Project Information

Project Acronym:	**PANACEA**
Project Number:	**246064**
Funding Scheme:	**Small or medium scale focused research project (STREP)**
Duration:	**36 months**
Start date:	**1 January 2010**
EU funding:	**€2.68 million**

THE PANACEA CONSORTIUM:

Coordinator	Institution
Núria Bel	**Universitat Pompeu Fabra (Spain)**
Nicoletta Calzolari	**Instituto di Linguistica Computazionale (Italy)**
Stelios Piperidis	**Institute for Language & Speech Processing (Greece)**
Anna Korhonen	**University of Cambridge (United Kingdom)**
Andy Way	**Dublin City University (Ireland)**
Gregor Thurmair	**Linguatec (Germany)**
Khalid Choukri	**Evaluation & LR Distribution Agency** (France)

www.panacea-lr.eu

SEVENTH FRAMEWORK PROGRAMME

If:
- You are a researcher on the area of automatic production of language resources;
- You have language resources whose production you want to automate;
- You are interested in this project and you want to receive information about its development.

Please, contact us at: *info@panacea-lr.eu*

Platform for **A**utomatic, **N**ormalized **A**nnotation and **C**ost-**E**ffective **A**cquisition

of Language Resources for Human Language Technologies

"A web service-based production line that automates the stages involved in the acquisition, production, updating and maintenance of the Language Resources required by MT and other Language Technologies"

SEVENTH FRAMEWORK PROGRAMME

www.panacea-lr.eu

(7FP-ITC-248064)

http://cordis.europa.eu/fp7/ict/language-technologies/

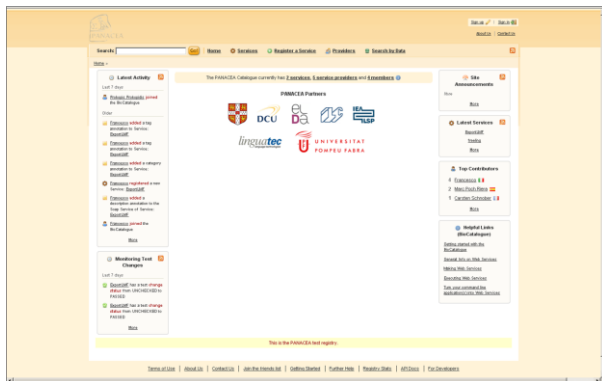**Information and Communication Technologies**

A Platform for the production of Language Resources

PANACEA is working to create a "Factory" of Language Resources (LRs) in the form of a production line that automates all the steps involved in the acquisition, production, updating and maintenance of LRs required by Machine Translation and by other Language Technologies. PANACEA's utilizes different advanced technological components, which can be smoothly assembled in an integrated process that still enables the quick integration of new or different components for:

- Monolingual and Parallel Text Acquisition and Pre-processing;
- Sentential and sub-sentential alignment;
- Bilingual Dictionary and Transfer Grammar production;
- Lexical Information Acquisition for rich information dictionary production.

Cost and time reduction by automation is the only way to ensure the continuous supply of LR's that can guarantee a LT industry covering all languages, all domains, for current and future needs, and in the time required by the market.
PANACEA's contribution and impact will be demonstrated by a significant reduction in time and cost in the production of LRs for a real life use case.
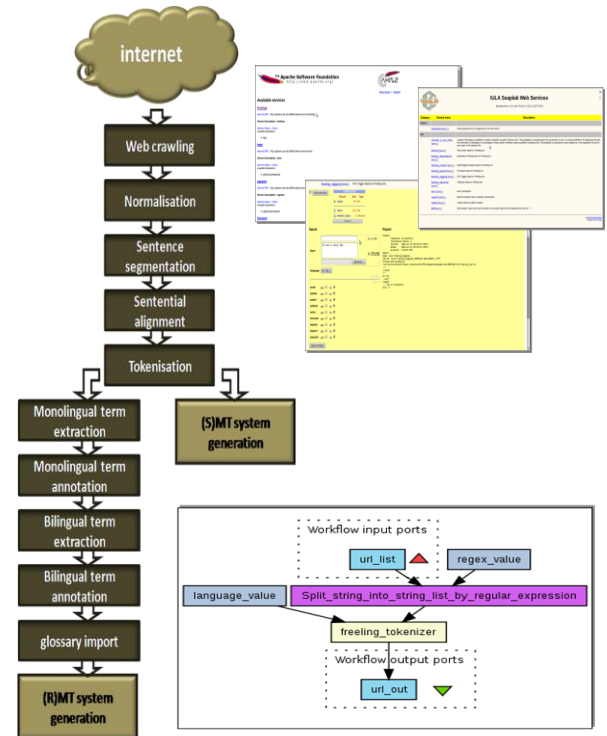
The first version of the platform and interoperable web Visit www.panacea-lr.eu to access PANACEA's platform and web services. The services will be available by February 2011.

1.   First versión of PANACEA webservice registry

Project Challenges

- To propose the definition of an interoperable space for the integration of robust, scalable, web service-deployed components;
- To handle the impact of dealing with massive data;
- To be convincing on the industrial use of the available automatic acquisition technologies by introducing ready-to-use tools, with confidence indicators, which give priority to high precision;
- To carry out further research to increase accuracy in the automatic acquisition and production of Language Resources.

2.   Different perspectives of web service workflows

Impact on Human Language Technologies

PANACEA's goal is to build a Language Resource "Factory" that ensures the continuous supply of LRs needed by the Language Technology industry to solve current and future needs and problems. Its necessity is seen, for example, with the ability of Machine Translation systems to accurately cover all languages and domains in the time required by the market.

This document is part of dissemination material generated in the PANACEA Project, **P**latform for **A**utomatic, **N**ormalized **A**nnotation and **C**ost-**E**ffective **A**cquisition (Grant Agreement no. 248064).

Please send feedback and questions on this document to: iulatrl@upf.edu

TRL Group (Tecnologies dels Recursos Lingüístics), Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra (IULA-UPF)

Author: Núria Bel. Barcelona, 2010