

Communications
in Computer and Information Science

91

Maristella Agosti Floriana Esposito
Costantino Thanos (Eds.)

Digital Libraries

6th Italian Research Conference, IRCDL 2010
Padua, Italy, January 28-29, 2010
Revised Selected Papers

Volume Editors

Maristella Agosti
Università degli Studi di Padova
Padua, Italy
E-mail: maristella.agosti@unipd.it

Floriana Esposito
Università di Bari
Bari, Italy
E-mail: esposito@di.uniba.it

Costantino Thanos
Istituto di Scienza e Tecnologie
dell'Informazione "Alessandro Faedo"
Pisa, Italy
E-mail: costantino.thanos@isti.cnr.it

Library of Congress Control Number: Applied for

CR Subject Classification (1998): H.3, H.5, H.4, J.1, H.2, H.2.8

ISSN	1865-0929
ISBN-10	3-642-15849-8 Springer Berlin Heidelberg New York
ISBN-13	978-3-642-15849-0 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 2010
Printed in Germany

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

Preface

This volume contains the revised accepted papers from among those presented at the 6th Italian Research Conference on Digital Libraries (IRCDL 2010), which was held in the Department of Information Engineering of the University of Padua, Italy, during January 28–29, 2010.

The well-established aim of IRCDL is to bring together Italian researchers interested in the different methods and techniques that allow the building and operation of Digital Libraries. A national Program Committee was set up composed of 15 members, with representatives of the most active Italian research groups on Digital Libraries.

Seventeen of the papers presented at the conference were accepted for inclusion in this volume. Selected authors submitted expanded versions of their workshop papers. Those papers were reviewed again and the results of the selection are the papers appearing in these proceedings. The covered topics testifying to the bright interests of the community are:

- System Interoperability and Data Integration including emerging interoperability issues and technologies promoting interoperability
- Infrastructures, Metadata Creation and Content Management for Digital Libraries, such as methods and techniques for text summarization and key phrase extraction, ontology-based annotation, event-centric provenance models, semantic relatedness
- Information Access and Search for Digital Library Systems from mathematical symbol indexing to term-based text retrieval and disambiguation techniques, from audio-based information retrieval to video content-based identification, classification and retrieval
- User Interfaces for Digital Libraries including interactive visual representations, collaborative user interfaces, multimodal user interface for multimedia Digital Libraries, personalization models

The volume also contains the reports on the two keynote addresses and a communication on a relevant national project.

This year is the first time that the IRCDL proceedings have been published in the Springer *CCIS* series.

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 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 ongoing research, to be presented in a friendly and relaxed atmosphere, to facilitate constructive discussion and exchange of opinions.

Thanks to the initial support of DELOS, and later on to the support of both the DELOS Association and the Department of Information Engineering of the University of Padua, IRCDL continued in the subsequent years and has become a fixture 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 also contains links to IRCDL editions.

Here we would like to thank those institutions and individuals who made this conference possible: the Program Committee members, the Department of Information Engineering of the University of Padua, the members of the same department who contributed to the organization of the event, namely, Maria Bernini, Emanuele Di Buccio, Marco Dussin, and Nicola Montecchio, and the members of the University of Padua Library Centre who contributed to the organization of the registration and management of the on-line presentations, namely, Yuri Carrer, Francesca Moro, and Ornella Volpato. Finally, we take this opportunity to also thank Fabrizio Falchi of ISTI CNR of Pisa who helped in the revision of the final papers.

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 the IRCDL series of conferences today 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 Fifth Framework Program and after that as a Network of Excellence under the Sixth Framework Program. It is generally recognized that during these years DELOS has made 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 came to an end. In order to keep the “DELOS spirit” alive, a DELOS Association was 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 to supporting the new edition of IRCDL, which, as customary, will be held in January 2011.

A call for participation for IRCDL 2011 will be circulated, but meanwhile we invite all researchers having research interests in Digital Libraries to start thinking about possible contributions to next year’s conference.

June 2010

Maristella Agosti
Floriana Esposito
Costantino Thanos

Organization

General Chair

Costantino Thanos	ISTI CNR, Pisa
-------------------	----------------

Program Chairs

Maristella Agosti	University of Padua
Floriana Esposito	University of Bari

Program Committee

Giuseppe Amato	ISTI CNR, Pisa
Marco Bertini	University of Florence
Leonardo Candela	ISTI CNR, Pisa
Tiziana Catarci	University of Rome “La Sapienza”
Alberto Del Bimbo	University of Florence
Stefano Ferilli	University of Bari
Nicola Ferro	University of Padua
Maria Guercio	University of Urbino “Carlo Bo”
Carlo Meghini	ISTI CNR, Pisa
Nicola Orio	University of Padua
Fausto Rabitti	ISTI CNR, Pisa
Pasquale Savino	ISTI CNR, Pisa
Anna Maria Tammaro	University of Parma
Letizia Tanca	Politecnico di Milano
Carlo Tasso	University of Udine

Organizing Committee

Department of Information Engineering, University of Padua:

Maria Bernini
Emanuele Di Buccio
Marco Dussin
Nicola Montecchio

Library Centre - CAB, University of Padua:

Yuri Carrer
Francesca Moro
Ornella Volpato

Supporting Institutions

IRCDL 2010 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 Padua, Italy

Table of Contents

Keynote Addresses

Digital Cultural Content: National and European Projects and Strategies	1
<i>Rossella Caffo</i>	
Archival Information Systems in Italy and the National Archival Portal	5
<i>Stefano Vitali</i>	

System Interoperability and Data Integration

Making Digital Library Content Interoperable	13
<i>Leonardo Candela, Donatella Castelli, and Costantino Thanos</i>	
Integrating a Content-Based Recommender System into Digital Libraries for Cultural Heritage	27
<i>Cataldo Musto, Fedelucio Narducci, Pasquale Lops, Marco de Gemmis, and Giovanni Semeraro</i>	

Infrastructures, Metadata Creation and Management

Digital Stacks: Turning a Current Prototype into an Operational Service	39
<i>Giovanni Bergamin and Maurizio Messina</i>	
A First National Italian Register for Digital Resources for Both Cultural and Scientific Communities (Communication)	47
<i>Maurizio Lunghi</i>	
FAST and NESTOR: How to Exploit Annotation Hierarchies	55
<i>Nicola Ferro and Gianmaria Silvello</i>	
A New Domain Independent Keyphrase Extraction System	67
<i>Nirmala Pudota, Antonina Dattolo, Andrea Baruzzo, and Carlo Tasso</i>	
An Event-Centric Provenance Model for Digital Libraries	79
<i>Donatella Castelli, Leonardo Candela, Paolo Manghi, Pasquale Pagano, Cristina Tang, and Costantino Thanos</i>	

A Digital Library Effort to Support the Building of Grammatical Resources for Italian Dialects	89
<i>Maristella Agosti, Paola Benincà, Giorgio Maria Di Nunzio, Riccardo Miotto, and Diego Pescarini</i>	

Representation, Indexing and Retrieval in Digital Libraries

Interactive Visual Representations of Complex Information Structures	101
<i>Gianpaolo D'Amico, Alberto Del Bimbo, and Marco Meoni</i>	
Mathematical Symbol Indexing for Digital Libraries	113
<i>Simone Marinai, Beatrice Miotti, and Giovanni Soda</i>	
Using Explicit Word Co-occurrences to Improve Term-Based Text Retrieval	125
<i>Stefano Ferilli, Marenglen Biba, Teresa M.A. Basile, and Floriana Esposito</i>	
Semantic Relatedness Approach for Named Entity Disambiguation	137
<i>Anna Lisa Gentile, Ziqi Zhang, Lei Xia, and José Iria</i>	
Merging Structural and Taxonomic Similarity for Text Retrieval Using Relational Descriptions	149
<i>Stefano Ferilli, Marenglen Biba, Nicola Di Mauro, Teresa M.A. Basile, and Floriana Esposito</i>	

Handling Audio-Visual and Non-traditional Objects

Audio Objects Access: Tools for the Preservation of the Cultural Heritage	161
<i>Sergio Canazza and Nicola Orio</i>	
Toward Conversation Retrieval	173
<i>Matteo Magnani and Danilo Montesì</i>	
Improving Classification and Retrieval of Illuminated Manuscript with Semantic Information	183
<i>Costantino Grana, Daniele Borghesani, and Rita Cucchiara</i>	
Content-Based Cover Song Identification in Music Digital Libraries	195
<i>Riccardo Miotto, Nicola Montecchio, and Nicola Orio</i>	
Toward an Audio Digital Library 2.0: Smash, a Social Music Archive of SHellac Phonographic Discs	205
<i>Sergio Canazza and Antonina Dattolo</i>	
Author Index	219