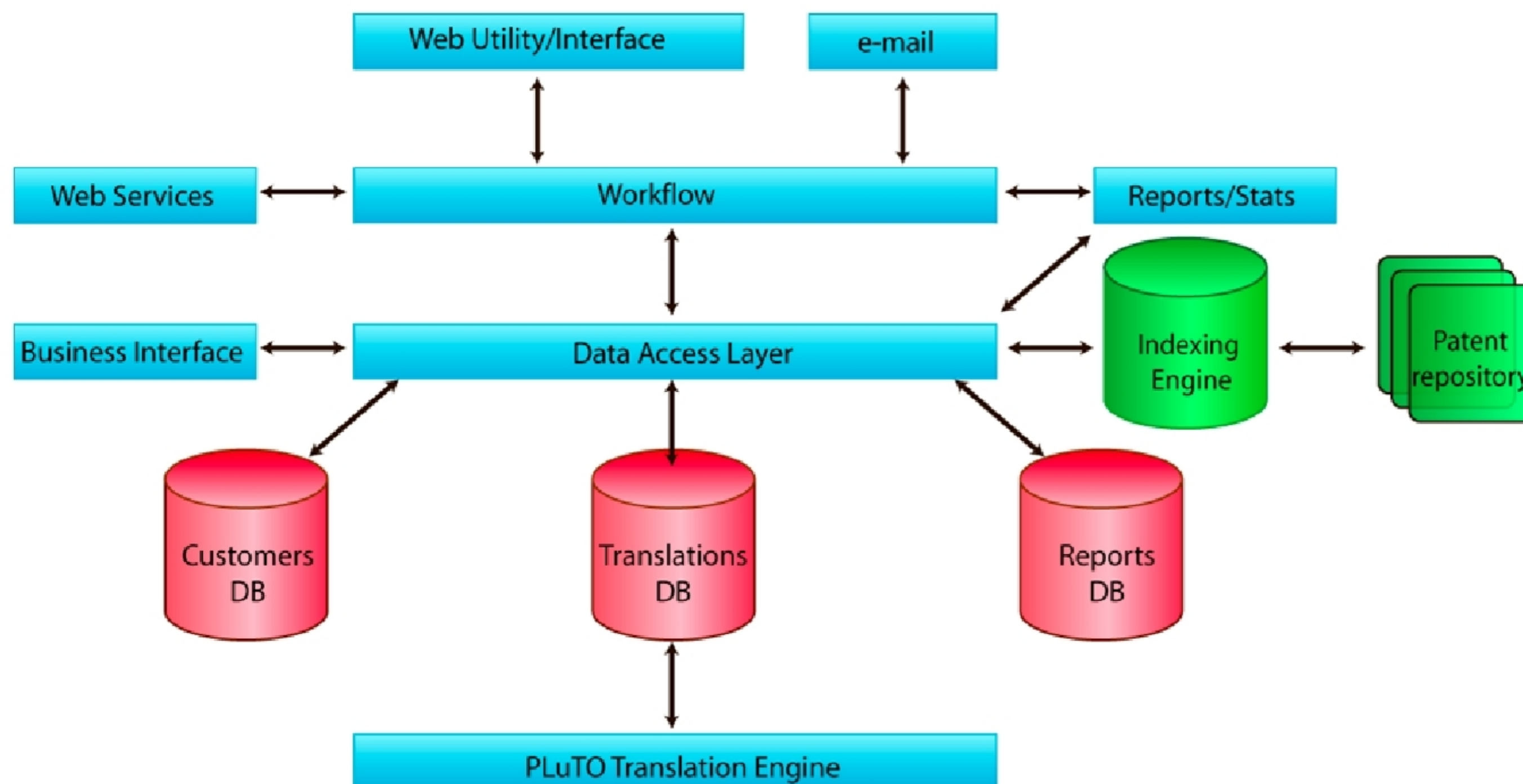
PLuTO: Patent Language Translations Online

Executive Summary

The problem addressed by this project is the increase, in a large number of countries, in IP protection activities, particularly related to patents (applications, oppositions, and infringement lawsuits). This has resulted in an increased and urgent need for tools to operate cross language barriers, as language differences are no excuse in cases of patent infringement. The specific language of patents – both technical and legal at the same time – makes it particularly difficult to overcome these language barriers. Furthermore, the multi-national distribution of patent sources adds additional difficulties because of the need to identify and connect to different data sources in multiple languages.

PLuTO aims to overcome these barriers by providing an integrated, online translation tool, where several human experts (technical, legal, consultants) can take advantage of existing web-content, state-of-the-art data-driven machine translation and information retrieval tools to collaboratively retrieve and translate patents. PLuTO will also result in the commoditisation of domain-adaptable machine translation systems, making them more visible 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

PLuTO System Architecture



Components

MaTrEx (CNGL/DCU)

Hybrid statistical, example-based and hierarchical machine translation framework.

ESTeam Translator

Comprehensive translation software environment including translation memory.

Lucene/SOLR Indexing Engine (IRF)

State-of-the-art open-source platform from IR, including multi-lingual analysers and stemmers.

Patent Repository

Including data from the IRF, WON, and the European Patent Office.

Languages

PLuTO will initially provide patent search and translation for 12 language pairs: English <=> French, German, Spanish, Dutch, Swedish, and Russian.

Language choice is flexible based on user statistics from the EPO's esp@cenet service

Objectives

- Facilitate cross-language retrieval of patent information
- Provide online translation service
- Accelerate translations of patent documents
- Provide an all-encompassing tool for prior art search to both SMEs and individual inventors

Impact

- More inclusive innovation society: Individual inventors and SMEs will have access to data and automatic translations
- Expansion of business opportunities by overcoming language barriers
- Broad user base will stimulate discussion and innovation
- Collaborative, data-driven platform can be easily extended to other, less well represented (European) languages => enhanced business visibility of such nations

For more info

Project URL:
<http://www.pluto-patenttranslation.eu>

Contact:
Dr. Páraic Sheridan
Phone: +353 (1) 7006706
Email: psheridan@computing.dcu.ie

Project Partners

Centre for Next Generation Localisation, Dublin City University, Ireland



- Project co-ordinator;
- State-of-the-art machine translation (MT) engine;
- Expertise in development, evaluation and deployment of MT;
- Significant information retrieval (IR) experience.
- URL: <http://www.cngl.ie>

Information Retrieval Facility, Austria



- IR expertise – state-of-the-art retrieval engine;
- Experience in organisation and execution of IR evaluation;
- Substantial patent data collections.
- URL: <http://www.ir-facility.org/>

ESTeam, Sweden



- Translation memory (TM) expertise including large TM database;
- Own in-house MT system;
- Significant experience in the deployment of commercial grade MT solutions.
- URL: <http://www.esteam.se/>

Cross Language, Belgium



- Expertise in applied translation solutions;
- Human evaluation of translated documents from a linguistic perspective;
- URL: <http://www.crosslang.com>

Dutch Patent Information User Group (WON), the Netherlands



- Patent experts;
- Human evaluation of translated documents from a patent perspective;
- URL: <http://www.won-nl.org>