

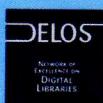
Cross Language Evaluation Forum

ABSTRACTS

CLEF 2007 Workshop, 19-21 September, Budapest, Hungary

Results of the **CLEF 2007**

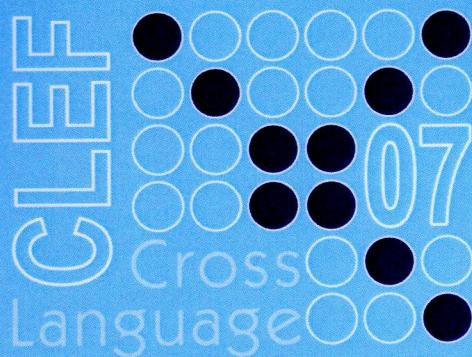
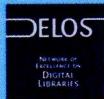
Cross-Language
System
Evaluation
Campaign



ISSN: 1818-8044

ISBN: 2-912335-31-0

Editors: Alessandro Nardi, Carol Peters,



Cross Language Evaluation Forum

About CLEF

The Cross-Language Evaluation Forum (CLEF) supports global digital library applications by (i) developing an infrastructure for the testing, tuning and evaluation of information retrieval systems operating on European languages in both monolingual and cross-language contexts, (ii) providing a forum for the discussion of results, and (iii) creating test-suites of reusable data which can be employed by system developers for benchmarking purposes.

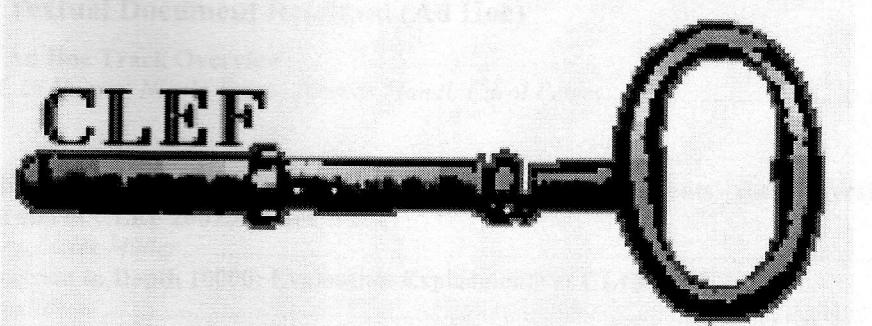
CLEF 2007 is partially supported by the
DELOS Network of Excellence on Digital Libraries
www.delos.info

CLEF Website: www.clef-campaign.org

Coordination:

CLEF is coordinated by the Istituto di Scienza e Tecnologie dell'Informazione, Consiglio Nazionale delle Ricerche, Pisa, Italy. The following Institutions have contributed to the organisation of the different tracks of the CLEF 2007 campaign:

Centre for the Evaluation of Human Language and Multimodal Communication Technologies (CELCT), Trento, Italy - College of Information Studies and Institute for Advanced Computer Studies, University of Maryland, USA - Department of Computer Science, University of Indonesia - Department of Computer Science, RWTH Aachen University, Germany - Department of Computer Science and Information Systems, University of Limerick, Ireland - Department of Computer Science and Information Engineering, National University of Taiwan - Department of Information Engineering, University of Padua, Italy - Department of Information Science, University of Hildesheim, Germany - Department of Information Studies, University of Sheffield, UK - Department of Medical Informatics, Aachen University of Technology (RWTH), Germany - Evaluations and Language Resources Distribution Agency Sarl, Paris, France - Fondazione "Bruno Kessler FBK ist", Trento, Italy - German Research Centre for Artificial Intelligence, DFKI, Saarbrücken, Germany - Information and Language Processing Systems, University of Amsterdam, Netherlands - InformationsZentrum Sozialwissenschaften, Bonn, Germany - Institute for Information Technology, Hyderabad, India - Institute of Formal and Applied Linguistics, Charles University, Czech Rep - Lenguajes y Sistemas Informáticos, Universidad Nacional de Educación a Distancia, Madrid, Spain - Linguateca, Sintef, Oslo, Norway - Linguistic Modelling Laboratory, Bulgarian Academy of Sciences, Bulgaria - Microsoft Research Asia - National Institute of Standards and Technology, Gaithersburg MD, USA - Oregon Health and Science University, USA - Research Computing Center of Moscow State University - Research Institute for Linguistics, Hungarian Academy of Sciences, Hungary - School of Computer Science and Mathematics, Victoria University, Australia - School of Computing, Dublin City University, Ireland - UC Data Archive and School of Information Management and Systems, UC Berkeley, USA - University "Alexandru Ioan Cuza", IASI, Romania - University Hospitals and University of Geneva, Switzerland - Vienna University of Technology, Austria



Cross-Language System Evaluation Campaign

CLEF 2007

Book of Abstracts

Edited by

Alessandro Nardi, Carol Peters

ISTI-CNR, Pisa, Italy

and

Valeria Quochi

ISSN: 1818-8044
ISBN: 2-912335-31-0

L. ALFONSO URQUA López

CONTENTS

What Happened in CLEF 2007	
Carol Peters.....	I
Multilingual Textual Document Retrieval (Ad Hoc)	
CLEF 2007: Ad Hoc Track Overview	
Giorgio M. Di Nunzio, Nicola Ferro, Thomas Mandl, Carol Peters.....	2
Monolingual	
Czech Monolingual Information Retrieval Using Off-the-Shelf Components - the University of West Bohemia at CLEF 2007 Ad-Hoc track	
Pavel Ircing, Luděk Müller.....	4
Sampling Precision to Depth 10000: Evaluation Experiments at CLEF 2007	
Stephen Tomlinson.....	4
Applying Query Expansion Techniques to Ad Hoc Monolingual tasks with the IR-n system	
Elisa Noguera, Fernando Llopis.....	4
Stemming Approaches for East European Languages	
Ljiljana Dolamic, Jacques Savoy.....	5
Hungarian and Czech Stemming using YASS	
Prasenjit Majumder, Mandar Mitra, Dipasree Pal	5
Charles University at CLEF 2007 Ad-Hoc Track	
Pavel Češka, Pavel Pecina	5
Cross-Language	
Ambiguity and Unknown Term Translation in CLIR	
Dong Zhou, Mark Truran, Tim Brailsford.....	6
Evaluating Language Resources for CLEF 2007	
Herika Hayurani, Syandra Sari, Mirna Adriani	6
Amharic-English Information Retrieval with Pseudo Relevance Feedback	
Atelach Alemu Argaw	6
Performing Cross-Language Retrieval with Wikipedia	
Péter Schönhofen, András Benczúr, István Bíró, Károly Csalogány.....	7
Oromo-English Information Retrieval Experiments at CLEF 2007	
Kula Kekeba Tune, Vasudeva Varma	7
Indian Language	
Cross-Lingual Information Retrieval System for Indian Languages	
Jagadeesh Jagarlamudi, A. Kumaran.....	7
Bengali, Hindi and Telugu to English Ad-hoc Bilingual task at CLEF 2007	
Sivaji Bandyopadhyay, Tapabrata Mondal, Sudip Kumar Naskar, Asif Ekbal, Rejwanul Haque, Srinivasa Rao Godavarthy.....	8
IIT Hyderabad at CLEF 2007 - Adhoc Indian Language CLIR task	
Prasad Pingali, Vasudeva Varma.....	8
Hindi and Marathi to English Cross Language Information Retrieval at CLEF 2007	
Manoj Kumar Chinnakotla, Sagar Ranadive, Pushpak Bhattacharyya, Om P. Damani.....	8
Bengali and Hindi to English Cross-language Text Retrieval under Limited Resources	
Debasis Mandal, Sandipan Dandapat, Mayank Gupta, Pratyush Banerjee, Sudeshna Sarkar	9
Robust and More	
Report of MIRACLE team for the Ad-Hoc track in CLEF 2007	
José Miguel Goñi-Menoyo, José C. González-Cristóbal, Julio Villena-Román, Sara Lana-Serrano	9
SINAI at CLEF Ad-Hoc Robust Track 2007: Applying Google Search Engine for Robust Cross-lingual Retrieval	
Fernando Martínez-Santiago, Arturo Montejo-Ráez, Miguel A. García-Cumbreras	9
CoLesIR at CLEF 2007: from English to French via Character N-Grams	
Jesús Vilares, Michael P. Oakes, Manuel Vilares	10
Robust Retrieval Experiments at the University of Hildesheim	
Ben Heuwing, Thomas Mandl.....	10

University of Hagen at QA@CLEF 2007: Coreference Resolution for Questions and Answer Merging	
<i>Sven Hartrumpf, Ingo Glöckner, Johannes Leveling</i>	21
Priberam's Question Answering System in QA@CLEF 2007	
<i>Carlos Amaral, Adán Cassan, Helena Figueira, André Martins, Afonso Mendes, Pedro Mendes, Cláudia Pinto, Daniel Vidal</i>	21
The University of Amsterdam at CLEF@QA 2007	
<i>Valentin Jijkoun, Katja Hofmann, David Ahn, Mahboob Alam Khalid, Joris van Rantwijk, Maarten de Rijke, Erik Tjong, Kim Sang</i>	22
FBK-irst at CLEF 2007	
<i>Milen Kouylekov, Matteo Negri, Bernardo Magnini, Bonaventura Coppola</i>	22
Finding Answers Using Resources in the Internet	
<i>Septian Adiwibowo, Mirna Adriani</i>	22
Making RAPOSA (FOX) smarter	
<i>Luís Sarmento, Eugénio Oliveira</i>	22
Answer Validation Exercise	
UNED at Answer Validation Exercise 2007	
<i>Álvaro Rodrigo, Anselmo Peñas, Felisa Verdejo</i>	23
DFKI-LT at AVE 2007: Using Recognizing Textual Entailment for Answer Validation	
<i>Rui Wang, Günter Neumann</i>	23
SINAI at QA@CLEF 2007. Answer Validation Exercise	
<i>M.A. García-Cumbreras, J. M. Perea-Ortega, F. Martínez Santiago, L.A. Ureña-López</i>	23
The Contribution of the University of Alicante to AVE 2007	
<i>Óscar Ferrández, Daniel Micol, Rafael Muñoz, Manuel Palomar</i>	24
INAOE at AVE 2007: Experiments in Spanish Answer Validation	
<i>Alberto Téllez-Valero, Manuel Montes-y-Gómez, Luis Villaseñor-Pineda</i>	24
University of Hagen at CLEF 2007: Answer Validation Exercise	
<i>Ingo Glöckner</i>	24
Question Answering on Speech Transcription (QAST)	
Robust Question Answering for Speech Transcripts Using Minimal Syntactic Analysis	
<i>Pere R. Comas, Jordi Turmo, Mihai Surdeanu</i>	25
The LIMSI Participation to the QAst track	
<i>Sophie Rosset, Olivier Galibert, Gilles Adda, Eric Bilinski</i>	25
AnswerFinder at QAst 2007: Named Entity Recognition for QA on Speech Transcripts	
<i>Diego Mollá, Steve Cassidy, Menno van Zaanen</i>	25
CLEF2007 Question Answering Experiments at Tokyo Institute of Technology	
<i>E.W.D. Whittaker, J.R. Novak, M. Heie, S. Furui</i>	26
DFKI-LT at QAST 2007: Adapting QA Components to Mine Answers in Speech Transcripts	
<i>Günter Neumann, Rui Wang</i>	26
Cross-Language Retrieval in Image Collections	
Overview of the ImageCLEFphoto 2007 Photographic Retrieval Task	
<i>Michael Grubinger, Paul Clough, Allan Hanbury, Henning Müller</i>	27
Overview of the ImageCLEF 2007 Object Retrieval Task	
<i>Thomas Deselaers, Allan Hanbury, Ville Viitaniemi, András Benczúr, Mátyás Brendel, Bálint Daróczy, Hugo Jair Escalante Balderas, Theo Gevers, Carlos Arturo Hernández Gracidas, Steven C. H. Hoi, Jorma Laaksonen, Mingjing Li, Heidy Marisol Marin Castro, Hermann Ney, Xiaoguang Rui, Nicu Sebe, Julian Stöttinger, Lei Wu</i>	28
Overview of the ImageCLEF 2007 Medical Retrieval and Annotation Tasks	
<i>Henning Müller, Thomas Deselaers, Eugene Kim, Jayashree Kalpathy-Cramer, Thomas M. Deserno, William Hersh</i>	29
ImageCLEFphoto	
Linked Relevance Feedback for the ImageCLEF Photo Task	
<i>Ray R. Larson</i>	29
Exploring Image, Text and Geographic Evidences in ImageCLEF 2007	
<i>João Magalhães, Simon Overell, Stefan Rüger</i>	30

MIRACLE at ImageCLEFphoto 2007: Evaluation of Merging Strategies for Multilingual and Multimedia Information Retrieval	
<i>Julio Villena-Román, Sara Lana-Serrano, José Luis Martínez-Fernández, José Carlos González-Cristóbal</i>	30
Cross-Language and Cross-Media Image Retrieval: An Empirical Study at ImageCLEF2007	
<i>Steven C. H. Hoi</i>	31
XRCE's Participation to ImageCLEFphoto 2007	
<i>Stephane Clinchant, Jean-Michel Renders, Gabriela Csurka</i>	31
Experiment for Using Web Information to do Query and Document Expansion	
<i>Yih-Chen Chang, Hsin-Hsi Chen</i>	32
Experiments with Clustering the Collection at ImageCLEF 2007	
<i>Osama El Demerdash, Leila Kosseim, Sabine Bergler</i>	32
TIA-INAOE's Participation at ImageCLEF 2007	
<i>H. Jair Escalante, Carlos A. Hernández, Aurelio López, Heidy M. Marín, Manuel Montes, Eduardo Morales, Luis E. Sucar, Luis Villaseñor</i>	32
IPAL at ImageClef 2007 Mixing Features, Models and Knowledge.	
<i>Sheng Gao, Jean-Pierre Chevallet, Thi Hoang Diem Le, Trong Ton Pham, Joo Hwee Lim</i>	33
DCU and UTA at ImageCLEFPhoto 2007	
<i>Anni Järvelin, Peter Wilkins, Tomasz Adamek, Eija Airio, Gareth J. F. Jones, Alan F. Smeaton, Eero Sormunen</i>	33
Information Retrieval of Visual Descriptions with IR-n System Based on Passages	
<i>Sergio Navarro, Fernando Llopis, Rafael Muñoz Guillena, Elisa Noguera</i>	34
Multi-Modal Interactive Approach to ImageCLEF 2007 Photographic and Medical Retrieval Tasks by CINDI	
<i>M. M. Rahman, Bipin C. Desai, Prabir Bhattacharya</i>	34
Cross-Modal Retrieval by Text and Image Feature Bioclustering	
<i>András Benczúr, István Bíró, Mátyás Brendel, Károly Csalogány, Bálint Daróczy, Dávid Siklósi</i>	34
Using Pseudo-Relevance Feedback to Improve Image Retrieval Results	
<i>Mouna Torjmen, Karen Pinel-Sauvagnat, Mohand Boughanem</i>	35
Content-based Image Retrieval Using Shape-Size Pattern Spectra	
<i>Florence Tushabe, Michael. H. F. Wilkinson</i>	35
Experiments for the ImageCLEF 2007 Photographic Retrieval Task	
<i>Thomas Wilhelm, Jens Kürsten, Maximilian Eibl</i>	35
ImageCLEFmed	
Baseline Results for the CLEF 2007 Medical Automatic Annotation Task	
<i>Mark O. Güld, Thomas M. Deserno</i>	36
CLEF2007 Image Annotation Task: an SVM-based Cue Integration Approach	
<i>Tatiana Tommasi, Francesco Orabona, Barbara Caputo</i>	36
CYU_IM@ImageCLEF 2007: Medical image annotation task	
<i>Pei-Cheng Cheng, Wei-Pang Yang</i>	37
Multiplying Concept Sources for Graph Modeling	
<i>Loïc Maisonnasse, Eric Gaussier, Jean-Pierre Chevallet</i>	37
MIRACLE at ImageCLEFmed 2007: Merging Textual and Visual Strategies to Improve Medical Image Retrieval	
<i>Julio Villena-Román, Sara Lana-Serrano, José Carlos González-Cristóbal</i>	38
MIRACLE at ImageCLEFanot 2007: Machine Learning Experiments on Medical Image Annotation	
<i>Sara Lana-Serrano, Julio Villena-Román, José Carlos González-Cristóbal, José Miguel Goñi-Menoyo</i>	38
Medical Image Retrieval and Automatic Annotation: OHSU at ImageCLEF 2007	
<i>Jayashree Kalpathy-Cramer, William Hersh</i>	39
Speeding up IDM without Degradation of Retrieval Quality	
<i>Michael Springmann, Heiko Schuldt</i>	39
ImageCLEFphoto&med	
SINAI at ImageCLEF 2007	
<i>M.C. Díaz-Galiano, M.A. García-Cumbreras, M.T. Martín-Valdivia, A. Montejo-Raez, L.A. Ureña-López</i>	40
University and Hospitals of Geneva at ImageCLEF 2007	
<i>Xin Zhou, Julien Gobeill, Patrick Ruch, Henning Müller</i>	40

FIRE in ImageCLEF 2007

Thomas Deselaers, Tobias Gass, Tobias Weyand, Hermann Ney 40

Cross-Language Speech Retrieval

Overview of the CLEF-2007 Cross-Language Speech Retrieval

Pavel Pecina, Petra Hoffmannová, Gareth Jones, Ying Zhang, Douglas W. Oard 41

SINAI at CL-SR task at CLEF 2007

M.C. Díaz-Galiano, M.T. Martín-Valdivia, M.A. García-Cumbreras, L.A. Ureña-López 44

Model Fusion Experiments for the Cross Language Speech Retrieval Task at CLEF 2007

Muath Alzghool, Diana Inkpen 44

Attempts to Search Czech Spontaneous Spoken Interviews - the University of West Bohemia at CLEF 2007 CL-SR track

Pavel Ircing, Luděk Müller 44

University of Chicago at the CLEF 2007 Cross-language Speech Retrieval Track

Gina-Anne Levow 45

Dublin City University at CLEF 2007: Cross-Language Speech Retrieval (CL-SR) Experiments

Ying Zhang, Gareth J. F. Jones, Ke Zhang 45

Brown at CL-SR'07: Retrieving Conversational Speech in English and Czech

Matthew Lease, Eugene Charniak 45

Charles University at CLEF 2007 CL-SR Track

Pavel Češka, Pavel Pecina 46

The University of Amsterdam at the CLEF Cross Language Speech Retrieval Track 2007

Bouke Huurnink 46

Multilingual Web Track

Overview of WebCLEF 2007

Valentin Jijkoun, Maarten de Rijke 47

REINA at WebCLEF 2007. Selecting Useful Snippets

Carlos G. Figuerola, José L. Alonso Berrocal, Ángel F. Zazo Rodríguez, Emilio Rodríguez 47

The University of Amsterdam at WebCLEF 2007: Using Centrality to Rank Web Snippets

Valentin Jijkoun, Maarten de Rijke 47

Cross-Language Geographical Retrieval

GeoCLEF 2007: the CLEF 2007 Cross-Language Geographic Information Retrieval Track

Overview

Thomas Mandl, Fredric Gey, Giorgio Di Nunzio, Nicola Ferro, Ray Larson, Mark Sanderson, Diana Santos, Christa Womser-Hacker, Xing Xie 49

Query Parsing Task for GeoCLEF2007 Report

Zhisheng Li, Chong Wang, Xing Xie, Wei-Ying Ma 50

University of Groningen at GeoCLEF 2007

Geoffrey Andogah, Gosse Bouma 50

MSRA Columbus at GeoCLEF2007

Zhisheng Li, Chong Wang, Xing Xie, Wei-Ying Ma 50

GIR Experiments with Forostar at GeoCLEF 2007

Simon Overell, João Magalhães, Stefan Rüger 51

MIRACLE at GeoCLEF Query Parsing 2007: Extraction and Classification of Geographical Information

Sara Lana-Serrano, Julio Villena-Román, José Miguel Goñi-Menoyo 51

Cheshire at GeoCLEF 2007: Retesting Text Retrieval Baselines

Ray R. Larson 51

GEOUJA System. University of Jaén at GEOCLEF 2007

José M. Perea-Ortega, Miguel A. García-Cumbreras, Manuel García-Vega, Arturo Montejo-Ráez 52

Monolingual Retrieval Experiments with Spatial Restrictions at GeoCLEF 2007

Ralph Kölle, Ben Heuwing, Thomas Mandl, Christa Womser-Hacker 52

GeoCLEF2007 Experiments in Query Parsing and Cross-Language GIR

Rocío Guillén 52

The UPV at GeoCLEF 2007

Davide Buscaldi, Paolo Rosso 53

TALP at GeoCLEF 2007: Using Terrier with Geographical Knowledge Filtering

Daniel Ferrés, Horacio Rodríguez 53