

DELOS

AN ASSOCIATION FOR  
DIGITAL  
LIBRARIES

**Post-proceedings of the  
5<sup>th</sup> Italian Research Conference on  
Digital Libraries - IRCDL 2009**

*A Conference of the DELOS Association and the  
Department of Information Engineering of the  
University of Padova*

**Padova, 29-30 January 2009**

Editors: M. Agosti, F. Esposito and C. Thanos  
July 2009



DELOS: an Association for Digital Libraries



Maristella Agosti Floriana Esposito Costantino Thanos  
(Editors)

**Post-proceedings of the  
5<sup>th</sup> Italian Research Conference on  
Digital Libraries - IRCDL 2009**

*A Conference of the DELOS Association and the  
Department of Information Engineering of the University  
of Padova*

Padova, Italy, 29-30 January 2009

Revised Selected Papers

## Volume Editors

Maristella Agosti  
Department of Information Engineering  
University of Padua  
Via Gradenigo, 6/a  
35131 Padova - Italy  
Email: [agosti@dei.unipd.it](mailto:agosti@dei.unipd.it)

Floriana Esposito  
Department of Computer Science  
University of Bari  
Via Orabona, 4  
70126 Bari - Italy  
Email: [esposito@di.uniba.it](mailto:esposito@di.uniba.it)

Costantino Thanos  
Istituto di Scienza e Tecnologie dell'Informazione "Alessandro Faedo"  
Area della Ricerca CNR di Pisa  
Via G. Moruzzi, 1  
56124 Pisa - Italy  
Email: [Costantino.Thanos@isti.cnr.it](mailto:Costantino.Thanos@isti.cnr.it)

## Preface

This volume contains the revised accepted papers presented at the *5th Italian Research Conference on Digital Libraries* (IRCDL 2009), which was held in the Department of Information Engineering of the University of Padova, Italy, from the 29<sup>th</sup> to the 30<sup>th</sup> January 2009. The well-established aim of IRCDL is to bring together the Italian researchers interested in the different methods and techniques that allow to build and operate Digital Libraries and to present ongoing research projects. A national Programme Committee was set up composed of 12 members, with representatives of the most active Italian research groups on digital libraries. Among the papers presented at the conference, 12 were accepted for inclusion in these proceedings together with 9 project presentations.

The IRCDL series of national conferences was originally conceived and organized in the context of the activities of DELOS, the Network of Excellence on Digital Libraries (<http://www.delos.info/>), partially funded by the European Union under the Sixth Framework Program from 2004 to 2007. The first IRCDL conference took place in 2005, as an opportunity for the Italian researchers to present recent results on their research activities related to the wide world of Digital Libraries. In particular, young researchers were (and still are) invited to submit the results of their on going research, to be presented in a friendly and relaxed atmosphere, to facilitate constructive discussions and exchange of opinions.

Thanks to the initial support of DELOS, and later on thanks to the support of both the DELOS Association and the Department of Information Engineering of the University of Padua, IRCDL has continued in the subsequent years and is confirming itself as a yearly meeting point for the Italian researchers on digital libraries and related topics. Detailed information about IRCDL can be found at the conference home page (<http://ims.dei.unipd.it/ircdl/home.html>), which contains links to the past editions, and also at the DBLP web site, which contains information on proceedings and presented papers (<http://www.informatik.uni-trier.de/~ley/db/conf/ircdl/>).

We should like here to thank those institutions and individuals who have made this conference possible: the Programme Committee members, the Department of Information Engineering of the University of Padova, the members of the same department who have contributed to the organization of the event, namely Maria Bernini, Emanuele Di Buccio, Marco Dussin, Riccardo Miotto, and Gianmaria Silvello, and the members of the University of Padova Library Centre who have contributed to the organization of the registration and management of the on-line presentations and publications, namely Yuri Carrer, and Ornella Volpato. Finally, we take this opportunity to thank also the additional reviewers who have helped in the revision of the final papers, and who have contributed to the improvement of the papers presented in this volume.

To conclude, we would like to point out that, in addition to the enthusiastic participation of the young researchers and the good will of the members of the various committees, much of the credit for having today the IRCDL series of conferences goes to DELOS. As a matter of fact, DELOS started its activities more than ten years ago as a working group under the ESPRIT Program, then continued as a Thematic Network under the 5th Framework Program and after that went on as a Network of Excellence under the 6th Framework Program. It is generally recognized that during these years DELOS has given a substantial contribution to the establishment in Europe of a research community on Digital Libraries. At the end of 2007 the funding of the DELOS Network of Excellence has come to an end. In order to keep the “DELOS spirit” alive, a DELOS Association has been established as a not-for-profit organization, with the main aim of continuing as much as possible the DELOS activities by promoting research activities in the field of digital libraries. In this vein, there is also the commitment of supporting the new edition of IRCDL, which, as customary, will be held in Padova in January 2010. A Call for Participation will be circulated, but meanwhile we invite all researchers (Italian or non Italian, young or not so young), having research interests in Digital Libraries, to start thinking about possible contributions to the next year conference.

Maristella Agosti, Floriana Esposito and Costantino Thanos  
Editors of the Proceedings of the 5<sup>th</sup> IRCDL

July 2009

# Organization and Support

## General Chair

Costantino Thanos, ISTI CNR, Pisa

## Program Chair

Maristella Agosti, University of Padova

## Program Committee

Giuseppe Amato, ISTI CNR, Pisa  
Giorgetta Bonfiglio Dosio, University of Padova  
Donatella Castelli, ISTI CNR, Pisa  
Tiziana Catarci, University of Rome "La Sapienza"  
Alberto Del Bimbo, University of Florence  
Floriana Esposito, University of Bari  
Nicola Ferro, University of Padova  
Maria Guercio, University of Urbino "Carlo Bo"  
Nicola Orio, University of Padova  
Fausto Rabitti, ISTI CNR, Pisa  
Pasquale Savino, ISTI CNR, Pisa  
Carlo Tasso, University of Udine

## Local Organization - University of Padova

### Department of Information Engineering:

Maria Bernini  
Emanuele Di Buccio  
Marco Dussin  
Ivano Masiero  
Riccardo Miotto  
Gianmaria Silvello

### Library Centre - CAB:

Yuri Carrer  
Ornella Volpato

## Additional reviewers

Stefano Ferilli, University of Bari  
Raffaele Perego, ISTI CNR, Pisa  
Fabrizio Sebastiani, ISTI CNR, Pisa  
Giuseppe Serra, University of Florence

## **Supporting Institutions**

IRCDL 2009 benefited from the support of the following organizations:

- DELOS Association
- Institute for Information Science and Technologies of the Italian National Research Council (ISTI-CNR), Pisa, Italy
- Department of Information Engineering, University of Padova, Italy.





## Table of Contents

### Models for Digital Libraries

On Foundations of Typed Data Models for Digital Libraries .....	1
<i>L. Candela, D. Castelli, P. Manghi, M. Mikulicic, P. Pagano</i>	
Design and Development of the Data Model of a Distributed DLS Architecture for Archive Metadata .....	12
<i>N. Ferro, G. Silvello</i>	
Combining Qualitative and Quantitative Keyword Extraction Methods with Document Layout Analysis .....	22
<i>S. Ferilli, M. Biba, T.M.A. Basile, F. Esposito</i>	
Handling Evolution in Digital Libraries .....	34
<i>A. Baruzzo, P. Casoto, A. Dattolo, C. Tasso</i>	

### Content Description

WibNED Wikipedia Based Named Entity Disambiguation .....	51
<i>A.L. Gentile, P. Basile, G. Semeraro</i>	
A Continuous Language Modelling Approach for Assessing Real-valued Attributes of Documents .....	60
<i>R. Bache, F. Crestani</i>	
CLEF Ad-hoc: A Perspective on the Evolution of the Cross-Language Evaluation Forum .....	72
<i>N. Ferro, C. Peters</i>	
Towards an Integrated Approach to Music Retrieval .....	80
<i>E. Di Buccio, I. Masiero, Y. Mass, M. Melucci, R. Miotto, N. Orio, B. Sznajder</i>	

### Information Access

Searching 100M Images by Content Similarity .....	88
<i>P. Bolettieri, F. Falchi, C. Lucchese, Y. Mass, R. Perego, F. Rabitti, M. Shmueli-Scheuer</i>	
Design of an Information Retrieval System Based on the Peer-to-Peer Paradigm: An Application to Music Retrieval .....	100
<i>E. Di Buccio, N. Ferro, M. Melucci, R. Miotto, N. Orio</i>	
A Hybrid Strategy for Italian Word Sense Disambiguation .....	108
<i>P. Basile, M. de Gemmis, P. Lops, G. Semeraro</i>	
Searching and Browsing Digital Library Catalogues: A Combined Log Analysis for The European Library .....	120
<i>M. Agosti, F. Crivellari, G.M. Di Nunzio, Y. Ioannidis, E. Stamatogiannakis, M.L. Triantafyllidi, M. Vayanou</i>	

### Relevant Projects Presentations

The On-TIME Project .....	136
<i>T. Catarci, A. Dix, R. Giuliano, M. Piva, A. Poggi, F. Terella, E. Tracanna</i>	

Major Preservation Projects under the 6 <sup>th</sup> Framework Program .....	143
<i>V. Casarosa</i>	
Digital Repository Infrastructure Vision for European Research – DRIVER .....	150
<i>S. Jones, P. Manghi</i>	
Europeana: Towards The European Digital Library .....	154
<i>N. Aloia, C. Concordia, C. Meghini</i>	
TrebleCLEF: Evaluation, Best Practices and Collaboration for Multilingual Access .....	158
<i>C. Peters</i>	
MultiMatch: Multilingual/Multimedia Access to Cultural Heritage .....	162
<i>G. Amato, F. Debole, C. Peters, P. Savino</i>	
D4Science: an e-Infrastructure for Supporting Virtual Research Environments .....	166
<i>L. Candela, D. Castelli, P. Pagano</i>	
TELplus: Aimed at Strengthening, Extending and Improving The European Library Service .....	170
<i>M. Agosti</i>	
<b>Author Index</b> .....	<b>175</b>