

Commenced Publication in 1973

Founding and Former Series Editors:

Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board

David Hutchison

Lancaster University, UK

Takeo Kanade

Carnegie Mellon University, Pittsburgh, PA, USA

Josef Kittler

University of Surrey, Guildford, UK

Jon M. Kleinberg

Cornell University, Ithaca, NY, USA

Alfred Kobsa

University of California, Irvine, CA, USA

Friedemann Mattern

ETH Zurich, Switzerland

John C. Mitchell

Stanford University, CA, USA

Moni Naor

Weizmann Institute of Science, Rehovot, Israel

Oscar Nierstrasz

University of Bern, Switzerland

C. Pandu Rangan

Indian Institute of Technology, Madras, India

Bernhard Steffen

University of Dortmund, Germany

Madhu Sudan

Microsoft Research, Cambridge, MA, USA

Demetri Terzopoulos

University of California, Los Angeles, CA, USA

Doug Tygar

University of California, Berkeley, CA, USA

Gerhard Weikum

Max-Planck Institute of Computer Science, Saarbruecken, Germany

Carol Peters Thomas Deselaers
Nicola Ferro Julio Gonzalo
Gareth J.F. Jones Mikko Kurimo
Thomas Mandl Anselmo Peñas
Vivien Petras (Eds.)

Evaluating Systems for Multilingual and Multimodal Information Access

9th Workshop of the Cross-Language Evaluation Forum,
CLEF 2008
Aarhus, Denmark, September 17-19, 2008
Revised Selected Papers

Volume Editors

Carol Peters
ISTI, CNR, Pisa, Italy; carol.peters@isti.cnr.it

Thomas Deselaers
RWTH Aachen University, Germany; deselaers@informatik.rwth-aachen.de

Nicola Ferro
University of Padua, Italy; nicola.ferro@unipd.it

Julio Gonzalo
Anselmo Peñas
LSI-UNED, Madrid, Spain; {julio,anselmo}@lsi.uned.es

Gareth J.F. Jones
Dublin City University, Ireland; gareth.jones@computing.dcu.ie

Mikko Kurimo
Helsinki University of Technology, Finland; mikko.kurimo@tkk.fi

Thomas Mandl
University of Hildesheim, Germany; mandl@uni-hildesheim.de

Vivien Petras
Humboldt University Berlin, Germany; vivien.petras@ibi.hu-berlin.de

Managing Editor
Danilo Giampiccolo, CELCT, Trento, Italy; giampiccolo@celct.it

Library of Congress Control Number: 2009934437

CR Subject Classification (1998): I.2.7, H.2.8, I.7, H.4, H.5, H.5.2, I.1.3

LNCS Sublibrary: SL 3 – Information Systems and Application, incl. Internet/Web and HCI

ISSN 0302-9743
ISBN-10 3-642-04446-8 Springer Berlin Heidelberg New York
ISBN-13 978-3-642-04446-5 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

springer.com

© Springer-Verlag Berlin Heidelberg 2009
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 12753536 06/3180 5 4 3 2 1 0

Preface

The ninth campaign of the Cross-Language Evaluation Forum (CLEF) for European languages was held from January to September 2008. There were seven main evaluation tracks in CLEF 2008 plus two pilot tasks. The aim, as usual, was to test the performance of a wide range of multilingual information access (MLIA) systems or system components. This year, 100 groups, mainly but not only from academia, participated in the campaign. Most of the groups were from Europe but there was also a good contingent from North America and Asia plus a few participants from South America and Africa. Full details regarding the design of the tracks, the methodologies used for evaluation, and the results obtained by the participants can be found in the different sections of these proceedings.

The results of the CLEF 2008 campaign were presented at a two-and-a-half day workshop held in Aarhus, Denmark, September 17–19, and attended by 150 researchers and system developers. The annual workshop, held in conjunction with the European Conference on Digital Libraries, plays an important role by providing the opportunity for all the groups that have participated in the evaluation campaign to get together comparing approaches and exchanging ideas.

The schedule of the workshop was divided between plenary track overviews, and parallel, poster and breakout sessions presenting this year's experiments and discussing ideas for the future. There were several invited talks. Noriko Kando, National Institute of Informatics Tokyo, reported on the activities of NTCIR-7 (NTCIR is an evaluation initiative focussed on testing IR systems for Asian languages), while John Tait of the Information Retrieval Facility, Vienna, presented a proposal for an Intellectual Property track which would focus on cross-language retrieval of legal patents in CLEF 2009. In the final session, Donna Harman, US National Institute of Standards and Technology, presented her impressions of the main trends emerging from the 2008 workshop and campaign, and Martin Braschler of Zurich University of Applied Sciences gave a talk describing a survey he had made on the search functionality of enterprise websites. The presentations given at the CLEF workshop can be found on the CLEF website at www.clef-campaign.org.

The workshop was preceded by two related events. On September 16, the ImageCLEF group, with the sponsorship of the Quaero program (www.quaero.org), organized a one-day workshop on Multimedia Information Retrieval Evaluation. The workshop included presentations of the activities of both Quaero and Theseus, two international projects working on the development of next-generation Internet search engines. The Morpho Challenge 2008 meeting on "Unsupervised Morpheme Analysis" was held on the morning of September 17. Morpho Challenge 2008 was part of the EU Network of Excellence PASCAL Programme and was run in collaboration with CLEF.

The CLEF 2008 and 2009 campaigns were organized as activities of TrebleCLEF, a Coordination Action of the Seventh Framework Programme. TrebleCLEF is building on and extending the results achieved by CLEF. The objective is to support the development and consolidation of expertise in the multidisciplinary research area of

multilingual information access and to promote a dissemination action in the relevant application communities. TrebleCLEF is also attempting to promote more user-and usage-focused investigations within CLEF.

At the time of writing the organization of CLEF 2009 is well underway. In line with the TrebleCLEF philosophy, the campaign this year includes three new tracks focused on analyzing user behavior in a multilingual context (LogCLEF), on studying the requirements of multilingual patent search (CLEF-IP), and on improving our understanding of MLIA systems and their behavior with respect to languages (GridCLEF).

These post-campaign proceedings represent extended and revised versions of the initial working notes distributed at the workshop. All papers were subjected to a reviewing procedure. The final volume was prepared with the assistance of the Center for the Evaluation of Language and Communication Technologies (CELCT), Trento, Italy, under the coordination of Danilo Giampiccolo. The support of CELCT is gratefully acknowledged. We should also like to thank all our reviewers for their careful refereeing.

May 2009

Carol Peters
Thomas Deselaers
Nicola Ferro
Julio Gonzalo
Gareth J. F. Jones
Mikko Kurimo
Thomas Mandl
Anselmo Peñas
Vivien Petras

Reviewers

The Editors express their gratitude to the colleagues listed below for their assistance in reviewing the papers in this volume:

- Eneko Agirre, University of the Basque Country, Spain
- Abolfazl AleAhmad, University of Tehran, Iran
- Hadi Amiri, University of Tehran, Iran
- Ebru Arisoy, Bogazici University, Turkey
- Stefan Baerisch, GESIS Leibniz-Institut for Social Sciences, Bonn, Germany
- Delphine Bernhard, Darmstadt University of Technology, Germany
- Johan Bos, University of Rome "La Sapienza", Italy
- Burcu Can, University of York, UK
- Nuno Cardoso, University of Lisbon, Portugal
- Paula Carvalho, Linguatca and University of Lisbon, Portugal
- Leda Casanova, CELCT, Italy
- Tolga Ciloglu, Middle East Technical University, Turkey
- Paul D. Clough, University of Sheffield, UK
- Luis F. Costa, SINTEF ICT, Portugal
- Thomas M. Deserno, RWTH Aachen University, Germany
- Giorgio Di Nunzio, University of Padua, Italy
- Corina Forascu, Institute for Research in Artificial Intelligence, Romania
- Miguel Garcia-Cumbreras, University of Jaen, Spain
- Fredric C. Gey, University of California at Berkeley, USA
- Ingo Glöckner, FernUniversität in Hagen, Germany
- Harald Hammarström, Chalmers University, Sweden
- Allan Hanbury, Technical University of Vienna, Austria
- Donna Harman, National Institute of Standards and Technology, USA
- Sven Hartrumpf, FernUniversität in Hagen, Germany
- Jesús Herrera, Universidad Complutense de Madrid, Spain
- William Hersh, Oregon Health and Science University, Portland, USA
- Jayashree Kalpathy-Cramer, Oregon Health and Science University, USA
- Chunyu Kit, Hong Kong City University, China
- Dietrich Klakow, University of Saarland, Germany
- Jana Kludas, University of Geneva, Switzerland
- Zornitsa Kozareva, USC Information Sciences Institute, USA
- Martha Larson, Delft University of Technology, The Netherlands
- Ray Larson, University of California at Berkeley, USA
- Johannes Leveling, FernUniversität in Hagen, Germany
- Patricio Martínez, University of Alicante, Spain
- Paul McNamee, Johns Hopkins University, USA

- Henning Müller, University of Applied Sciences Western Switzerland, Sierre and University of Geneva, Switzerland
- Diego Molla, Macquarie University, Australia
- Manuel Montes, INAOE, Mexico
- Günter Neumann, German Research Centre for Artificial Intelligence, Germany
- Eamonn Newman, Dublin City University, Ireland
- Petya Osenova, Bulgarian Academy of Sciences, Bulgaria
- Simon Overell, Imperial College London, UK
- Alvaro Rodrigo, UNED, Madrid, Spain
- Paolo Rosso, Polytechnic University of Valencia, Spain
- Andrew Salway, Dublin City University, Ireland
- Mark Sanderson, University of Sheffield, UK
- Diana Santos, Linguatca and SINTEF ICT, Norway
- Murat Saraclar, Bogazici University, Turkey
- Jacques Savoy, University of Neuchâtel, Switzerland
- Gianmaria Silvello, University of Padua, Italy
- Theodora Tsikrika, CWI, Amsterdam, The Netherlands
- Jordi Turmo, Polytechnic of Catalonia, Spain
- Christa Womser-Hacker, University of Hildesheim, Germany
- Fabio Massimo Zanzotto, University of Rome “Tor Vergata”, Italy

CLEF 2008 Coordination

CLEF is coordinated by the Istituto di Scienza e Tecnologie dell'Informazione, Consiglio Nazionale delle Ricerche, Pisa. The following institutions contributed to the organization of the different tracks of the CLEF 2008 campaign:

- Adaptive Informatics Research Centre, Helsinki University of Technology, Finland
- Athena Research Center, Athens, Greece
- Business Information Systems, Univ. of Applied Sciences Western Switzerland, Sierre, Switzerland
- Centre for the Evaluation of Human Language and Multimodal Communication Technologies (CELCT), Trento, Italy
- Centrum voor Wiskunde en Informatica, Amsterdam, The Netherlands
- Computer Science Department, University of the Basque Country, Spain
- Computer Vision and Multimedia Lab, University of Geneva, Switzerland
- Database Research Group, University of Tehran, Iran
- Department of Computer Science, Aachen University of Technology, Germany
- Department of Computer Science and Information Systems, University of Limerick, Ireland
- 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 and Clinical Epidemiology, Oregon Health and Science University, USA
- Department of Medical Informatics, Aachen University of Technology, Germany
- Department of Medical Informatics, University Hospitals and University of Geneva, Switzerland
- Evaluations and Language Resources Distribution Agency Sarl, Paris, France
- German Research Centre for Artificial Intelligence, Saarbrücken, Germany
- GESIS Leibniz-Institut for the Social Sciences, Bonn, Germany
- Information Science, University of Groningen, The Netherlands
- Institute of Computer Aided Automation, Vienna University of Technology, Austria
- Intelligent Systems Lab Amsterdam, University of Amsterdam, The Netherlands
- Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur (LIMSI), Orsay, France

- Lenguajes y Sistemas Informáticos, Universidad Nacional de Educación a Distancia, Madrid, Spain
- Linguateca, Sintef, Oslo, Norway
- Linguateca, CISUC, Department of Information Engineering, University of Coimbra, Portugal
- Linguateca, XLDB, Department of Information Engineering, University of Lisbon, Portugal
- Linguistic Modelling Laboratory, Bulgarian Academy of Sciences, Bulgaria
- Microsoft Research Asia
- National Institute of Standards and Technology, Gaithersburg, USA
- Research Computing Center of Moscow State University, Russia
- Romanian Institute for Computer Science, Romania
- School of Computing, Dublin City University, Ireland
- School of Computer Science and Mathematics, Victoria University, Australia
- TALP Research Center, Universitat Politècnica de Catalunya, Barcelona, Spain
- UC Data Archive and School of Information Management and Systems, UC Berkeley, USA

CLEF 2008 Steering Committee

- Maristella Agosti, University of Padua, Italy
- Martin Braschler, Zurich University of Applied Sciences, Switzerland
- Amedeo Cappelli, ISTI-CNR and CELCT, Italy
- Hsin-Hsi Chen, National Taiwan University, Taipei, Taiwan
- Khalid Choukri, Evaluations and Language Resources Distribution Agency, Paris, France
- Paul Clough, University of Sheffield, UK
- Thomas Deselaers, Aachen University of Technology, Germany
- Giorgio Di Nunzio, University of Padua, Italy
- David A. Evans, Clairvoyance Corporation, USA
- Marcello Federico, Fondazione Bruno Kessler, Trento, Italy
- Nicola Ferro, University of Padua, Italy
- Christian Fluhr, CEA-LIST, Fontenay-aux-Roses, France
- Norbert Fuhr, University of Duisburg, Germany
- Frederic C. Gey, U.C. Berkeley, USA
- Julio Gonzalo, LSI-UNED, Madrid, Spain
- Donna Harman, National Institute of Standards and Technology, USA
- Gareth Jones, Dublin City University, Ireland
- Franciska de Jong, University of Twente, The Netherlands
- Noriko Kando, National Institute of Informatics, Tokyo, Japan
- Jussi Karlgren, Swedish Institute of Computer Science, Sweden
- Michael Kluck, German Institute for International and Security Affairs, Berlin, Germany
- Natalia Loukachevitch, Moscow State University, Russia
- Bernardo Magnini, Fondazione Bruno Kessler, Trento, Italy
- Paul McNamee, Johns Hopkins University, USA
- Henning Müller, University of Applied Sciences Western Switzerland, Sierre and University of Geneva, Switzerland
- Douglas W. Oard, University of Maryland, USA
- Anselmo Peñas, LSI-UNED, Madrid, Spain
- Vivien Petras, GESIS Leibniz Institute for the Social Sciences, Bonn, Germany
- Maarten de Rijke, University of Amsterdam, The Netherlands
- Diana Santos, Linguatca, Sintef, Oslo, Norway
- Jacques Savoy, University of Neuchâtel, Switzerland
- Peter Schäuble, Eurospider Information Technologies, Switzerland
- Richard Sutcliffe, University of Limerick, Ireland

- Hans Uszkoreit, German Research Center for Artificial Intelligence, Germany
- Felisa Verdejo, LSI-UNED, Madrid, Spain
- José Luis Vicedo, University of Alicante, Spain
- Ellen Voorhees, National Institute of Standards and Technology, USA
- Christa Womser-Hacker, University of Hildesheim, Germany

Table of Contents

What Happened in CLEF 2008	1
<i>Carol Peters</i>	
Part I: Multilingual Textual Document Retrieval (Ad Hoc)	
CLEF 2008: Ad Hoc Track Overview	15
<i>Eneko Agirre, Giorgio Maria Di Nunzio, Nicola Ferro, Thomas Mandl, and Carol Peters</i>	
TEL@CLEF	
Logistic Regression for Metadata: Cheshire Takes on Adhoc-TEL	38
<i>Ray R. Larson</i>	
Query Expansion via Library Classification System	42
<i>Alessio Bosca and Luca Dini</i>	
Experiments on a Multinomial Language Model versus Lucene's Off-the-Shelf Ranking Scheme and Rocchio Query Expansion (TEL@CLEF Monolingual Task)	50
<i>Jorge Machado, Bruno Martins, and José Borbinha</i>	
WikiTranslate: Query Translation for Cross-Lingual Information Retrieval Using Only Wikipedia	58
<i>Dong Nguyen, Arnold Overwijk, Claudia Hauff, Dolf R.B. Trieschnigg, Djoerd Hiemstra, and Franciska de Jong</i>	
UFRGS@CLEF2008: Using Association Rules for Cross-Language Information Retrieval	66
<i>André Pinto Geraldo and Viviane P. Moreira</i>	
CLEF 2008 Ad-Hoc Track: Comparing and Combining Different IR Approaches	75
<i>Jens Kürsten, Thomas Wilhelm, and Maximilian Eibl</i>	
Multi-language Models and Meta-dictionary Adaptation for Accessing Multilingual Digital Libraries	83
<i>Stephane Clinchant and Jean-Michel Renders</i>	

Persian@CLEF

Improving Persian Information Retrieval Systems Using Stemming and Part of Speech Tagging 89
Reza Karimpour, Amineh Ghorbani, Azadeh Pishdad, Mitra Mohtarami, Abolfazl AleAhmad, Hadi Amiri, and Farhad Oroumchian

Fusion of Retrieval Models at CLEF 2008 Ad Hoc Persian Track 97
Zahra Aghazade, Nazanin Dehghani, Leili Farzinvasht, Razieh Rahimi, Abolfazl AleAhmad, Hadi Amiri, and Farhad Oroumchian

Cross Language Experiments at Persian@CLEF 2008 105
Abolfazl AleAhmad, Ehsan Kamalloo, Arash Zareh, Masoud Rahgozar, and Farhad Oroumchian

Robust-WSD

Evaluating Word Sense Disambiguation Tools for Information Retrieval Task 113
Fernando Martínez-Santiago, José M. Perea-Ortega, and Miguel A. García-Cumbreras

IXA at CLEF 2008 Robust-WSD Task: Using Word Sense Disambiguation for (Cross Lingual) Information Retrieval 118
Eneko Agirre, Arantxa Otegi, and German Rigau

SENSE: SEMantic N-levels Search Engine at CLEF2008 Ad Hoc Robust-WSD Track 126
Annalina Caputo, Pierpaolo Basile, and Giovanni Semeraro

IR-n in the CLEF Robust WSD Task 2008 134
Sergio Navarro, Fernando Llopis, and Rafael Muñoz

Query Clauses and Term Independence 138
José R. Pérez-Agüera and Hugo Zaragoza

Analysis of Word Sense Disambiguation-Based Information Retrieval ... 146
Jacques Guyot, Gilles Falquet, Saïd Radhouani, and Karim Benzineb

Crosslanguage Retrieval Based on Wikipedia Statistics..... 155
Andreas Juffinger, Roman Kern, and Michael Granitzer

Ad Hoc Mixed: TEL and Persian

Sampling Precision to Depth 10000 at CLEF 2008 163
Stephen Tomlinson

JHU Ad Hoc Experiments at CLEF 2008	170
<i>Paul McNamee</i>	

UniNE at CLEF 2008: TEL, and Persian IR	178
<i>Ljiljana Dolamic, Claire Fautsch, and Jacques Savoy</i>	

Part II: Mono- and Cross-Language Scientific Data Retrieval (Domain-Specific)

The Domain-Specific Track at CLEF 2008	186
<i>Vivien Petras and Stefan Baerisch</i>	

UniNE at Domain-Specific IR - CLEF 2008	199
<i>Claire Fautsch, Ljiljana Dolamic, and Jacques Savoy</i>	

Back to Basics – Again – for Domain-Specific Retrieval	203
<i>Ray R. Larson</i>	

Concept Models for Domain-Specific Search	207
<i>Edgar Meij and Maarten de Rijke</i>	

The Xtrieval Framework at CLEF 2008: Domain-Specific Track	215
<i>Jens Kürsten, Thomas Wilhelm, and Maximilian Eibl</i>	

Using Wikipedia and Wiktionary in Domain-Specific Information Retrieval	219
<i>Christof Müller and Iryna Gurevych</i>	

Part III: Interactive Cross-Language Retrieval (iCLEF)

Overview of iCLEF 2008: Search Log Analysis for Multilingual Image Retrieval	227
<i>Julio Gonzalo, Paul Clough, and Jussi Karlgren</i>	

Log Analysis of Multilingual Image Searches in Flickr	236
<i>Victor Peinado, Julio Gonzalo, Javier Artiles, and Fernando López-Ostenero</i>	

Cross-Lingual Image Retrieval Interactions Based on a Game Competition	243
<i>Giorgio Maria Di Nunzio</i>	

A Study of Users' Image Seeking Behaviour in FlickLing	251
<i>Evgenia Vassilakaki, Frances Johnson, Richard J. Hartley, and David Randall</i>	

SICS at iCLEF 2008: User Confidence and Satisfaction Tentatively Inferred from iCLEF Logs.....	260
<i>Jussi Karlgren</i>	

Part IV: Multiple Language Question Answering (QA@CLEF)

Overview of the Clef 2008 Multilingual Question Answering Track.....	262
<i>Pamela Forner, Anselmo Peñas, Eneko Agirre, Iñaki Alegria, Corina Forăscu, Nicolas Moreau, Petya Osenova, Prokopis Prokopidis, Paulo Rocha, Bogdan Sacaleanu, Richard Sutcliffe, and Erik Tjong Kim Sang</i>	
Overview of the Answer Validation Exercise 2008	296
<i>Álvaro Rodrigo, Anselmo Peñas, and Felisa Verdejo</i>	
Overview of QAST 2008	314
<i>Jordi Turmo, Pere R. Comas, Sophie Rosset, Lori Lamel, Nicolas Moreau, and Djamel Mostefa</i>	

Mono and Bilingual QA

Assessing the Impact of Thesaurus-Based Expansion Techniques in QA-Centric IR.....	325
<i>Luís Sarmiento, Jorge Teixeira, and Eugénio Oliveira</i>	
Using AliQAn in Monolingual QA@CLEF 2008	333
<i>Sandra Roger, Katia Vila, Antonio Ferrández, María Pardiño, José Manuel Gómez, Marcel Puchol-Blasco, and Jesús Peral</i>	
Priberam's Question Answering System in QA@CLEF 2008	337
<i>Carlos Amaral, Adán Cassan, Helena Figueira, André Martins, Afonso Mendes, Pedro Mendes, José Pina, and Cláudia Pinto</i>	
IdSay: Question Answering for Portuguese	345
<i>Gracinda Carvalho, David Martins de Matos, and Vitor Rocio</i>	
Dublin City University at QA@CLEF 2008	353
<i>Sisay Fissaha Adafre and Josef van Genabith</i>	
Using Answer Retrieval Patterns to Answer Portuguese Questions.....	361
<i>Luís Fernando Costa</i>	
Ihardetsi: A Basque Question Answering System at QA@CLEF 2008 ...	369
<i>Olatz Ansa, Xabier Arregi, Arantxa Otegi, and Ander Soraluze</i>	
Question Interpretation in QA@L ² F	377
<i>Luísa Coheur, Ana Mendes, João Guimarães, Nuno J. Mamede, and Ricardo Ribeiro</i>	

UAIC Participation at QA@CLEF2008	385
<i>Adrian Iftene, Diana Trandabăț, Ionuț Pistol, Alex-Mihai Moruz, Maria Husarciuc, and Dan Cristea</i>	
RACAI's QA System at the Romanian–Romanian QA@CLEF2008 Main Task	393
<i>Radu Ion, Dan Ștefănescu, Alexandru Ceaușu, and Dan Tufiș</i>	
Combining Logic and Machine Learning for Answering Questions	401
<i>Ingo Glöckner and Björn Pelzer</i>	
The MIRACLE Team at the CLEF 2008 Multilingual Question Answering Track	409
<i>Ángel Martínez-González, César de Pablo-Sánchez, Concepción Polo-Bayo, María Teresa Vicente-Díez, Paloma Martínez-Fernández, and José Luís Martínez-Fernández</i>	
Efficient Question Answering with Question Decomposition and Multiple Answer Streams	421
<i>Sven Hartrumpf, Ingo Glöckner, and Johannes Leveling</i>	
DFKI-LT at QA@CLEF 2008	429
<i>Bogdan Sacaleanu, Günter Neumann, and Christian Spurk</i>	
Integrating Logic Forms and Anaphora Resolution in the AliQAn System	438
<i>Rafael Muñoz-Terol, Marcel Puchol-Blasco, María Pardiño, José Manuel Gómez, Sandra Roger, Katia Vila, Antonio Ferrández, Jesús Peral, and Patricio Martínez-Barco</i>	
Some Experiments in Question Answering with a Disambiguated Document Collection	442
<i>Davide Buscaldi and Paolo Rosso</i>	
Answer Validation Exercise (AVE)	
Answer Validation on English and Romanian Languages	448
<i>Adrian Iftene and Alexandra Balahur</i>	
The Answer Validation System ProdicosAV Dedicated to French	452
<i>Christine Jacquín, Laura Monceaux, and Emmanuel Desmontils</i>	
Studying the Influence of Semantic Constraints in AVE	460
<i>Óscar Ferrández, Rafael Muñoz, and Manuel Palomar</i>	
RAVE: A Fast Logic-Based Answer Validator	468
<i>Ingo Glöckner</i>	

Information Synthesis for Answer Validation	472
<i>Rui Wang and Günter Neumann</i>	

Analyzing the Use of Non-overlap Features for Supervised Answer Validation	476
<i>Alberto Téllez-Valero, Antonio Juárez-González, Manuel Montes-y-Gómez, and Luis Villaseñor-Pineda</i>	

Question Answering on Script Transcription (QAST)

The LIMSI Multilingual, Multitask QAst System	480
<i>Sophie Rosset, Olivier Galibert, Guillaume Bernard, Eric Bilinski, and Gilles Adda</i>	

IBQAst: A Question Answering System for Text Transcriptions	488
<i>María Pardiño, José M. Gómez, Héctor Llorens, Rafael Muñoz-Terol, Borja Navarro-Colorado, Estela Saquete, Patricio Martínez-Barco, Paloma Moreda, and Manuel Palomar</i>	

Robust Question Answering for Speech Transcripts: UPC Experience in QAst 2008	492
<i>Pere R. Comas and Jordi Turmo</i>	

Part V: Cross-Language Retrieval in Image Collections (ImageCLEF)

Overview of the ImageCLEFphoto 2008 Photographic Retrieval Task . . .	500
<i>Thomas Arni, Paul Clough, Mark Sanderson, and Michael Grubinger</i>	

Overview of the ImageCLEFmed 2008 Medical Image Retrieval Task . . .	512
<i>Henning Müller, Jayashree Kalpathy-Cramer, Charles E. Kahn Jr., William Hatt, Steven Bedrick, and William Hersh</i>	

Medical Image Annotation in ImageCLEF 2008	523
<i>Thomas Deselaers and Thomas M. Deserno</i>	

The Visual Concept Detection Task in ImageCLEF 2008	531
<i>Thomas Deselaers and Allan Hanbury</i>	

Overview of the WikipediaMM Task at ImageCLEF 2008	539
<i>Theodora Tsikrika and Jana Kludas</i>	

ImageCLEFphoto

Meiji University at ImageCLEF2008 Photo Retrieval Task: Evaluation of Image Retrieval Methods Integrating Different Media	551
<i>Kosuke Yamauchi, Takuya Nomura, Keiko Usui, Yusuke Kamo, and Tomohiro Takagi</i>	

Building a Diversity Featured Search System by Fusing Existing Tools	560
<i>Jiayu Tang, Thomas Arni, Mark Sanderson, and Paul Clough</i>	
Some Results Using Different Approaches to Merge Visual and Text-Based Features in CLEF'08 Photo Collection	568
<i>Ana García-Serrano, Xaro Benavent, Ruben Granados, and José Miguel Goñi-Menoyo</i>	
MIRACLE-GSI at ImageCLEFphoto 2008: Different Strategies for Automatic Topic Expansion	572
<i>Julio Villena-Román, Sara Lana-Serrano, and José Carlos González-Cristóbal</i>	
Using Visual Concepts and Fast Visual Diversity to Improve Image Retrieval	577
<i>Sabrina Tollari, Marcin Detyniecki, Ali Fakeri-Tabrizi, Christophe Marsala, Massih-Reza Amini, and Patrick Gallinari</i>	
A Comparative Study of Diversity Methods for Hybrid Text and Image Retrieval Approaches	585
<i>Sabrina Tollari, Philippe Mulhem, Marin Ferecatu, Hervé Glotin, Marcin Detyniecki, Patrick Gallinari, Hichem Sahbi, and Zhong-Qiu Zhao</i>	
University of Jaén at ImagePhoto 2008: Filtering the Results with the Cluster Term	593
<i>Miguel Angel García-Cumbreras, Manuel Carlos Díaz-Galiano, María Teresa Martín-Valdivia, and L. Alfonso Ureña-López</i>	
Combining TEXT-MESS Systems at ImageCLEF 2008	597
<i>Sergio Navarro, Miguel Angel García-Cumbreras, Fernando Llopis, Manuel Carlos Díaz-Galiano, Rafael Muñoz, María Teresa Martín-Valdivia, L. Alfonso Ureña-López, and Arturo Montejo-Ráez</i>	
Image Retrieval by Inter-media Fusion and Pseudo-relevance Feedback	605
<i>Osama El Demerdash, Leila Kosseim, and Sabine Bergler</i>	
Increasing Precision and Diversity in Photo Retrieval by Result Fusion	612
<i>Yih-Chen Chang and Hsin-Hsi Chen</i>	
Diversity in Image Retrieval: DCU at ImageCLEFPhoto 2008	620
<i>Neil O'Hare, Peter Wilkins, Cathal Gurrin, Eamonn Newman, Gareth J.F. Jones, and Alan F. Smeaton</i>	

Visual Affinity Propagation Improves Sub-topics Diversity without Loss of Precision in Web Photo Retrieval	628
<i>Hervé Glotin and Zhong-Qiu Zhao</i>	
Exploiting Term Co-occurrence for Enhancing Automated Image Annotation.....	632
<i>Ainhoa Llorente, Simon Overell, Haiming Liu, Rui Hu, Adam Rae, Jianhan Zhu, Dawei Song, and Stefan Ruger</i>	
Enhancing Visual Concept Detection by a Novel Matrix Modular Scheme on SVM	640
<i>Zhong-Qiu Zhao and Herve Glotin</i>	
SZTAKI @ ImageCLEF 2008: Visual Feature Analysis in Segmented Images	644
<i>Balint Daroczy, Zsolt Fekete, Matyas Brendel, Simon Racz, Andras Benczur, David Siklosi, and Attila Pereszlenyi</i>	
THESEUS Meets ImageCLEF: Combining Evaluation Strategies for a New Visual Concept Detection Task 2009	652
<i>Stefanie Nowak, Peter Dunker, and Ronny Paduschek</i>	
Query Types and Visual Concept-Based Post-retrieval Clustering	661
<i>Masashi Inoue and Piyush Grover</i>	
Annotation-Based Expansion and Late Fusion of Mixed Methods for Multimedia Image Retrieval	669
<i>Hugo Jair Escalante, Jesus A. Gonzalez, Carlos A. Hernandez, Aurelio Lopez, Manuel Montes, Eduardo Morales, Luis E. Sucar, and Luis Villasenor-Pineda</i>	
Evaluation of Diversity-Focused Strategies for Multimedia Retrieval	677
<i>Julien Ah-Pine, Gabriela Csurka, and Jean-Michel Renders</i>	
Clustering for Photo Retrieval at Image CLEF 2008	685
<i>Diana Inkpen, Marc Stogaitis, Franois DeGuire, and Muath Alzghool</i>	
ImageCLEFmed	
Methods for Combining Content-Based and Textual-Based Approaches in Medical Image Retrieval	691
<i>Mouna Torjmen, Karen Pinel-Sauvagnat, and Mohand Boughanem</i>	
An SVM Confidence-Based Approach to Medical Image Annotation	696
<i>Tatiana Tommasi, Francesco Orabona, and Barbara Caputo</i>	
LIG at ImageCLEF 2008.....	704
<i>Loic Maisonnasse, Philippe Mulhem, Eric Gaussier, and Jean Pierre Chevallet</i>	

The MedGIFT Group at ImageCLEF 2008	712
<i>Xin Zhou, Julien Gobeill, and Henning Müller</i>	
MIRACLE at ImageCLEFmed 2008: Semantic vs. Statistical Strategies for Topic Expansion	719
<i>Sara Lana-Serrano, Julio Villena-Román, and José Carlos González-Cristóbal</i>	
Experiments in Calibration and Validation for Medical Content-Based Images Retrieval	724
<i>Jose L. Delgado, Covadonga Rodrigo, and Gonzalo León</i>	
MIRACLE at ImageCLEFannot 2008: Nearest Neighbour Classification of Image Feature Vectors for Medical Image Annotation	728
<i>Sara Lana-Serrano, Julio Villena-Román, José Carlos González-Cristóbal, and José Miguel Goñi-Menoyo</i>	
Query Expansion on Medical Image Retrieval: MeSH vs. UMLS	732
<i>Manuel Carlos Díaz-Galiano, Miguel Angel García-Cumbreras, María Teresa Martín-Valdivia, L. Alfonso Ureña-López, and Arturo Montejo-Rúez</i>	
Query and Document Expansion with Medical Subject Headings Terms at Medical Imageclef 2008	736
<i>Julien Gobeill, Patrick Ruch, and Xin Zhou</i>	
Multimodal Medical Image Retrieval OHSU at ImageCLEF 2008	744
<i>Jayashree Kalpathy-Cramer, Steven Bedrick, William Hatt, and William Hersh</i>	
Baseline Results for the ImageCLEF 2008 Medical Automatic Annotation Task in Comparison over the Years	752
<i>Mark O. Güld, Petra Welter, and Thomas M. Deserno</i>	
ImageCLEFWiki	
Evaluating the Impact of Image Names in Context-Based Image Retrieval	756
<i>Mouna Torjmen, Karen Pinel-Sauvagnat, and Mohand Boughanem</i>	
Large-Scale Cross-Media Retrieval of WikipediaMM Images with Textual and Visual Query Expansion	763
<i>Zhi Zhou, Yonghong Tian, Yuanning Li, Tiejun Huang, and Wen Gao</i>	
Conceptual Image Retrieval over a Large Scale Database	771
<i>Adrian Popescu, Hervé Le Borgne, and Pierre-Alain Moëllic</i>	

UJM at ImageCLEFwiki 2008	779
<i>Christophe Moulin, Cécile Barat, Mathias Géry, Christophe Ducottet, and Christine Largeron</i>	

Part VI: Multilingual Web Track (WebCLEF)

Overview of WebCLEF 2008	787
<i>Valentin Jijkoun and Maarten de Rijke</i>	
On the Evaluation of Snippet Selection for WebCLEF	794
<i>Arnold Overwijk, Dong Nguyen, Claudia Hauff, Dolf Trieschnigg, Djoerd Hiemstra, and Franciska de Jong</i>	
UNED at WebCLEF 2008: Applying High Restrictive Summarization, Low Restrictive Information Retrieval and Multilingual Techniques	798
<i>Enrique Amigó, Juan Martinez-Romo, Lourdes Araujo, and V́ctor Peinado</i>	
Retrieval of Snippets of Web Pages Converted to Plain Text. More Questions Than Answers	802
<i>Carlos G. Figuerola, José Luis Alonso Berrocal, Ángel F. Zazo Rodríguez, and Montserrat Mateos</i>	

Part VII: Cross-Language Geographical Retrieval (GeoCLEF)

GeoCLEF 2008: The CLEF 2008 Cross-Language Geographic Information Retrieval Track Overview	808
<i>Thomas Mandl, Paula Carvalho, Giorgio Maria Di Nunzio, Fredric Gey, Ray R. Larson, Diana Santos, and Christa Womser-Hacker</i>	
GIR with Language Modeling and DFR Using Terrier	822
<i>Rocio Guillén</i>	
Cheshire at GeoCLEF 2008: Text and Fusion Approaches for GIR	830
<i>Ray R. Larson</i>	
Geographic and Textual Data Fusion in Forostar	838
<i>Simon Overell, Adam Rae, and Stefan Rűger</i>	
Query Expansion for Effective Geographic Information Retrieval	843
<i>Qiang Pu, Daqing He, and Qi Li</i>	
Integrating Methods from IR and QA for Geographic Information Retrieval	851
<i>Johannes Leveling and Sven Hartrumpf</i>	

Using Query Reformulation and Keywords in the Geographic Information Retrieval Task	855
<i>José Manuel Perea-Ortega, L. Alfonso Ureña-López, Manuel García-Vega, and Miguel Angel García-Cumbreras</i>	
Using GeoWordNet for Geographical Information Retrieval	863
<i>Davide Buscaldi and Paolo Rosso</i>	
GeoTextMESS: Result Fusion with Fuzzy Borda Ranking in Geographical Information Retrieval	867
<i>Davide Buscaldi, José Manuel Perea Ortega, Paolo Rosso, L. Alfonso Ureña López, Daniel Ferrés, and Horacio Rodríguez</i>	
A Ranking Approach Based on Example Texts for Geographic Information Retrieval	875
<i>Esaú Villatoro-Tello, Manuel Montes-y-Gómez, and Luis Villaseñor-Pineda</i>	
Ontology-Based Query Construction for GeoCLEF	880
<i>Rui Wang and Günter Neumann</i>	
Experiments with Geographic Evidence Extracted from Documents	885
<i>Nuno Cardoso, Patrícia Sousa, and Mário J. Silva</i>	
GikiP at GeoCLEF 2008: Joining GIR and QA Forces for Querying Wikipedia	894
<i>Diana Santos, Nuno Cardoso, Paula Carvalho, Iustin Dornescu, Sven Hartrumpf, Johannes Leveling, and Yvonne Skalban</i>	
Part VIII: Cross-Language Video Retrieval (VideoCLEF)	
Overview of VideoCLEF 2008: Automatic Generation of Topic-Based Feeds for Dual Language Audio-Visual Content	906
<i>Martha Larson, Eamonn Newman, and Gareth J.F. Jones</i>	
MIRACLE at VideoCLEF 2008: Topic Identification and Keyframe Extraction in Dual Language Videos	918
<i>Julio Villena-Román and Sara Lana-Serrano</i>	
DCU at VideoClef 2008	923
<i>Eamonn Newman and Gareth J.F. Jones</i>	
Using an Information Retrieval System for Video Classification	927
<i>José Manuel Perea-Ortega, Arturo Montejo-Ráez, Manuel Carlos Díaz-Galiano, María Teresa Martín-Valdivia, and L. Alfonso Ureña-López</i>	

VideoCLEF 2008: ASR Classification with Wikipedia Categories 931
Jens Kürsten, Daniel Richter, and Maximilian Eibl

Metadata and Multilinguality in Video Classification 935
Jiyin He, Xu Zhang, Wouter Weerkamp, and Martha Larson

**Part IX: Multilingual Information Filtering
 (INFILE@CLEF)**

Overview of CLEF 2008 INFILE Pilot Track 939
*Romaric Besançon, Stéphane Chaudiron, Djamel Mostefa,
 Olivier Hamon, Ismaïl Timimi, and Khalid Choukri*

Online Document Filtering Using Adaptive k-NN 947
Vincent Bodinier, Ali Mustafa Qamar, and Eric Gaussier

Part X: Morpho Challenge at CLEF 2008

Overview of Morpho Challenge 2008 951
Mikko Kurimo, Ville Turunen, and Matti Varjokallio

ParaMor and Morpho Challenge 2008 967
Christian Monson, Jaime Carbonell, Alon Lavie, and Lori Levin

AllomorfeSSor: Towards Unsupervised Morpheme Analysis 975
Oskar Kohonen, Sami Virpioja, and Mikaela Klami

Using Unsupervised Paradigm Acquisition for Prefixes 983
Daniel Zeman

Morpho Challenge Evaluation by Information Retrieval Experiments . . . 991
Mikko Kurimo, Mathias Creutz, and Ville Turunen

Author Index 999