

Table of Contents

Introduction

What Happened in CLEF 2007	1
<i>Carol Peters</i>	

Part I: Multilingual Textual Document Retrieval (Ad Hoc)

CLEF 2007: Ad Hoc Track Overview	13
<i>Giorgio M. Di Nunzio, Nicola Ferro, Thomas Mandl, and Carol Peters</i>	

Monolingual

Charles University at CLEF 2007 Ad-Hoc Track	33
<i>Pavel Češka and Pavel Pecina</i>	
Stemming Approaches for East European Languages	37
<i>Ljiljana Dolamic and Jacques Savoy</i>	
Applying Query Expansion Techniques to Ad Hoc Monolingual Tasks with the IR-n System	45
<i>Elisa Noguera and Fernando Llopis</i>	
Bulgarian, Hungarian and Czech Stemming Using YASS	49
<i>Prasenjit Majumder, Mandar Mitra, and Dipasree Pal</i>	
Sampling Precision to Depth 10000 at CLEF 2007	57
<i>Stephen Tomlinson</i>	

Cross-Language: European

Disambiguation and Unknown Term Translation in Cross Language Information Retrieval	64
<i>Dong Zhou, Mark Truran, and Tim Brailsford</i>	
Cross-Language Retrieval with Wikipedia	72
<i>Péter Schönhofen, András Benczúr, István Bíró, and Károly Csalogány</i>	

Cross-Language: Non-European

Cross-Lingual Information Retrieval System for Indian Languages	80
<i>Jagadeesh Jagarlamudi and A. Kumaran</i>	

Bengali, Hindi and Telugu to English Ad-Hoc Bilingual Task at CLEF 2007	88
<i>Sivaji Bandyopadhyay, Tapabrata Mondal, Sudip Kumar Naskar, Asif Ekbal, Rejwanul Haque, and Srinivasa Rao Godhavarthy</i>	
Bengali and Hindi to English CLIR Evaluation	95
<i>Debasis Mandal, Mayank Gupta, Sandipan Dandapat, Pratyush Banerjee, and Sudeshna Sarkar</i>	
Improving Recall for Hindi, Telugu, Oromo to English CLIR	103
<i>Prasad Pingali, Kula Kekeba Tune, and Vasudeva Varma</i>	
Hindi to English and Marathi to English Cross Language Information Retrieval Evaluation	111
<i>Manoj Kumar Chinnakotla, Sagar Ranadive, Om P. Damani, and Pushpak Bhattacharyya</i>	
Amharic-English Information Retrieval with Pseudo Relevance Feedback	119
<i>Atelach Alemu Argaw</i>	
Indonesian-English Transitive Translation for Cross-Language Information Retrieval	127
<i>Mirna Adriani, Herika Hayurani, and Syandra Sari</i>	
Robust	
Robust Retrieval Experiments at the University of Hildesheim	134
<i>Ben Heuwing and Thomas Mandl</i>	
SINAI at CLEF Ad-Hoc Robust Track 2007: Applying Google Search Engine for Robust Cross-Lingual Retrieval	137
<i>F. Martínez-Santiago, A. Montejo-Ráez, and M.Á. García-Cumbreras</i>	
Improving Robustness Using Query Expansion	143
<i>Angel F. Zazo, José L. Alonso Berrocal, and Carlos G. Figuerola</i>	
English-to-French CLIR: A Knowledge-Light Approach through Character <i>N</i> -Grams Alignment	148
<i>Jesús Vilares, Michael P. Oakes, and Manuel Vilares</i>	
MIRACLE Progress in Monolingual Information Retrieval at Ad-Hoc CLEF 2007	156
<i>José-Carlos González-Cristóbal, José Miguel Goñi-Menoyo, Julio Villena-Román, and Sara Lana-Serrano</i>	
Part II: Domain-Specific Information Retrieval (Domain-Specific)	
The Domain-Specific Track at CLEF 2007	160
<i>Vivien Petras, Stefan Baerisch, and Maximillian Stempfhuber</i>	

The XTRIEVAL Framework at CLEF 2007: Domain-Specific Track	174
<i>Jens Kürsten, Thomas Wilhelm, and Maximilian Eibl</i>	
Query Translation through Dictionary Adaptation	182
<i>Stéphane Clinchant and Jean-Michel Renders</i>	
Experiments in Classification Clustering and Thesaurus Expansion for Domain Specific Cross-Language Retrieval	188
<i>Ray R. Larson</i>	
Domain-Specific IR for German, English and Russian Languages	196
<i>Claire Fautsch, Ljiljana Dolamic, Samir Abdou, and Jacques Savoy</i>	
Part III: Multiple Language Question Answering (QA@CLEF)	
Overview of the CLEF 2007 Multilingual Question Answering Track	200
<i>Danilo Giampiccolo, Pamela Forner, Jesús Herrera, Anselmo Peñas, Christelle Ayache, Corina Forascu, Valentin Jijkoun, Petya Osenova, Paulo Rocha, Bogdan Sacaleanu, and Richard Sutcliffe</i>	
Overview of the Answer Validation Exercise 2007	237
<i>Anselmo Peñas, Álvaro Rodrigo, and Felisa Verdejo</i>	
Overview of QAST 2007	249
<i>Jordi Turmo, Pere R. Comas, Christelle Ayache, Djamel Mostefa, Sophie Rosset, and Lori Lamel</i>	
Main Task: Mono- and Bilingual QA	
Question Answering with Joost at CLEF 2007	257
<i>Gosse Bouma, Geert Kloosterman, Jori Mur, Gertjan van Noord, Lonneke van der Plas, and Jörg Tiedemann</i>	
What Happened to Esfinge in 2007?	261
<i>Luís Miguel Cabral, Luís Fernando Costa, and Diana Santos</i>	
Coreference Resolution for Questions and Answer Merging by Validation	269
<i>Sven Hartrumpf, Ingo Glöckner, and Johannes Leveling</i>	
Multilingual Question Answering through Intermediate Translation: LCC's PowerAnswer at QA@CLEF 2007	273
<i>Mitchell Bowden, Marian Olteanu, Pasin Suriyentrakorn, Thomas d'Silva, and Dan Moldovan</i>	
RACAI's Question Answering System at QA@CLEF2007	284
<i>Dan Tufiş, Dan Ştefănescu, Radu Ion, and Alexandru Ceauşu</i>	

DFKI-LT at QA@CLEF 2007	292
<i>Bogdan Sacaleanu, Günter Neumann, and Christian Spurk</i>	
University of Wolverhampton at CLEF 2007.....	300
<i>Georgiana Puşcaşu and Constantin Orăsan</i>	
Bilingual Question Answering Using CINDI-QA at QA@CLEF 2007....	308
<i>Chedid Haddad and Bipin C. Desai</i>	
The University of Évora's Participation in QA@CLEF-2007	316
<i>José Saias and Paulo Quaresma</i>	
Web-Based Anaphora Resolution for the QUASAR Question Answering System	324
<i>Davide Buscaldi, Yassine Benajiba, Paolo Rosso, and Emilio Sanchis</i>	
A Lexical Approach for Spanish Question Answering	328
<i>Alberto Téllez, Antonio Juárez, Gustavo Hernández, Claudia Denicia, Esaú Villatoro, Manuel Montes, and Luis Villaseñor</i>	
Finding Answers Using Resources in the Internet	332
<i>Mirna Adriani and Septian Adiwibowo</i>	
UAIC Romanian QA System for QA@CLEF	336
<i>Adrian Iftene, Diana Trandabăt, Ionuţ Pistol, Alex Moruz, Alexandra Balahur, Diana Cotelea, Iustin Dornescu, Iuliana Drăghici, and Dan Cristea</i>	
The University of Amsterdam's Question Answering System at QA@CLEF 2007	344
<i>Valentin Jijkoun, Katja Hofmann, David Ahn, Mahboob Alam Khalid, Joris van Rantwijk, Maarten de Rijke, and Erik Tjong Kim Sang</i>	
Combining Wikipedia and Newswire Texts for Question Answering in Spanish.....	352
<i>César de Pablo-Sánchez, José L. Martínez-Fernández, Ana González-Ledesma, Doaa Samy, Paloma Martínez, Antonio Moreno-Sandoval, and Harith Al-Jumaily</i>	
QA@L ² F, First Steps at QA@CLEF	356
<i>Ana Mendes, Luísa Coheur, Nuno J. Mamede, Ricardo Ribeiro, Fernando Batista, and David Martins de Matos</i>	
Príberam's Question Answering System in QA@CLEF 2007	364
<i>Carlos Amaral, Adán Cassan, Helena Figueira, André Martins, Afonso Mendes, Pedro Mendes, Cláudia Pinto, and Daniel Vidal</i>	

Answer Validation Exercise (AVE)

Combining Logic and Aggregation for Answer Selection	372
<i>Ingo Glöckner</i>	
On the Application of Lexical-Syntactic Knowledge to the Answer Validation Exercise	377
<i>Óscar Ferrández, Daniel Micol, Rafael Muñoz, and Manuel Palomar</i>	
Combining Lexical Information with Machine Learning for Answer Validation at QA@CLEF 2007	381
<i>M.Á. García-Cumbreras, J.M. Perea-Ortega, F. Martínez-Santiago, and L. Alfonso Ureña-López</i>	
Using Recognizing Textual Entailment as a Core Engine for Answer Validation	387
<i>Rui Wang and Günter Neumann</i>	
A Supervised Learning Approach to Spanish Answer Validation	391
<i>Alberto Téllez-Valero, Manuel Montes-y-Gómez, and Luis Villaseñor-Pineda</i>	
UAIC Participation at AVE 2007	395
<i>Adrian Iftene and Alexandra Balahur-Dobrescu</i>	
UNED at Answer Validation Exercise 2007	404
<i>Álvaro Rodrigo, Anselmo Peñas, and Felisa Verdejo</i>	

Question Answering on Speech Transcription (QAST)

Adapting QA Components to Mine Answers in Speech Transcripts	410
<i>Günter Neumann and Rui Wang</i>	
The LIMSI Participation in the QAst Track	414
<i>Sophie Rosset, Olivier Galibert, Gilles Adda, and Eric Bilinski</i>	
Robust Question Answering for Speech Transcripts Using Minimal Syntactic Analysis	424
<i>Pere R. Comas, Jordi Turmo, and Mihai Surdeanu</i>	

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

Overview of the ImageCLEFphoto 2007 Photographic Retrieval Task . . .	433
<i>Michael Grubinger, Paul Clough, Allan Hanbury, and Henning Müller</i>	

Overview of the ImageCLEF 2007 Object Retrieval Task	445
<i>Thomas Deselaers, Allan Hanbury, Ville Viitaniemi, Andras Benczur, Matyas Brendel, Balint Daroczy, Hugo Jair Escalante Balderas, Theo Gevers, Carlos Arturo Hernandez Gracidas, Steven C.H. Hoi, Jorma Laaksonen, Mingjing Li, Heidy Marisol Marın Castro, Hermann Ney, Xiaoguang Rui, Nicu Sebe, Julian Stottinger, and Lei Wu</i>	
Overview of the ImageCLEFmed 2007 Medical Retrieval and Medical Annotation Tasks	472
<i>Henning Muller, Thomas Deselaers, Thomas M. Deserno, Jayashree Kalpathy-Cramer, Eugene Kim, and William Hersh</i>	
ImageCLEFphoto	
FIRE in ImageCLEF 2007: Support Vector Machines and Logistic Models to Fuse Image Descriptors for Photo Retrieval	492
<i>Tobias Gass, Tobias Weyand, Thomas Deselaers, and Hermann Ney</i>	
MIRACLE at ImageCLEFphoto 2007: Evaluation of Merging Strategies for Multilingual and Multimedia Information Retrieval	500
<i>Julio Villena-Roman, Sara Lana-Serrano, Jose Luis Martınez-Fernandez, and Jose Carlos Gonzalez-Cristobal</i>	
Using an Image-Text Parallel Corpus and the Web for Query Expansion in Cross-Language Image Retrieval	504
<i>Yih-Chen Chang and Hsin-Hsi Chen</i>	
SINAI System: Combining IR Systems at ImageCLEFPhoto 2007	512
<i>M.. Garcıa-Cumbreras, M.C. Dıaz-Galiano, M.T. Martın-Valdivia, A. Montejo-Ruez, and L.A. Urena-Lopez</i>	
Multimodal Retrieval by Text-Segment Biclustering	518
<i>Andras Benczur, Istvan Bıro, Matyas Brendel, Karoly Csalogany, Balint Daroczy, and David Siklosi</i>	
Analysing an Approach to Information Retrieval of Visual Descriptions with IR-n, a System Based on Passages	522
<i>Sergio Navarro, Fernando Llopis, Rafael Munoz Guillena, and Elisa Noguera</i>	
DCU and UTA at ImageCLEFPhoto 2007	530
<i>Anni Jarvelin, Peter Wilkins, Tomasz Adamek, Eija Airio, Gareth J.F. Jones, Alan F. Smeaton, and Eero Sormunen</i>	

Cross-Language and Cross-Media Image Retrieval: An Empirical Study at ImageCLEF2007	538
<i>Steven C.H. Hoi</i>	
Towards Annotation-Based Query and Document Expansion for Image Retrieval.....	546
<i>Hugo Jair Escalante, Carlos Hernández, Aurelio López, Heidy Marín, Manuel Montes, Eduardo Morales, Enrique Sucar, and Luis Villaseñor</i>	
Content-Based Image Retrieval Using Combined 2D Attribute Pattern Spectra	554
<i>Florence Tushabe and Michael. H.F. Wilkinson</i>	
Text-Based Clustering of the ImageCLEFphoto Collection for Augmenting the Retrieved Results	562
<i>Osama El Demerdash, Leila Kosseim, and Sabine Bergler</i>	
Trans-Media Pseudo-Relevance Feedback Methods in Multimedia Retrieval.....	569
<i>Stéphane Clinchant, Jean-Michel Renders, and Gabriela Csurka</i>	
ImageCLEFmed	
Cue Integration for Medical Image Annotation.....	577
<i>Tatiana Tommasi, Francesco Orabona, and Barbara Caputo</i>	
Multiplying Concept Sources for Graph Modeling	585
<i>Loïc Maisonnasse, Eric Gaussier, and Jean Pierre Chevillet</i>	
MIRACLE at ImageCLEFmed 2007: Merging Textual and Visual Strategies to Improve Medical Image Retrieval.....	593
<i>Julio Villena-Román, Sara Lana-Serrano, and José Carlos González-Cristóbal</i>	
MIRACLE at ImageCLEFanot 2007: Machine Learning Experiments on Medical Image Annotation	597
<i>Sara Lana-Serrano, Julio Villena-Román, José Carlos González-Cristóbal, and José Miguel Goñi-Menoyo</i>	
Integrating MeSH Ontology to Improve Medical Information Retrieval.....	601
<i>M.C. Díaz-Galiano, M.Á. García-Cumbreras, M.T. Martín-Valdivia, A. Montejo-Ráez, and L.A. Ureña-López</i>	
Speeding Up IDM without Degradation of Retrieval Quality	607
<i>Michael Springmann and Heiko Schuldt</i>	

Content-Based Medical Image Retrieval Using Low-Level Visual Features and Modality Identification	615
<i>Juan C. Caicedo, Fabio A. Gonzalez, and Eduardo Romero</i>	
Medical Image Retrieval and Automatic Annotation: OHSU at ImageCLEF 2007	623
<i>Jayashree Kalpathy-Cramer and William Hersh</i>	
Using Bayesian Network for Conceptual Indexing: Application to Medical Document Indexing with UMLS Metathesaurus	631
<i>Thi Hoang Diem Le, Jean-Pierre Chevallet, and Joo Hwee Lim</i>	
Baseline Results for the ImageCLEF 2007 Medical Automatic Annotation Task Using Global Image Features	637
<i>Mark O. Güld and Thomas M. Deserno</i>	
Evaluation of Automatically Assigned MeSH Terms for Retrieval of Medical Images	641
<i>Miguel E. Ruiz and Aurélie Névéol</i>	
ImageCLEF photo and med	
University and Hospitals of Geneva Participating at ImageCLEF 2007	649
<i>Xin Zhou, Julien Gobeill, Patrick Ruch, and Henning Müller</i>	
An Interactive and Dynamic Fusion-Based Image Retrieval Approach by CINDI	657
<i>M.M. Rahman, B.C. Desai, and P. Bhattacharya</i>	
Using Pseudo-Relevance Feedback to Improve Image Retrieval Results	665
<i>Mouna Torjmen, Karen Pinel-Sauvagnat, and Mohand Boughanem</i>	
Part V: Cross-Language Speech Retrieval (CL-SR)	
Overview of the CLEF-2007 Cross-Language Speech Retrieval Track	674
<i>Pavel Pecina, Petra Hoffmannová, Gareth J.F. Jones, Ying Zhang, and Douglas W. Oard</i>	
A Dirichlet-Smoothed Bigram Model for Retrieving Spontaneous Speech	687
<i>Matthew Lease and Eugene Charniak</i>	
Model Fusion Experiments for the CLSR Task at CLEF 2007	695
<i>Muath Alzghool and Diana Inkpen</i>	

Dublin City University at CLEF 2007: Cross-Language Speech Retrieval Experiments	703
<i>Ying Zhang, Gareth J.F. Jones, and Ke Zhang</i>	
What Can and Cannot Be Found in Czech Spontaneous Speech Using Document-Oriented IR Methods—UWB at CLEF 2007 CL-SR Track.....	712
<i>Pavel Ircing, Josef Psutka, and Jan Vavruška</i>	
Using Information Gain to Filter Information in CLEF CL-SR Track ...	719
<i>M.C. Díaz-Galiano, M.T. Martín-Valdivia, M.Á. García-Cumbreras, and L.A. Ureña-López</i>	
Part VI: Multilingual Web Retrieval (WebCLEF)	
Overview of WebCLEF 2007	725
<i>Valentin Jijkoun and Maarten de Rijke</i>	
Segmentation of Web Documents and Retrieval of Useful Passages	732
<i>Carlos G. Figuerola, José L. Alonso Berrocal, and Angel F. Zazo Rodríguez</i>	
Using Centrality to Rank Web Snippets.....	737
<i>Valentin Jijkoun and Maarten de Rijke</i>	
Using Web-Content for Retrieving Snippets	742
<i>Okky Hendriansyah, Tri Firgantoro, and Mirna Adriani</i>	
Part VII: Cross-Language Geographical Retrieval (GeoCLEF)	
GeoCLEF 2007: The CLEF 2007 Cross-Language Geographic Information Retrieval Track Overview	745
<i>Thomas Mandl, Fredric Gey, Giorgio Di Nunzio, Nicola Ferro, Ray Larson, Mark Sanderson, Diana Santos, Christa Womser-Hacker, and Xing Xie</i>	
Inferring Location Names for Geographic Information Retrieval	773
<i>Johannes Leveling and Sven Hartrumpf</i>	
GeoParsing Web Queries	781
<i>Rocio Guillén</i>	
MIRACLE at GeoCLEF Query Parsing 2007: Extraction and Classification of Geographical Information.....	786
<i>Sara Lana-Serrano, Julio Villena-Román, José Carlos González-Cristóbal, and José Miguel Goñi-Menoyo</i>	

Relevance Measures Using Geographic Scopes and Types	794
<i>Geoffrey Andogah and Gosse Bouma</i>	
Using Geographic Signatures as Query and Document Scopes in Geographic IR	802
<i>Nuno Cardoso, David Cruz, Marcirio Chaves, and Mário J. Silva</i>	
Cheshire at GeoCLEF 2007: Retesting Text Retrieval Baselines	811
<i>Ray R. Larson</i>	
On the Relative Importance of Toponyms in GeoCLEF	815
<i>Davide Buscaldi and Paolo Rosso</i>	
Filtering for Improving the Geographic Information Search	823
<i>José M. Perea-Ortega, Miguel A. García-Cumbreras, Manuel García-Vega, and L.A. Ureña-López</i>	
TALP at GeoCLEF 2007: Results of a Geographical Knowledge Filtering Approach with Terrier	830
<i>Daniel Ferrés and Horacio Rodríguez</i>	
TALP at GeoQuery 2007: Linguistic and Geographical Analysis for Query Parsing	834
<i>Daniel Ferrés and Horacio Rodríguez</i>	
Applying Geo-feedback to Geographic Information Retrieval	838
<i>Mirna Adriani and Nasikhin</i>	
Exploring LDA-Based Document Model for Geographic Information Retrieval	842
<i>Zhisheng Li, Chong Wang, Xing Xie, Xufa Wang, and Wei-Ying Ma</i>	
Mono- and Crosslingual Retrieval Experiments with Spatial Restrictions at GeoCLEF 2007	850
<i>Ralph Kõlle, Ben Heuwing, Thomas Mandl, and Christa Womser-Hacker</i>	
GIR Experiments with Forostar	856
<i>Simon Overell, João Magalhães, and Stefan Rüger</i>	

Part VIII: CLEF in Other Evaluations

CLEF at MorphoChallenge

Morpho Challenge Evaluation Using a Linguistic Gold Standard	864
<i>Mikko Kurimo, Mathias Creutz, and Matti Varjokallio</i>	
Simple Morpheme Labelling in Unsupervised Morpheme Analysis	873
<i>Delphine Bernhard</i>	

Unsupervised and Knowledge-Free Morpheme Segmentation and Analysis	881
<i>Stefan Bordag</i>	
Unsupervised Acquiring of Morphological Paradigms from Tokenized Text	892
<i>Daniel Zeman</i>	
ParaMor: Finding Paradigms across Morphology	900
<i>Christian Monson, Jaime Carbonell, Alon Lavie, and Lori Levin</i>	
CLEF at SemEval 2007	
SemEval-2007 Task 01: Evaluating WSD on Cross-Language Information Retrieval.....	908
<i>Eneko Agirre, Oier Lopez de Lacalle, Bernardo Magnini, Arantxa Otegi, German Rigau, and Piek Vossen</i>	
Author Index	919