

**DELOS NoE  
on Digital Libraries**

**WP7: Evaluation**

**WP Leader: Norbert Fuhr  
University of Duisburg-Essen, Germany**

## WP7 tasks

- 7.3: INEX - Initiative for the Evaluation of XML Retrieval
- 7.4 CLEF – Cross-Language Evaluation Forum
- 7.5 A Digital Library Testbed Framework for the Evaluation of Architectures, Services and Execution Dynamics

# **Task 7.3: INEX - Initiative for the Evaluation of XML Retrieval**

**Task leader: Mounia Lalmas  
Queen Mary University of London, UK**

# INEX Evaluation Initiative: Activities 2005

## Objectives of INEX

Promote research and stimulate development of XML information access and retrieval, through

- Creation of evaluation infrastructure and organisation of regular evaluation campaigns for system testing
- Building of an XML information access and retrieval research community
- Construction of test-suites

**INEX 2005** has allowed a new community in XML information access to emerge, as shown by the number of publications (64+ in 2005, 37 in 2004 and 13 in 2003). INEX is allowing research in XML information access to thrive.

## INEX 2005 Results: Expansion of Test Collections

- **Ad-hoc Track:**
  - fulltexts in XML from 18 IEEE-CS journals
  - 12,107 +4,712 articles (1995-2004), 494 +241 MB
- **Multimedia track: Lonely Planet World Guide**
  - 462 XML documents with 1800 photos, 452 maps
- **Document mining**
  - WIPO corpus: 75,250 XML documents
  - MovieDB corpus (based on the Internet Movie Database): 9463 XML documents
- **Heterogeneous track**
  - 7 corpora with 900,000+ XML documents

# INEX 2005: Results

## Research in New Directions I

- **Interactive track**
  - 3 different tasks
  - 11 participating groups with ~100 test persons
  - How do users interact with the search system (basis for the definition of metrics)
  - How can interactive retrieval become more effective? (functionality, UI design)
- **Heterogeneous track**
  - new collections with different DTDs and their effect on XML IR system effectiveness.
- **Relevance feedback**
  - queries that also include structural hints (rather than content-only queries in 2004).

# INEX 2005: Results

## Research in New Directions II

- **Natural language query processing track**
  - New task allowing new participants with NLP expertise to join INEX without the need to develop a search engine, and thus encouraging wider accessibility.
- **Document mining track**
  - Collaboration with the PASCAL network of Excellence
  - Development of machine learning methods for structured data mining
  - This year: classification and clustering for XML documents.
- **Multimedia track**
  - Access methods for structured multimedia documents
  - This year: text and image retrieval

# INEX 2005: Results

## Research in New Directions IV

- **Ad-hoc track** a simulation of how a XML content-based library might be used.  
assumptions about user requirements led to three different strategies:
  - **focussed**: user prefers a single element that most exhaustively and specifically discusses the topic of the query
  - **thorough**: user prefers all highly exhaustive and specific elements.
  - **fetch and browse**: user is interested in highly exhaustive and specific elements that are contained only within highly relevant articles.



# INEX 2005: Results

## Research in New Directions V

- New set of metrics:  
eXtended Cumulated Gain (XCG)
  - evaluation framework that considers dependency among XML document components
  - incorporate mechanisms to reward the retrieval of near-misses and to address issues of overlap.
  - system and user-oriented evaluation aspects and recall and precision-like qualities are measured.

# INEX 2005: Overall Results

- stimulation of research activity in new, previously unexplored areas, such as interactive aspects of XML information access and multimedia XML access,
- study and implementation of evaluation methodologies for diverse types of XML access systems
- novel evaluation methodology where user aspects can be formalised in the evaluation measures (XCG metrics)
- development of test collections and evaluation methodologies for XML access system evaluation
- building of a strong, multidisciplinary research community
- participation is up: 64 groups in 2005

# Plans for INEX 2006

## INEX 2005 Tracks

- XML retrieval (ad-hoc)
- Interactive track
- Heterogeneous track
- Relevance feedback track
- Natural language query processing track
- Multimedia XML retrieval track
- Document mining track
- **New tracks**
  - XML entity search track (under discussion)

**Details of new tracks and new tasks and evaluation methodologies under discussion**

**Further development of the infrastructure for results analysis and testing of methodology.**

# **Task 7.4: CLEF – Cross language Evaluation Forum**

**Task Leader: Carol Peters  
CNR-ISTI, Italy**

# Cross Language Evaluation Forum: Activity 2005

## Objectives of CLEF

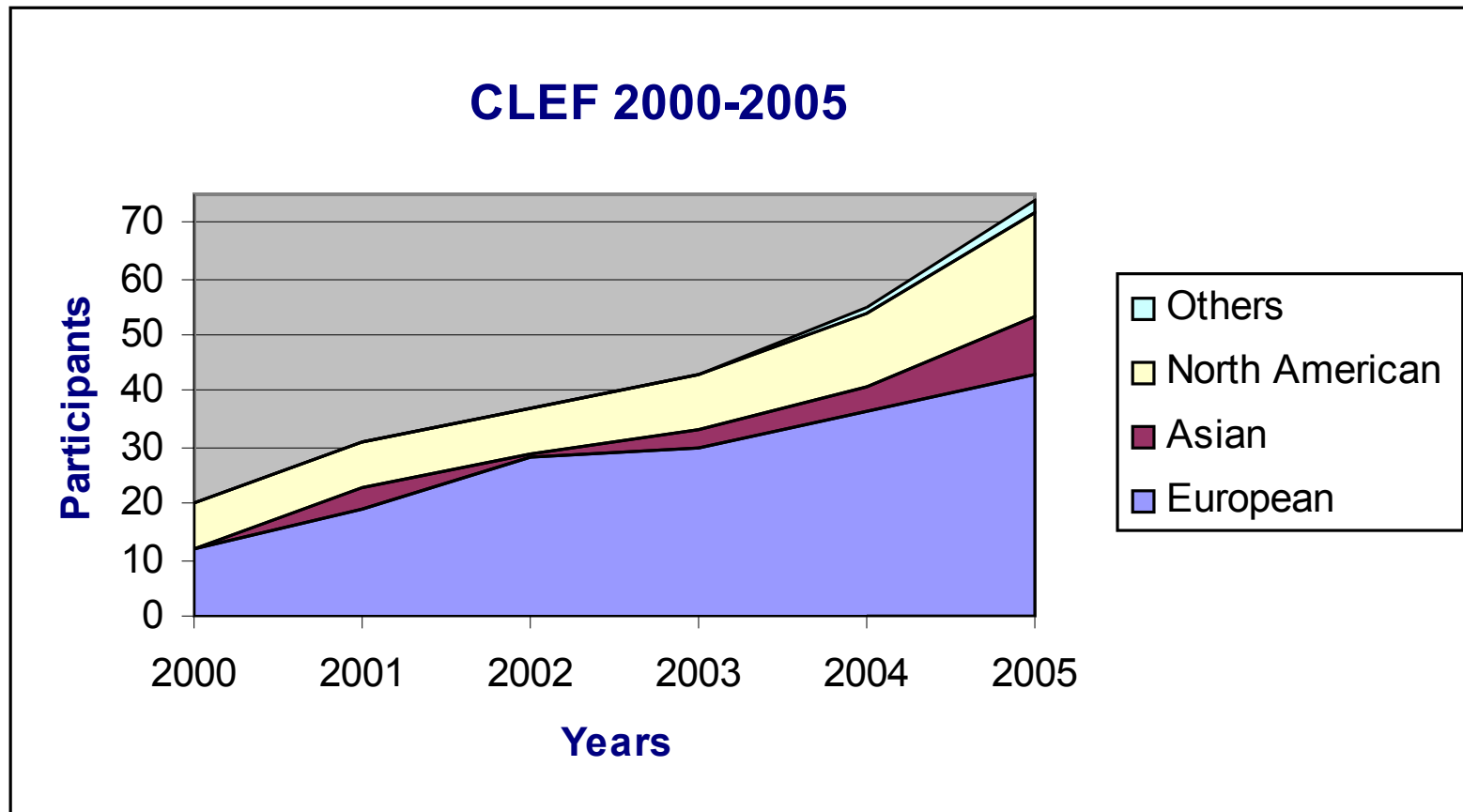
Promote research and stimulate development of multilingual IR systems for European languages, through

- Creation of evaluation infrastructure and organisation of regular evaluation campaigns for system testing
- Building of an MLIA/CLIR research community
- Construction of publicly available test-suites

## CLEF 2005

- continued expansion of participation, tasks, and test collections
- Focus shifted from cross-language textual document retrieval to include various types of information extraction in multilingual/multimedia context.

# CLEF: Growth in Participation



# CLEF 2005 Results: Research in New Directions I

- **Ad-hoc track:**
  - results merging over collections/over languages and at measuring progress in multilingual IR system development over time;
- **Interactive track:**
  - comparative study of user-inclusive cross-language search strategies
- **Question-answering track:**
  - mono- and cross-language QA systems.
  - new types of natural language questions
  - new evaluation measures – namely the K1 value and r coefficient
- **Image retrieval track:**
  - combining text and content-based retrieval for cross-language image retrieval
  - effectiveness of combining text and image for retrieval

# CLEF 2005: Results

## Research in New Directions II

- **Speech track**
  - searching spontaneous speech from oral historical interviews (Shoah archives)
- **Web track**
  - Construction of a multilingual web corpus
  - Resource for investigating multilingual web retrieval.
- **Geographic information retrieval track**
  - retrieval of multilingual documents with an emphasis on geographic search
  - first time that GIR systems have been evaluated in a multilingual context
  - strong interest, especially from industry, was encouraging.



# CLEF 2005: Overall Results

- Stimulation of research activity in new, previously unexplored areas, such as cross-language question answering, image and geographic information retrieval
- Study and implementation of evaluation methodologies for diverse types of cross-language IR systems
- Documented improvement in system performance for cross-language text retrieval systems
- Creation of a large set of empirical data about multilingual information access from the user perspective;
- Quantitative and qualitative evidence with respect to best practice in cross-language system development
- R&D activity in new areas such as cross-language question answering, multilingual retrieval for mixed media, and cross-language geographic information retrieval
- Creation of important, reusable test collections for system benchmarking building of a strong, multidisciplinary research community

# Plans for CLEF 2006

## **CLEF 2006 Tracks (same tracks – new tasks)**

- Multilingual, Bilingual and Monolingual (non-English) Text Retrieval on News Collections
- Domain-Specific Retrieval on Social Science Docs
- Interactive Cross-Language IR
- Multilingual Question Answering
- Cross-Language Image Retrieval
- Cross-Language Speech Retrieval
- Multilingual Web Track
- Cross-Language Geographical IR

**Details of new tasks and evaluation methodologies now under discussion**

**Further development of the infrastructure for results analysis and testing of pooling methodology.**

# **Task 7.5: A Digital Library Testbed Framework for the Evaluation of Architectures, Services and Execution Dynamics**

**Task Leader: Claus-Peter Klas  
University of Duisburg-Essen**

# DL Testbed: Objectives

- Standard Testbed Framework for DL Evaluation
  - Guide, Repository, Testbed-Framework Daffodil
- Standard Event Model of DL users and services
  - Logging standard schema and corresponding evaluation tools
- Evaluation Actions:
  - INEX Baseline System
  - The European Library
  - Collaboration in Computer Science DLs

# Standard Event Model for DL Usage and Services

- Digital Library Life Cycle  
Discover - Retrieve - Collate - Interpret - Re-Present
- Logging levels
  1. **User behavior level:** context, task, user info, etc.
  2. **Concept level:** browse - store - annotate - help - navigate - inspect – communicate – author
  3. **Service level:** metadata search, detail search, journal browse, .....
  4. **HCI (keystroke) level:** mouse movement/clicking, keyboard input, eye tracking
  5. **System level:** server load, resource, network traffic etc.

# Comparative Evaluation: The European Library

Ευρωπαϊκή Βιβλιοθήκη La Biblioteca Europea Det europeiske bibliotek La bibliothèque européenne De Europese Bibliotheek  
 k The European Library Det europeiska biblioteket Avrupa Kütüphanesi Euroopska knjižnica Eiropas bibliotēka  
 a Evropska knjižnica Az Európai Könyvtár Die Europäische Bibliothek Euroopan kirjasto Euroopa raamatukogu

**SEARCH** COLLECTIONS LIBRARIES TREASURES ABOUT US  
 Simple search **Advanced search** Results

## ADVANCED SEARCH

On start up you search in the [default list of collections](#). For more specific results [choose your own collections](#).

**1 SELECT COLLECTIONS** ?

▶ [choose your own collections](#)  
 ▶ [reset: default list of collections](#)

- EU Online books, images, maps, music...
- UK British Library integrated catalogue
- PT Biblioteca Nacional (Portugal)
- FR BN-OPALE PLUS
- DE Online catalogue of the Deutsche Bibliothek
- FI HELKA
- IT SBN OPAC
- NL General Catalogue Koninklijke Bibliotheek
- CH HELVETICAT
- SI National Library Catalogue (KatNUK)
- KR Croatian National Bibliographies of Books : 1990-2003
- SR Serbian Union Catalogue COBIB.SR

**2 SPECIFY YOUR QUERY** ?

Specify a query for one or more fields.

Any field

and  Title

and  Author

and  Subject

▶ [Add more fields](#) **RESET**

**SEARCH**

**SEARCH RESULTS** ?

Matches for '(digital libraries)' .

[Stop the search](#)

EU	<a href="#">Online books, images, maps, music...</a>	0
UK	<b><a href="#">British Library integrated catalogue</a></b>	<b>390</b>
PT	<a href="#">Biblioteca Nacional (Portugal)</a>	1
FR	<a href="#">BN-OPALE PLUS</a>	217
DE	<a href="#">Online catalogue of the Deutsche Bibliothek</a>	...
FI	<a href="#">HELKA</a>	0
IT	<a href="#">SBN OPAC</a>	46
NL	<a href="#">General Catalogue Koninklijke Bibliotheek</a>	305
CH	<a href="#">HELVETICAT</a>	...
SI	<a href="#">National Library Catalogue (KatNUK)</a>	...
KR	<a href="#">Croatian National Bibliographies of Books : 1990-2003</a>	...
SR	<a href="#">Serbian Union Catalogue COBIB.SR</a>	...

**British Library integrated catalogue - UK -**

390 objects with (digital libraries) have been found in 'British Library integrated catalogue'.

jump to

[next](#)

1	<p>▶ <a href="#">Colleges, code, and copyright : the impact of digital networks and technological controls on copyright and the dissemination of information in higher education / Center for Intellectual Property, University of Maryland, University College.</a></p> <p>Document type: text , language materials , monograph   Language: eng</p>	
2	<p>▶ <a href="#">Joint conference on digital libraries (ICDL) 2004 [electronic resource] : global reach and diverse impact.</a></p> <p>Document type: text , language materials , part   Language: eng</p>	
3	<p>▶ <a href="#">LIDA 2005.</a></p> <p>Document type: text , language materials , part   Language: eng</p>	
4	<p>▶ <a href="#">Archives in the society : ICA pre-congress meeting, Vienna, 2004 : papers of the international conference, Elblag, May 22-24, 2003 / edited by Wladyslaw Stepniak.</a></p> <p>Colloquia Jerzy Skowronek dedicata (9th : 2003 : Elblag, Poland) / Stepniak, Wladyslaw.</p> <p>Document type: text , language materials , monograph   Language: eng</p>	
5	<p>▶ <a href="#">Research and advanced technology for digital libraries : 9th European conference, ECDL 2005, Vienna, Austria, September 18-23, 2005. proceedings / [edited by] Andreas Rauber, Stavros Christodoulakis, A Min Tjoa.</a></p> <p>ECDL 2005 (2005 : Vienna, Austria) / Rauber, Andreas. / Christodoulakis, Stavros, 1948- / Tjoa, A Min.</p> <p>Document type: text , language materials , monograph   Language: eng</p>	
6	<p>▶ <a href="#">Peer-to-peer, grid, and service-orientation in digital library architectures : 6th thematic workshop of the EU Network of Excellence DELOS , Cagliari, Italy, June 24-25, 2004 : revised selected papers / Can Türker, Maristella Agosti, Hans-Jörg Schek, eds.</a></p> <p>Thematic workshop of the EU Network of Excellence DELOS (6th : 2004 Jun : Cagliari, Italy) / Türker, Can. / Agosti, Maristella. / Schek, H. J.</p> <p>Document type: text , language materials , monograph   Language: eng</p>	
7	<p>▶ <a href="#">User-based evaluation of academic digital libraries : case studies, Social Science Information Gateway, Art, Design Architecture &amp; Media Gateway and the Electronic Journals Service of the University of Patras, Greece / by Maria Monopoli.</a></p> <p>Monopoli, Maria.</p> <p>Document type: text , language materials , monograph   Language: eng</p>	
8	<p>▶ <a href="#">Proceedings of the 5th ACM/IEEE Joint Conference on Digital Libraries : Denver, CO, USA, June 7-11, 2005 : digital libraries, cyberinfrastructure for research and education : / sponsors, ACM SIG on Information Retrieval, ACM SIG on Hypertext, Hypermedia and the Web, IEEE Technical Committee for Digital Libraries.</a></p> <p>ACM/IEEE Joint Conference on Digital Libraries (5th : 2005 : Denver, Colo.)</p> <p>Document type: text , language materials , monograph   Language: eng</p>	
9	<p>▶ <a href="#">Information retrieval design : principles and options for information description, organization, display, and access in information retrieval databases, digital libraries, catalogs, and indexes / by James D. Anderson and José Pérez-Carballo.</a></p> <p>Anderson, James D. (James Doig), 1940- / Pérez-Carballo, José</p> <p>Document type: text , language materials , monograph   Language: eng</p>	
10	<p>▶ <a href="#">Cataloging and organizing digital resources : a how-to-do-it manual for librarians / Anne M. Mitchell and Brian E. Surratt.</a></p> <p>Mitchell, Anne M., 1972- / Surratt, Brian E., 1972-</p> <p>Document type: text , language materials , monograph   Language: eng</p>	

LIBRARY  
HS111RB

search within results (digital libraries)

[▶ Advanced search \(more options\)](#)  
[▶ Change the collections selection](#)

Search

Filter

Author:

Title:

Year:

Free-Text:

Form Query

Reset  use cache Search

Hits: 45 Sort by: Ranked

Filter now...

1 **Barbara Buttenfield**  
Usability **Evaluation of Digital Libraries**.  
(1999) from Achilles.

2 **L. Covi; T. Saracevic**   
Challenges for **digital library evaluation**.  
(2000) from BibDB.

3 **Susan Buhr; Lecia Barker; Thomas C. Reeves**  
Tutorials: **Evaluating digital libraries**  
(2005) from ACM.

4 **Muniram Budhu; Anita Coleman**  
The Design and **Evaluation** of Interactivities in a **Digital Library**.  
(2002) from Achilles.

5 **Elaine Peterson**  
**Evaluation of digital libraries** using snowball sampling.  
(2005) from DBLPIS; DBLP2; Achilles.

6 **Ingrid Hsieh-Yee**  
**Digital Library Evaluation: Progress \& Next Steps**  
(2005) from Achilles.

7 **Judy H. Jeng**  
Usability of **digital libraries**: an **evaluation** model  
(2004) from DBLP2; Achilles.

8 **Pedro Isaias**   
PHIMA-U Model and **Digital Library Evaluation**  
(2002) from Achilles.

Extract:

Metadata G: PDF G: PS G: DOC G: PPT

Personal Lib

Browsing

- fuhr
- DELOS
- DPA
- FIZ Karlsruhe
- INEX
- Information Retrieval
  - PuertoRico
  - Guidos
  - XML IR
  - OWL, DL und Datalog
  - Effects of Annotations on Student Readers and Writers
  - Interactive Information Retrieval
  - Link to Evaluation
  - IR in P2P networks
  - Interaktive IR
  - Text summarization
  - Structural similarity in trees
  - Bertrand -> Henrik
  - DL Evaluation
    - Posters: Evaluating a digital video library web interface
    - Digital library evaluation by analysis of user retrieval patterns.
    - Terms
    - Queries
    - Evaluating Strategic Support for Information Access in the DAFFODIL Syst
    - Component evaluation
    - Evaluation Frameworks
    - Question types in digital reference: an evaluation of question taxonomies
    - Thesauri and ontologies in digital libraries: 1. structure and use in knowle
    - PHIMA-U Model and Digital Library Evaluation
    - Initiatives
    - Surveys
    - Usability
    - <http://www.ils.unc.edu/~march/perseus/lib-trends-final.pdf>
    - A study on the evaluation model for university libraries in digital environr
    - Use of DL Evaluated
    - General
    - How fast is too fast?: evaluating fast forward surrogates for digital video
    - The XML log standard for digital libraries: analysis, evolution, and deployn
    - Journals

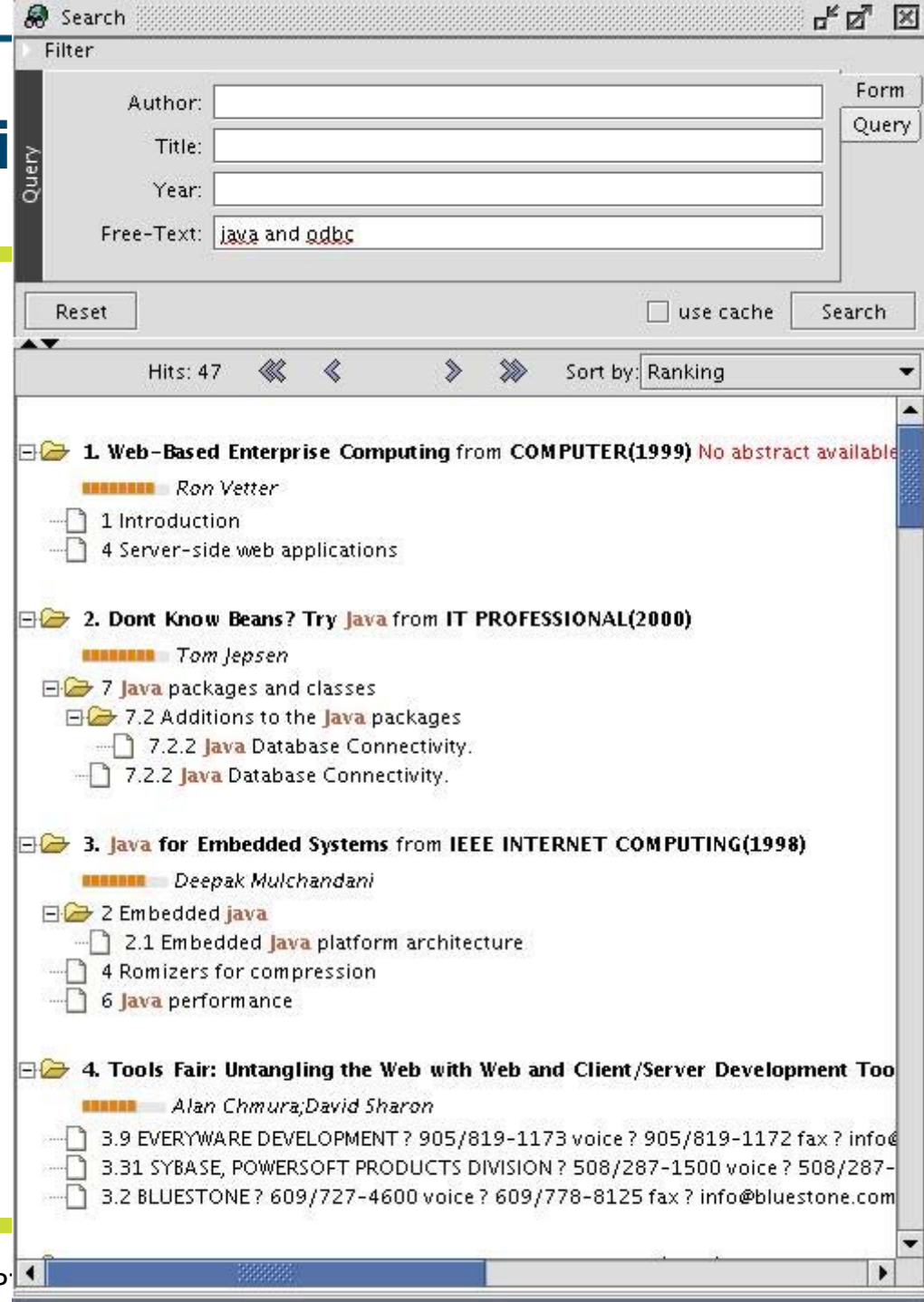


# Comparative Evaluation: TEL standard interface vs. DAFFODIL

## Analytical evaluation of functional similarities & differences

- Usability
  - Search & Browse
  - Display
  - Feedback / Help Functions
- 
- Empirical evaluation
    - How well does each tool support the users' needs?
    - Experts and End-Users

- Daffodil as baseline system for interactive XML retrieval
  - Integration of services and sources into Daffodil
  - Integration of new search tool with special result list handling and display of XML data



## Mobility Management for Cellul...

About the Article

1 Introduction

2 Mobility management

3 Handoff

3.1 INTER-BASE STATION HANDOFF

3.2 INTERSYSTEM HANDOFF

4 Roaming

4.1 THE INTERCONNECTION BETWEEN

4.2 REGISTRATION

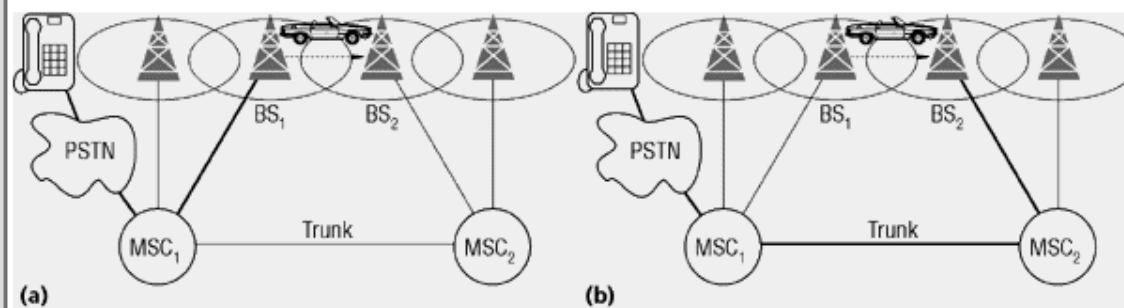
4.3 CALL DELIVERY

5 Conclusion

**INTERSYSTEM HANDOFF**

Our description of the intersystem handoff follows IS-41[2] (GSM follows similar procedures), and we assume network-controlled handoff. Figure 3 illustrates the trunk (voice or data circuit) connection before and after the handoff. A communicating **mobile user** moves out of the base station served by  $MSC_1$  and enters the area covered by  $MSC_2$ . The handoff follows these steps:

- $MSC_1$  requests  $MSC_2$  to perform handoff measurement.  $MSC_2$  then selects a candidate base station,  $BS_2$ , for handoff. That is,  $MSC_2$  finds a base station that covers the **mobile phone** and has a free radio channel to cover the call.  $MSC_2$  returns the signal-quality parameter values and other information to  $MSC_1$ .
- $MSC_1$  checks if the **mobile phone** has made too many handoffs or if intersystem trunks are not available. If so,  $MSC_1$  exits the procedure. Otherwise,  $MSC_1$  asks  $MSC_2$  to set up a voice channel. Suppose that a voice channel is available in  $BS_2$ .  $MSC_2$  asks  $MSC_1$  to start the radio link transfer.
- $MSC_1$  sends the **mobile phone** a handoff order. The **mobile phone** tries to synchronize to  $BS_2$ . After the **mobile phone** connects to  $BS_2$ ,  $MSC_2$  informs  $MSC_1$  that the handoff is successful.  $MSC_1$  then connects the call path (trunk) to  $MSC_2$  and completes the handoff.



**Figure 3:** Before (a) and after (b) an intersystem handoff.

In this process,  $MSC_1$  is called the anchor  $MSC$ , and is always in the call path after the handoff (see Figure 1a). If the **mobile phone** moves back to  $MSC_1$  again, the connection between  $MSC_1$  and  $MSC_2$  terminates.

## Next Activities

- Define standard for logging schema
  - January 2006
- Setup evaluation testbed framework
  - January 2006
- TEL Evaluation

## Next Activities

- Evaluation of cooperation in DL usage (computer science domain)
  - Evaluation of existing tools
    - Personal Library
    - Annotations/Discussion threads
    - Chat tool (text/voice)
    - Whiteboard
  - Study on tools required for collaboration
- Evaluation of Annotation-Service (Joint work with task 4.10)

## WP7 - Conclusion

- Evaluation initiatives:
  - Community building: growing participation, strong support from volunteering groups
  - Foci of international research on XML IR / CLIR
  - New tasks addressing more application-oriented settings
- DL Testbed Framework:
  - Testbed framework: new applications & services
  - Development of logging schema
  - Prototype evaluations: INEX, TEL, Cooperation in DLs