

# Proiecte europene

---

## PROIECTUL EUROPEAN IST-WORLD

### „BAZĂ DE CUNOȘTINȚE PRIVIND COMPETENȚELE DE CERCETARE ȘI DEZVOLTARE TEHNOLOGICĂ DIN DOMENIUL TEHNOLOGIILOR SOCIETĂȚII INFORMAȚIONALE”

Gabriel Neagu

Ileana Trandafir

Ana-Maria Borozan

*Institutul Național de Cercetare – Dezvoltare în Informatică, ICI, București*

*ist-world@ici.ro*

IST World este un proiect de tipul „acțiune suport specifică” din cadrul Programului cadru 6, care se desfășoară în perioada aprilie 2005 - septembrie 2007. Proiectul își propune realizarea și popularea unui portal informațional cu funcții inovative, care să promoveze competențele de cercetare-dezvoltare tehnologică din domeniul tehnologiilor societății informaționale (IST), la nivel local, național și european. Acest portal va permite:

- furnizarea de informații despre principalii actori care desfășoară activități de cercetare-dezvoltare tehnologică (CDT), cum sunt: persoane, organizații, proiecte, cât și experiența și expertiza acestora,
- promovarea competențelor dintr-un domeniu specific,
- furnizarea de servicii automatizate și inovative,
- sprijinirea căutării de parteneri pentru propunerea de proiecte IST sau de proiecte comerciale.

La realizarea proiectului participă un consorțiu larg, format din: Centrul German de Cercetare pentru Inteligență Artificială (Germania), Institutul Jozef Stefan (Slovenia), Ontotext Lab - Sirma Group Corp. (Bulgaria), Institutul de Teorie a Informației și Automatică (Cehia), RTD Talos (Cipru), Fundația Arhimede (Estonia), Institutul de Matematică și Calculatoare - Universitatea din Letonia, Centrul de Inovare din Lituania, Proiecte în Mișcare (Malta), Consiliul pentru Laboratorul Central al Consiliilor de Cercetare (Marea Britanie), Universitatea de Tehnologie din Silezia (Polonia), **Institutul Național de Cercetare - Dezvoltare în Informatică, ICI, București** (România), Universitatea de Tehnologie (Slovacia), TUBITAK (Turcia), Institutul de Cercetare pentru Calculatoare și Automatică - Academia de Științe din Ungaria.

## 1. Funcționalitatea portalului

Portalul IST World (o primă versiune poate fi accesată la <http://www.ist-world.org/>) va conține informații organizate într-o manieră care să permită reprezentarea competențelor persoanelor individuale și ale organizațiilor, expertiza acestora și rețelele sociale. Formularele web furnizate vor permite o înregistrare fără efort a actorilor CDT. Tehnicile de web mining utilizate vor facilita actualizarea automată a informațiilor. Portalul acesta va oferi facilități superioare celor existente atât prin punerea la dispoziția utilizatorilor a unor servicii avansate de analiză și de predicție, cât și prin concertarea mapping-ului de cunoștințe cu construirea de rețele sociale. Portalul va oferi următoarele funcționalități:

1. depozitarea informațiilor (**Information Repository**): informații despre parteneri, resurse, proiecte active, experiența acestora;
2. găsirea de parteneri (**Partner Finding Tool**): propunerea / predicția unui consorțiu de parteneri, optim bazat pe competențe, experiență și încredere;
3. identificarea rețelilor sociale (**Social Network Identification**): analiza activităților actuale de cercetare, actori, rețele sociale (conexiuni între actori) și rezultate vizualizate prin diferite tehnici;
4. previzionarea / predicția (**Forecasting / Prediction**): previzionarea tendințelor CDT, bazată pe monitorizarea inițiativelor curente de cercetare, a proiectelor și realizărilor și predicția viitoarelor teme posibile de cercetare bazată pe detectarea automată a tendințelor;
5. identificarea expertizei (**Expertize Identification**): sintetizarea și prezentarea diferitelor aspecte privind profilul complex de expertiză al persoanelor, pe baza extragerii și analizei unui volum important de căutări pe web site.

## 2. Sursele de date

Pentru a atinge masa critică de date de maximă acuratețe și cu un grad de acoperire corespunzător, se consideră că nu este suficientă utilizarea unei singure căi de colectare a informațiilor. Proiectul ia în considerare trei căi principale de populare a bazei de date:

- **baze existente** - în primul rând, importul de date existente din CORDIS, din portaluri pe domenii specifice IST, ca de exemplu LT World (pentru domeniul „language technology”) și din sistemele naționale și europene, conforme cu standardul CERIF (European Current Research Information Systems - CRISs);
- **baze constituite pe comunități** - portalul va oferi servicii pentru constituirea și întreținerea de comunități în care organizațiile, grupurile și experții se pot înregistra singuri, putând forma comunități profesionale virtuale. În acest scop, se vor organiza campanii de promovare, mai ales în noile state membre și state asociate candidate la UE, cu prioritate pentru IMM-uri, pentru a asigura o bună reprezentare a competențelor acestora în baza de date;
- **în mod automat** - tehnicile de web mining și text mining utilizate vor permite obținerea de date adiționale față de cele existente în baza de date.

## 3. Metodologii și tehnologii

Funcționalitatea portalului va fi realizată cu ajutorul unor tehnici și instrumente software, realizate de DFKI - Josef Stefan Institute (Slovenia) și Ontotext (Bulgaria) în cadrul unor alte proiecte de cercetare comune, finalizate sau în curs de desfășurare. Tehnicile avansate de achiziție automată a datelor, clasificare, extragere de cunoștințe, detectarea și predicția tendințelor vor fi implementate ca servicii care să susțină căutarea de parteneri pentru propunerea de proiecte IST și proiecte comerciale. Tehnicile de analiză a rețelelor sociale vor fi utilizate pentru monitorizarea relațiilor de încredere, identificare a zonelor de autoritate, identificare și monitorizare a comunităților și de predicție a evoluției acestora.

Aceste tehnologii au fost utilizate cu succes în portaluri informaționale operaționale în prezent: *Project Intelligence* (<http://pi.ijs.si/>), un serviciu de analiză și vizualizare a colaborărilor din proiectele europene de cercetare, *LT World* (<http://www.lt-world.org/>), un portal bazat pe ontologii din domeniul tehnologiei limbajului și *KIM* (<http://www.ontotext.com/kim>), o platformă pentru managementul informațiilor și cunoștințelor. Modelul datelor va fi bazat pe standardul CERIF (*Common European Research Information Format*), un set deschis de recomandări pentru sistemele informatice ale cercetării. Metodologia respectivă este scalabilă și poate fi ușor aplicată la o gamă largă de proiecte.

Tehnologiile de bază utilizate sunt prezentate în tabelul alăturat.

Tehnologii	Depozitarea informațiilor	Găsirea de parteneri	Identificarea rețelelor sociale	Previzionarea și predicția	Identificarea expertizei
Web Crawling	x				x
Cleaning and consolidating of multiple data sources	x		x		x
Language technology	x				x
Data / text / web mining	x		x	x	x
Social-network / link analysis		x	x	x	x
User modelling		x			(x)
Presentation / visualition Results		x	x	x	x
Community Building		x	x		

## 4. Grupurile țintă și impactul proiectului

Utilizatorii previzibili sunt indivizii și organizațiile interesate în obținerea informațiilor referitoare la experiența și competențele europene în domeniul IST. Grupurile țintă, vizate în mod special de acest proiect, sunt companiile și centrele de cercetare din noile și viitoare state membre ale UE pentru a le facilita participarea la activitățile CDT internaționale și a beneficia de rezultate. IST World dorește să ajute aceste organizații să-și găsească parteneri adecvați sau să-și promoveze capacitățile, serviciile și tehnologiile. Proiectul contribuie astfel la

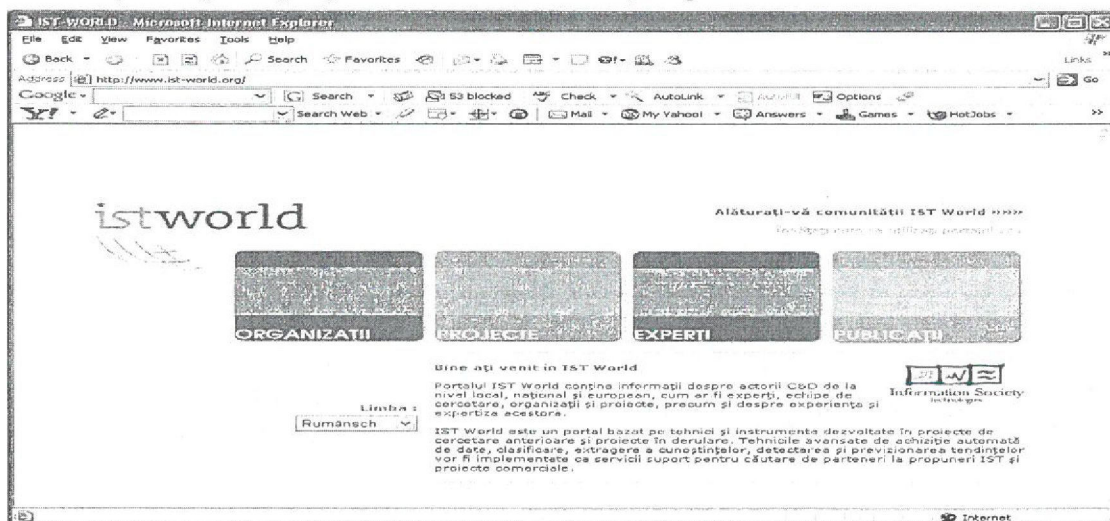


construcția spațiului de cercetare european (ERA), care funcționează ca o „piață comună” a serviciilor, tehnologiilor și experților CDT. Proiectul va face activitățile acestei piețe mai eficiente prin furnizarea de informații de înaltă calitate și prin mai buna echilibrare a cererii cu oferta. Un beneficiu important este acela că portalul va furniza o „hartă” a principalelor competențe de cercetare din Europa și, în mod particular, din noile și viitoarele state membre ale UE. Aceasta va evidenția atât punctele tari și cluster-ele de organizații inovative, cât și punctele slabe sau rămănerile în urmă din diferite zone. Tehnologiile de data mining și vizualizare vor permite detectarea și analiza datelor pentru găsirea de parteneri, investitori sau decidenți politici. Utilizatorii vor fi implicați în proiectarea, evaluarea și îmbunătățirea portalului prin intermediul *User Steering Group*, ca organism special creat în structurile de lucru și decizie ale proiectului.

## 5. Prima versiune a portalului

În primul an al proiectului, a fost finalizată o primă versiune a portalului, disponibilă la adresa <http://www.ist-world.org/>, care permite accesul (multilingv, inclusiv în limba română) la depozitul de date IST World integrat din trei surse distincte de date, așa cum se va vedea în continuare. Funcțiile operaționale sunt:

- căutare (căutare simplă și avansată după organizații, proiecte, experți, publicații) și navigare în baza de date integrată,
- tehnica de vizualizare analitică a diagramei de colaborare dintre diferite entități reprezentate prin informațiile stocate,
- interfața de import-export pentru seturile de date XML, compatibile cu CERIF.



Depozitul de date este administrat cu MS SQL Server 2005. Schema bazei de date este proiectată pe baza metodologiei CERIF și implementează modelul relațional CERIF extins, cu ajutorul script-urilor furnizate de grupul CERIF și a extensiilor dezvoltate de echipa proiectului. Schema bazei de date este deosebit de complexă și constă din 361 tabele.

Cele trei surse de date utilizate pentru popularea bazei de date au fost:

- **datele SICRIS** - baza de date CRIS (adaptată la CERIF) care conține informații de-facto despre activitățile de cercetare din Slovenia: 579 organizații, 11147 cercetători, 2982 proiecte de cercetare, 597 programe, suportate de Ministerul Educației și Sportului și de Ministerul Științei și Învățământului Superior din Slovenia;
- **datele GOOGLE SCHOLAR** - bază de date care conține informații despre literatură educațională, recenzii, teze de doctorat, cărți, manuscrise, rezumate, rapoarte tehnice din diferite domenii de cercetare. Google Scholar include tehnici de web crawling pentru găsirea publicațiilor de la o varietate de edituri academice, societăți profesionale, depozite de manuscrise, universități, accesibile pe web. Subsetul selectat în cadrul proiectului conține 63 474 persoane și 58 300 publicații;
- **Project Intelligence (PI) - date CORDIS** - portalul Project Intelligence utilizează datele publice din sistemul Cordis și din datele Comisiei Europene. Subsetul selectat conține 3181 organizații și 663 proiecte.

Toate sursele menționate au cumulativ aproximativ 141.000 entități, disponibile în prezent în baza de date.

Prima versiune a portalului permite căutarea în baza de date integrată și analiza rezultatelor căutării cu ajutorul diagramelor de colaborare, funcțiuni adoptate din cele două portaluri de referință menționate anterior: Project Intelligence și LT World. Interfața tehnică actuală cu depozitul de date IST World (back end) permite o integrare simplă și eficientă a seturilor de date XML, compatibile CERIF din diferite surse de date.

Prima versiune a portalului a fost realizată utilizând Microsoft Internet Information Server bazat pe .NET v2.0. Proiectarea paginii web urmează recomandările de programare ASP.NET.

Dezvoltarea portalului IST World continuă conform planului proiectului, urmând să se extindă cu noi informații și facilități; o nouă versiune cu îmbunătățiri majore fiind planificată să devină disponibilă în toamna 2006.

Informații de detaliu despre proiect și noi inițiative generate de acesta pot fi găsite pe web site-ul proiectului: <http://ist-world.dfki.de/>.

## 6. Activitatea părții române

Ca și ceilalți reprezentanți din noile și viitoarele state membre ale UE, **Institutul Național de Cercetare - Dezvoltare în Informatică, ICI**, București are ca principală sarcină asigurarea unui nivel de relevanță corespunzător (ca volum și arie de cuprindere) pentru datele referitoare la domeniul cercetării-dezvoltării IST în România. În acest sens, principalele sarcini se referă la: activarea surselor reprezentative de date la nivel național, asigurarea importului acestor date în baza de date IST World, asigurarea interfeței în limba română pentru serviciile portalului și desfășurarea unei campanii active de promovare și diseminare a informației privind oferta de servicii a proiectului.



# IST World

Aveți nevoie de parteneri de cercetare cu experiență –pentru proiecte internaționale?

- Sunteți în căutare de competențe specifice sau de noi tehnologii?
- Sunteți doar curioși să aflați potențialul creativ, distribuția regională și cooperările existente în cercetarea IST europeană și pe scena tehnologiilor?

<http://www.ist-world.org/>

Găsiți parteneri,  
tehnologii,  
competențe  
și tendințe ...

Portalul IST World oferă informații despre experți, grupuri de cercetare, centre și companii implicate în crearea de tehnologii pentru dezvoltarea societății informaționale.

Scopul principal al serviciilor IST World îl constituie expertiza și experiența jucătorilor relevanți din țările europene.

Coordonarea proiectului:  
Centrul German de Cercetare pentru Inteligență Artificială (DFKI GmbH)  
Prof. Hans Uszkoreit  
Saarbrücken, Germania  
Tel: +49 681 302 5282  
Fax: +49 681 302 5338

Punct de contact în România:  
Institutul Național de Cercetare-Dezvoltare în Informatică, ICI București  
Dr. ing. Gabriel Neagu  
Tel: +40 21 316 07 36  
Fax: +40 21 316 10 30  
Email: [ist-world@ictr.ro](mailto:ist-world@ictr.ro)

Parteneri:

Centrul German de Cercetare pentru Inteligență Artificială (Germania), Institutul Jozef Stefan (Slovenia), Ontotext Lab - Sirna Group Corp. (Bulgaria), Institutul de Teorie a Informației și Automatică (Cehia), RTD Talos (Cipru), Fundația Arfimed (Estonia), Institutul de Matematică și Calculatoare - Universitatea din Letonia, Centrul de Inovare din Lituania, Proiecte în Mișcare (Malta), Consiliul pentru Laboratorul Central al Consiliilor de Cercetare (Marea Britanie), Universitatea de Tehnologie din Silezia (Polonia), Institutul Național de Cercetare-Dezvoltare în Informatică (România), Universitatea de Tehnologie (Slovacia), TUBITAK (Turcia), Institutul de Cercetare pentru Calculatoare și Automatică - Academia de Științe din Ungaria.



Information Society  
Technologies

IST World este finanțat de Comisia Europeană în cadrul PC6 în IST - Contract nr.: FP6-2004-IST-3-015823.