

Carol Peters Martin Braschler
Julio Gonzalo Michael Kluck (Eds.)

Advances in Cross-Language Information Retrieval

Third Workshop of the
Cross-Language Evaluation Forum, CLEF 2002
Rome, Italy, September 19-20, 2002
Revised Papers



Springer

Table of Contents

Introduction	
<i>C. Peters</i>	1

I System Evaluation Experiments at CLEF 2002

CLEF 2002 – Overview of Results	
<i>M. Braschler</i>	9

Cross-Language and More

Cross-Language Retrieval Experiments at CLEF 2002	
<i>A. Chen</i>	28

ITC-irst at CLEF 2002: Using <i>N</i> -Best Query Translations for CLIR	
<i>N. Bertoldi and M. Federico</i>	49

Océ at CLEF 2002	
<i>R. Brand and M. Brünner</i>	59

Report on CLEF 2002 Experiments: Combining Multiple Sources of Evidence	
<i>J. Savoy</i>	66

UTACLIR @ CLEF 2002 – Bilingual and Multilingual Runs with a Unified Process	
<i>E. Airio, H. Keskustalo, T. Hedlund, and A. Pirkola</i>	91

A Multilingual Approach to Multilingual Information Retrieval	
<i>J.-Y. Nie and F. Jin</i>	101

Combining Evidence for Cross-Language Information Retrieval	
<i>J. Kamps, C. Monz, and M. de Rijke</i>	111

Exeter at CLEF 2002: Experiments with Machine Translation for Monolingual and Bilingual Retrieval	
<i>A.M. Lam-Adesina and G.J.F. Jones</i>	127

Portuguese-English Experiments Using Latent Semantic Indexing	
<i>V.M. Orenco and C. Huyck</i>	147

Thomson Legal and Regulatory Experiments for CLEF 2002 <i>I. Moulínier and H. Molina-Salgado</i>	155
Eurospider at CLEF 2002 <i>M. Braschler, A. Göhring, and P. Schäuble</i>	164
Merging Mechanisms in Multilingual Information Retrieval <i>W.-C. Lin and H.-H. Chen</i>	175
SINAI at CLEF 2002: Experiments with Merging Strategies <i>F. Martínez, L.A. Ureña, and M.T. Martín</i>	187
Cross-Language Retrieval at the University of Twente and TNO <i>D. Reidsma, D. Hiemstra, F. de Jong, and W. Kraaij</i>	197
Scalable Multilingual Information Access <i>P. McNamee and J. Mayfield</i>	207
Some Experiments with the Dutch Collection <i>A.P. de Vries and A. Diekema</i>	219
Resolving Translation Ambiguity Using Monolingual Corpora <i>Y. Qu, G. Grefenstette, and D.A. Evans</i>	223
Monolingual Experiments	
Experiments in 8 European Languages with Hummingbird SearchServer™ at CLEF 2002 <i>S. Tomlinson</i>	242
Italian Monolingual Information Retrieval with PROSIT <i>G. Amati, C. Carpineto, and G. Romano</i>	257
COLE Experiments in the CLEF 2002 Spanish Monolingual Track <i>J. Vilares, M.A. Alonso, F.J. Ribadas, and M. Vilares</i>	265
Improving the Automatic Retrieval of Text Documents <i>M. Agosti, M. Bacchin, N. Ferro, and M. Melucci</i>	279
IR-n System at CLEF-2002 <i>F. Llopis, J.L. Vicedo, and A. Ferrández</i>	291
Experiments in Term Expansion Using Thesauri in Spanish <i>Á.F. Zazo, C.G. Figuerola, J.L.A. Berrocal, E. Rodríguez, and R. Gómez</i>	301
SICS at CLEF 2002: Automatic Query Expansion Using Random Indexing <i>M. Sahlgren, J. Karlgren, R. Cöster, and T. Järvinen</i>	311

Pliers and Snowball at CLEF 2002	
<i>A. MacFarlane</i>	321
Experiments with a Chunker and Lucene	
<i>G. Francopoulo</i>	336
Information Retrieval with Language Knowledge	
<i>E. Dura and M. Drejak</i>	338

Mainly Domain-Specific Information Retrieval

Domain Specific Retrieval Experiments with MIMOR at the University of Hildesheim	
<i>R. Hackl, R. Kölle, T. Mandl, and C. Womser-Hacker</i>	343
Using Thesauri in Cross-Language Retrieval of German and French Indexed Collections	
<i>V. Petras, N. Perelman, and F. Gey</i>	349
Assessing Automatically Extracted Bilingual Lexicons for CLIR in Vertical Domains:	
XRCE Participation in the GIRT Track of CLEF 2002	
<i>J.-M. Renders, H. Déjean, and É. Gaussier</i>	363

Interactive Track

The CLEF 2002 Interactive Track	
<i>J. Gonzalo and D.W. Oard</i>	372
SICS at iCLEF 2002:	
Cross-Language Relevance Assessment and Task Context	
<i>J. Karlgren and P. Hansen</i>	383
Universities of Alicante and Jaen at iCLEF	
<i>F. Llopis, J.L. Vicedo, A. Ferrández, M.C. Díaz, and F. Martínez</i>	392
Comparing User-Assisted and Automatic Query Translation	
<i>D. He, J. Wang, D.W. Oard, and M. Nossal</i>	400
Interactive Cross-Language Searching:	
Phrases Are Better than Terms for Query Formulation and Refinement	
<i>F. López-Ostenero, J. Gonzalo, A. Peñas, and F. Verdejo</i>	416
Exploring the Effect of Query Translation when Searching Cross-Language	
<i>D. Petrelli, G. Demetriou, P. Herring, M. Beaulieu, and M. Sanderson</i> ...	430

Cross-Language Spoken Document Retrieval

CLEF 2002 Cross-Language Spoken Document Retrieval
Pilot Track Report
G.J.F. Jones and M. Federico 446

Exeter at CLEF 2002:
Cross-Language Spoken Document Retrieval Experiments
G.J.F. Jones and A.M. Lam-Adesina 458

Cross-Language Spoken Document Retrieval
on the TREC SDR Collection
N. Bertoldi and M. Federico 476

**II Cross-Language Systems Evaluation Initiatives,
Issues and Results**

CLIR at NTCIR Workshop 3: Cross-Language and Cross-Genre Retrieval
N. Kando 485

Linguistic and Statistical Analysis of the CLEF Topics
T. Mandl and C. Womser-Hacker 505

CLEF 2002 Methodology and Metrics
M. Braschler and C. Peters 512

III Appendix

List of Run Characteristics 529

Overview Graphs 533

Multilingual Runs 544

Bilingual to German Runs 580

Bilingual to English Runs 593

Bilingual to Spanish Runs 609

Bilingual to Finnish Runs 625

Bilingual to French Runs 627

Bilingual to Italian Runs 641

Bilingual to Dutch Runs 655

Bilingual to Swedish Runs 664

Monolingual German Runs	665
Monolingual Spanish Runs	666
Monolingual Finnish Runs	714
Monolingual French Runs	725
Monolingual Italian Runs	741
Monolingual Dutch Runs	766
Monolingual Swedish Runs	785
AMARYLLIS Domain-Specific Runs	794
GIRT Domain-Specific Runs	809
Author Index	827