



CACAO projekt katalógusok, digitális könyvtárak lekérdezése saját nyelven

Moldován István
Országos Széchényi Könyvtár



Áttekintés

- Helyzetkép és feladat
- A soknyelvűség problémái
- A projekt célja, keretei
- Eszközök, megközelítések
- Kivitelezés, technikai modell
- Eddigi eredmények
- Résztvevők
- Tervek



Helyzetkép és feladat

- Az Európai Unióban 23 hivatalos nyelv van.
- Sok nemzetiségű felhasználók, sokféle nyelvhasználat az európai könyvtárakban;
- Soknyelvű állomány;
- A feltárás, katalógusok jellemzően egy nyelven
- A többnyelvű keresés igénye.
Kezdeményezések: többnyelvű tezauruszok
-> CACAO Cross-language Access to catalogues and on-line libraries



A projekt célja, keretei

- Célja: Egy olyan innovatív infrastruktúra létrehozása, amely lehetővé teszi az elektronikus könyvtárakban és könyvtárkatalógusokban rendelkezésre álló többnyelvű szöveges tartalmak elérését, megértését valamint az ezekben való tájékozódást, ezáltal segítve az európai felhasználókat, hogy jobban kiaknázhassák a rendelkezésre álló információkat.
- Időtartam: 2007. 12. – 2009. 11.30.



A soknyelvűség problémái 1.

- „False friends” – ugyanannak a szónak más nyelven más jelentése lehet; pl. Gift – ajándék (en) vagy mérég (de)?!, cane – nád (en) vagy kutya (it)?
- A pontos kereséshez nem elegendő a szavak sima fordítása



La Fontaine fable ?

*Die Quelle
Fabel?*

La Fontana favola ?

The Fountain fable ?

Źródło legenda ?

forrás mese ?

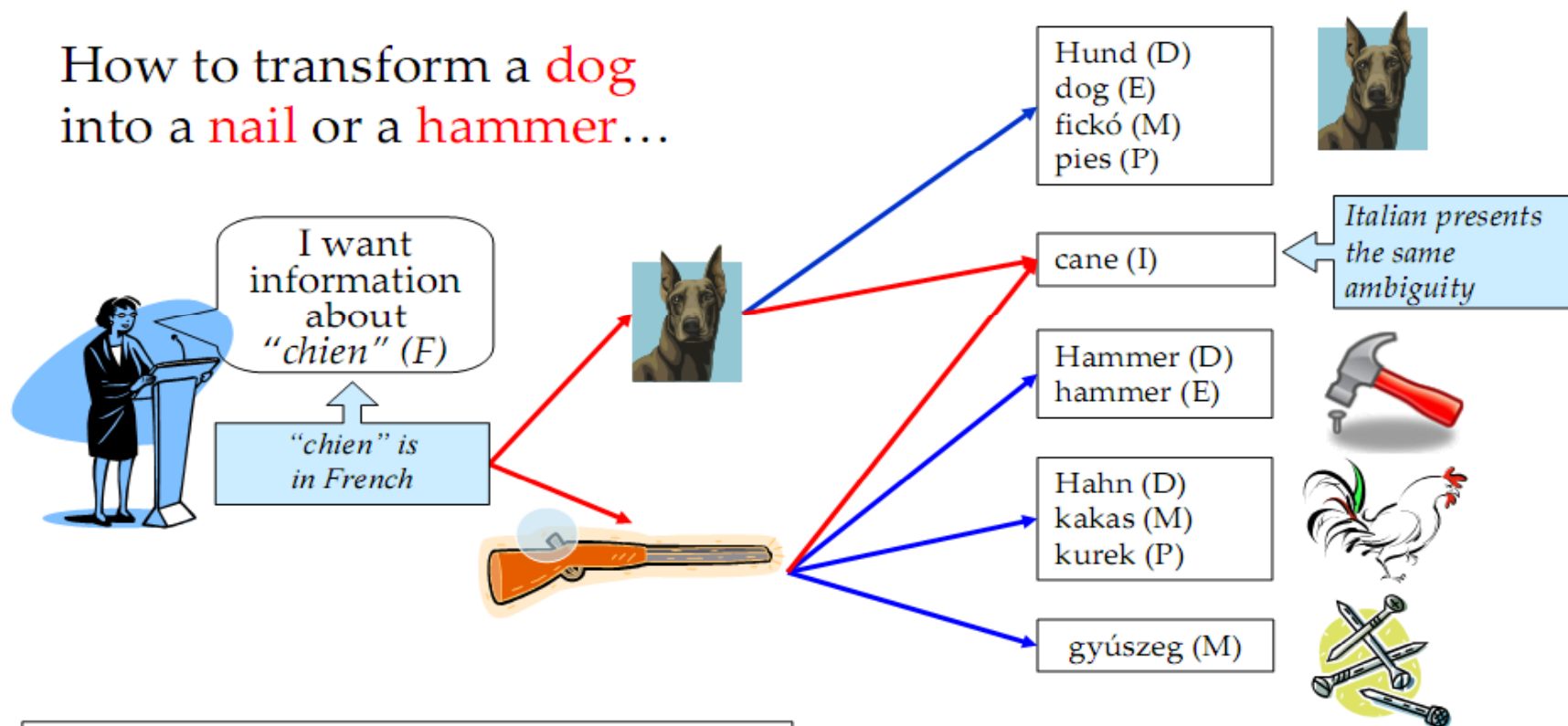
Proper names and locations should be identified to avoid wrong translations.



A soknyelvűség problémái 2.

Other issues: words are ambiguous

How to transform a **dog** into a **nail** or a **hammer**...



Translation does not reduce to a simple dictionary search...

D=Deutsch	M=Magyar	F=Français
E=English	P=Polski	I=Italiano

Eszközök, megközelítések



- A fordítás pontosítása;
 - metaadatok nyelvi kóddal való ellátása,
pl. `<dc:subject xml:lang="de">Gift</dc:subject>`
 - a többértelmű kifejezések szűkítése az osztályozási sémák segítségével.
- A tartalmi feltáró eszközök különbözősége, pl. ;
 - LC Subject Headings,
 - Schlagwortnormdatei,
 - Göttinger Online-Klassifikation,
 - Regensburger
 - Verbundklassifikation

Kivitelezés, technikai modell



- A leíró rekordok begyűjtése, aggregálás; OAI-PMH, Z39.50, SRU/SRW
- A keresőszavak lefordítása;
- Nyelvi elemző modul,
- A lefordított szavak összevetése a klasszifikációs rendszerekkel. PI
- Dog (en) = Cane (it) – kutya vagy ütőszeg!
+ állat vagy fegyverek csoport?



cane







Search term language: English German French Italian Polish

History

cane.

Total documents: 478 in: English, German, French and Italian.

English: 386 results for: **dog,bitch,blackguard** [Search info](#)



-  [The genetics of the dog](#)
Publisher: Wallingford [u.a.] : CABI
-  [Cadillac man : my life on the streets](#)
Publisher: New York : Bloomsbury
-  [Genetics for dog breeders / Robinson, Roy](#)
Publisher: Oxford [u.a.] : Pergamon Press
-  [Red dog / De Bernières, Louis](#)
Publisher: London : Secker & Warburg
-  [Bitch money / Traynor, Joanna](#)
Publisher: London : Bloomsbury
-  [A dog's life / Bailey, Paul](#)
Publisher: London [u.a.] : Hamish Hamilton


[Get more results...](#)

German: 38 results for: **Schuft,Schurke,Zicke** [Search info](#)

French: 46 results for: **bandit,chienne,goujat** [Search info](#)

Italian: 8 results for: **cane** [Search info](#)

Suchen | **Suchergebnis** | Erweiterte Suche | Zwischenablage | Benutzerkonto | Hilfe  



Suchen [und] [ALL] Alle Wörter

sortiert nach Erscheinungsjahr

Unscharfe Suche

Benutzerkennung: GAST | [Abmelden](#) | [Datenbankmenü](#)

[Suchgeschichte](#) | [Kurzliste](#) | [Titelanzeige](#) Copyright © 2009 OCLC

Ihre Suchanfrage: [Suchanfrage](#) | [Kopiebestellung](#) | [subito-Bestellung](#) | [Trefferanalyse](#) | [Speichern/Drucken](#)
Aktion [und] **331027593** 1 Treffer



Titel: Red dog / Louis De Bernières

Verfasser: De Bernières, Louis *1954-*

Erschienen: London : Secker & Warburg, 2001

Umfang: VII, 119 S. : Ill.

ISBN: 0-436-25617-7 : £10.00
0-436-20580-7

Schlagwörter: *Dogs / Australia / Fiction ; Australia / Social life and customs / Fiction

Sachgruppen: 18.07 ; Englische Literatur außerhalb Großbritanniens und der USA

Mehr zum Thema: 823.914 Dewey-Dezimalklassifikation



Résztevők

1. Koordinátor: Xerox Európai Kutatóközpont, Europea, Franciaország
2. CELI
3. Bolzanói Szabadegyetem, Olaszország
4. Lengyel Tudományos Akadémia, Kornik Könyvtár
5. Cité des Sciences et de l'Industrie, Franciaország
6. Gonetwerk, Olaszország
7. Magyar Tudományos Akadémia Nyelvtudományi Intézet, Magyarország
8. Göttingeni Állami Egyetemi Könyvtár, Németország
9. Országos Széchényi Könyvtár, Magyarország



Tervek

- Három tematikus portál kialakítása az EDL számára:
 - Földrajz,
 - Matematika,
 - Európai történelem,
- A CACAO algoritmusának alkalmazása
 - Könyvtári OPAC-okban,
 - Elektronikus könyvtárakban,
 - A TEL katalógusában.



Köszönöm a figyelmet

<http://www.cacaoprojekt.eu>

Moldován István

OSZK

moldovan@oszk.hu