

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

Evaluation of Cross-Language Information Retrieval Systems

Second Workshop of the
Cross-Language Evaluation Forum, CLEF 2001
Darmstadt, Germany, September 3-4, 2001
Revised Papers



Springer

Table of Contents

Introduction	1
<i>Carol Peters</i>	
Part I. System Evaluation Experiments at CLEF 2001	
CLEF 2001 - Overview of Results	9
<i>Martin Braschler</i>	
Mainly Cross-Language	
Report on CLEF-2001 Experiments: Effective Combined Query-Translation Approach	27
<i>Jacques Savoy</i>	
Multilingual Information Retrieval Using English and Chinese Queries	44
<i>Aitao Chen</i>	
Exeter at CLEF 2001: Experiments with Machine Translation for Bilingual Retrieval	59
<i>Gareth J.F. Jones, Adenike M. Lam-Adesina</i>	
TNO at CLEF-2001: Comparing Translation Resources	78
<i>Wessel Kraaij</i>	
ITC-irst at CLEF 2001: Monolingual and Bilingual Tracks	94
<i>Nicola Bertoldi, Marcello Federico</i>	
Experiments with the Eurospider Retrieval System for CLEF 2001	102
<i>Martin Braschler, Bärbel Ripplinger, Peter Schäuble</i>	
Using Co-occurrence, Augmented Restrictions, and C-E WordNet for Chinese-English Cross-Language Information Retrieval at CLEF 2001	111
<i>Wen-Cheng Lin, Hsin-Hsi Chen</i>	
Utaclir @ CLEF 2001 - Effects of Compound Splitting and N-Gram Techniques	118
<i>Turid Hedlund, Heikki Keskustalo, Ari Pirkola, Eija Airio, Kalervo Järvelin</i>	
Using Statistical Translation Models for Bilingual IR	137
<i>Jian-Yun Nie, Michel Simard</i>	
Cross-Lingual Pseudo-Relevance Feedback Using a Comparable Corpus ...	151
<i>Monica Rogati, Yiming Yang</i>	

Investigation on Disambiguation in CLIR: Aligned Corpus and Bi-directional Translation-Based Strategies	158
<i>M. Boughanem, C. Christment, N. Nassr</i>	
Vector-Based Semantic Analysis Using Random Indexing for Cross-Lingual Query Expansion	169
<i>Magnus Sahlgren, Jussi Karlgren</i>	
Query Expansion Techniques for the CLEF Bilingual Track	177
<i>Fatiha Sadat, Akira Maeda, Masatoshi Yoshikawa, Shunsuke Uemura</i>	
Intelligent Information Access Systems (SINAI) at CLEF 2001: Calculating Translation Probabilities with SemCor	185
<i>Fernando Martínez-Santiago, L. Alfonso Ureña-López, M. Carlos Díaz-Galiano, Maite Martín-Valdivia, Manuel García-Vega</i>	
JHU/APL Experiments at CLEF: Translation Resources and Score Normalization	193
<i>Paul McNamee, James Mayfield</i>	
Dictionary-Based Thai CLIR: An Experimental Survey of Thai CLIR.....	209
<i>Jaruskulchai Chuleerat</i>	
English-Dutch CLIR Using Query Translation Techniques	219
<i>Mirna Adriani</i>	
Thomson Legal and Regulatory at CLEF 2001: Monolingual and Bilingual Experiments	226
<i>Hugo Molina-Salgado, Isabelle Moulinier, Mark Knudson, Elizabeth Lund, Kirat Sekhon</i>	
Working with Russian Queries for the GIRT, Bilingual, and Multilingual CLEF Tasks	235
<i>Fredric C. Gey, Hailing Jiang, Natalia Perelman</i>	
IR-n: A Passage Retrieval System at CLEF-2001	244
<i>Fernando Llopis, José L. Vicedo</i>	
Monolingual Experiments	
Spanish Monolingual Track: The Impact of Stemming on Retrieval	253
<i>Carlos G. Figuerola, Raquel Gómez, Angel F. Zazo Rodríguez, José Luis Alonso Berrocal</i>	
Shallow Morphological Analysis in Monolingual Information Retrieval for Dutch, German, and Italian	262
<i>Christof Monz, Maarten de Rijke</i>	

Stemming Evaluated in 6 Languages by Hummingbird SearchServer™ at CLEF 2001	278
<i>Stephen Tomlinson</i>	
Minimalistic Test Runs of the Eidetica Indexer	288
<i>Teresita Frizzarin, Annius Groenink</i>	
Across the Bridge: CLEF 2001 - Non-english Monolingual Retrieval. The French Task	291
<i>Eugenia Matoyo, Tony Valsamidis</i>	
Mpro-IR in CLEF 2001	300
<i>Bärbel Ripplinger</i>	
Some Terms Are More Interchangeable than Others	303
<i>Jakob Klok, Samuel Driessen, Marvin Brunner</i>	
Interactive Track	
The CLEF 2001 Interactive Track	308
<i>Douglas W. Oard, Julio Gonzalo</i>	
Noun Phrase Translations for Cross-Language Document Selection	320
<i>Fernando López-Ostenero, Julio Gonzalo, Anselmo Peñas, Felisa Verdejo</i>	
iCLEF at Sheffield	332
<i>Mark Sanderson, Zoë Bathie</i>	
iCLEF 2001 at Maryland: Comparing Term-for-Term Gloss and MT	336
<i>Jianqiang Wang, Douglas W. Oard</i>	
Part II. Evaluation Issues and Results	
The Philosophy of Information Retrieval Evaluation	355
<i>Ellen M. Voorhees</i>	
CLIR System Evaluation at the Second NTCIR Workshop	371
<i>Noriko Kando</i>	
Multilingual Topic Generation within the CLEF 2001 Experiments	389
<i>Christa Womser-Hacker</i>	
CLEF Methodology and Metrics	394
<i>Martin Braschler, Carol Peters</i>	
Appendix	
List of Run Characteristics	405
Multilingual Runs	409

Bilingual to English Runs	435
Bilingual to Dutch Runs	496
Monolingual Dutch Runs	499
Monolingual French Runs	518
Monolingual German Runs	536
Monolingual Italian Runs	561
Monolingual Spanish Runs	575
GIRT Domain-Specific Runs	597
Author Index	601