
Knowledge Base for RTD Competencies in IST



Project Acronym: IST World
Project Full Title: Knowledge Base for RTD Competencies in IST
Project Coordinator: Prof. Hans Uszkoreit
German Research Center for Artificial Intelligence (DFKI)
Tel: +49 681 302 5282; Fax: +49 681 302 5338
Email: ist-world@dfki.de

IST World Portal: <http://www.ist-world.org/>
Project Website: <http://ist-world.dfki.de/>

Partners: German Research Center for Artificial Intelligence (DE), Institute Jozef Stefan (SI), Ontotext Lab - Sirma AI EAD (BG), RTD Talos (CY), Institute of Information Theory and Automation (CZ), Archimedes Foundation (EE), Computer and Automation Research Institute - Hungarian Academy of Sciences (HU), Institute of Mathematics and Computer Science -University of Latvia (LV), Lithuanian Innovation Centre (LT), Projects in Motion (MT), Technical University of Silesia (PL), National Institute for Research and Development in Informatics (RO), Slovak University of Technology (SK), TUBITAK (TK), The Council for the Central Laboratory of the Research Councils (UK)

Further Information: Specific Support Action – IST call 3: FP6
DG Information Society & Media - Unit F4

Project Officer: Leoni Schaefer

Objectives

The IST World project aims at setting up and populating an information portal with innovative functionalities that helps to promote RTD competencies in IST on a local, national and European level: The IST World portal will

- provide information about RTD actors, such as persons, organizations and projects and their experience and expertise
- promote RTD competencies in specific fields
- provide innovative and automated services
- support partner search for IST proposals and commercial projects

Functionalities of the Portal

The IST World portal will contain information about RTD. The information structure of the portal will be suitable for representing personal and organizational competencies, expertise and social networks. Web forms will be provided for easy self-registration of RTD actors. Web mining techniques will be employed for automated updates. The portal will overcome the shortcomings of the existing on-line services by offering advanced analytical and prediction services and by joining knowledge mapping with social networks building. The portal will offer the following functionalities:

1. **Information Repository:** information about partners, their resources, active projects and their expertise.
2. **Partner Finding Tool:** predicting the optimum consortia of partners based on their competences, experiences and trust.
3. **Social Network Identification:** analysis of present research activities, actors, social networks and results visualized by different techniques.
4. **Forecasting/Prediction:** forecasting of RTD trends based on monitoring of current research initiatives, projects and achievements, and predicting possible future research themes based on automatically detected trends.
5. **Expertise Identification:** summarizing and presenting different aspects of a person's complex expertise profile based on extraction of information from a potentially huge number of web search results.

Data Sources

In order to achieve a critical mass of accurate data and comprehensive coverage, one cannot rely on one single way of data acquisition. Therefore, there are three major ways to populate the database:

- **Base** - First of all, existing data will be imported from CORDIS, from domain-specific portals such as LT World, and from CERIF-conformant national and European Current Research Information Systems (CRISs).
- **Community based** - The portal will offer services for community building and maintenance in which organisations, groups and experts will register themselves, and their projects and build professional virtual communities. Promotion campaigns in the new member states (NMS) and associated candidate countries (ACC), aimed especially at SMEs, will be carried out in order to ensure that research competencies from these countries are well represented in the database.
- **Automatic** - Web and text mining techniques will be used to acquire additional data that is not yet present in the database.

Methodologies and Technologies

The portal functionalities are based on techniques and tools developed by DFKI, Jozef Stefan Institute, and Ontotext in previous and current research projects. Advanced techniques for automatic data acquisition, classification, knowledge extraction, trend detection and prediction will be implemented as services to support partner search for IST proposals and commercial projects. Techniques from social network analysis will be used for trust net monitoring, authority identification, community identification and monitoring, and prediction of community evolution.

The technologies have been utilized successfully in the established information portals, Project Intelligence (<http://pi.ijs.si/>), a service for analysis and visualization of collaborations in European research projects, LT World (<http://www.lt-world.org/>),

a comprehensive ontology-based information portal for the field of language technology, and KIM (<http://www.ontotext.com/kim/>) a platform for knowledge and information management. The data models will be based on CERIF (Common European Research Information Format), an open set of guidelines for research information systems. The methodology is scalable and can be easily applied to large collections of project descriptions.

Core Technologies

| | Information Repository | Partner Finding Tool | Social Network Identification | Forecasting and Prediction | Expertise Identification |
|---|------------------------|----------------------|-------------------------------|----------------------------|--------------------------|
| Web Crawling | X | | | | X |
| Cleaning and consolidating of multiple data sources | X | | X | | X |
| Language Technologies | X | | | | X |
| Data / Text / Web Mining | X | | X | X | X |
| Social-network / Link analysis | | X | X | X | X |
| User Modelling | | X | | | (X) |
| Presentation / Visualisation of Results | | X | X | X | X |
| Community Building | | X | X | | |

IST World Portal Milestones

- Month 5:** First version of the portal with basic functionality
- Month 11:** Enhanced version of the portal based on data mining and visualization
- Month 18:** Second version of the portal with full functionality
- Month 24:** Improved version of the portal

Target User Groups and Expected Impact

Foreseen users are all individuals and organizations in need of in-depth information on European IST experience and competencies. However the specially targeted user group is companies and research centers from new and future member states of the EU seeking to participate in international RTD activities or to benefit from their results. IST World will help these organizations to find suitable partners or to promote their capabilities, services and technologies.

The project contributes to the construction of the European Research Area, which functions as a "common market" for RTD services, technologies, and experts. The project will make market activities more efficient by providing high-quality information that enables an improved matching of demand with supply. A further benefit is that the portal provides a map of the research competencies in Europe and the NMS/ACC in particular. This will reveal local strengths and clusters of innovative organisations, as well as gaps and weaknesses in particular areas. The data mining and visualisation technologies will allow the detection and analysis of patterns in the data, and will be of use for partner search, investors and policy-makers. Users will be involved in the design, evaluation and improvement of the portal through a representative User Steering Group.