

Costantino Thanos Francesca Borri
Leonardo Candela (Eds.)

Digital Libraries: Research and Development

First International DELOS Conference
Pisa, Italy, February 13-14, 2007
Revised Selected Papers

Volume Editors

Costantino Thanos

Francesca Borri

Leonardo Candela

Consiglio Nazionale delle Ricerche

Istituto di Scienza e Tecnologie dell'Informazione

Via Moruzzi, 1, 56124, Pisa, Italy

E-mail: {costantino.thanos,francesca.borri,leonardo.candela}@isti.cnr.it

Library of Congress Control Number: 2007940368

CR Subject Classification (1998): H.2, H.4, H.3, H.2.4, H.5

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

ISSN 0302-9743

ISBN-10 3-540-77087-9 Springer Berlin Heidelberg New York

ISBN-13 978-3-540-77087-9 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 is a part of Springer Science+Business Media

springer.com

© Springer-Verlag Berlin Heidelberg 2007

Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 12199099 06/3180 5 4 3 2 1 0

Preface

Digital libraries represent the meeting point of a large number of technical areas within the field of informatics, i.e., information retrieval, document management, information systems, the Web, image processing, artificial intelligence, human – computer interaction, mass-storage systems, and others. Moreover, digital libraries draw upon other disciplines beyond informatics, such as library sciences, museum sciences, archives, sociology, psychology, etc. However, they constitute a relatively young scientific field, whose life spans roughly the last 15 years. During these years the DELOS Network of Excellence on Digital Libraries (<http://www.delos.info>) has represented a constant presence aiming to contribute to the consolidation of the field.

The activities of DELOS started many years ago, with the “DELOS Working Group” at the end of the 1990s, and the DELOS Thematic Network, under the Fifth Framework Program, from 2001 to 2003. Since the beginning, the main objective of DELOS has been to advance the state of the art in the field of digital libraries by coordinating the efforts of the major European research teams conducting activities in the major fields of interest.

Every year DELOS organizes the All-Tasks meeting, the annual appointment of the DELOS community where the scientific results achieved during the previous year are presented. Instead of the usual status reports from the various DELOS Tasks, in 2006 it was decided to issue a Call for Papers, soliciting papers reporting the scientific achievements of the DELOS members during 2006, and to organize a conference. The first DELOS conference was held during February 13–14, 2007, at the Grand Hotel Continental in Tirrenia, Pisa, Italy.

The conference represented a good coverage of the 27 research tasks in DELOS, with 38 papers being presented. In addition, two invited papers: “Semantic Digital Libraries” (John Mylopoulos, University of Toronto) and “Digital Libraries: From Proposals to Projects to Systems to Theory to Curricula” (Edward Fox, Virginia Tech) completed the program. The conference was open to the larger digital library community and not just to the DELOS partners. About 120 people attended, half of whom were not from DELOS partners. We believe that this is an indication of the increased interest in digital libraries, and the recognition that the DELOS research activities have played and are playing an important role in the European digital library scene.

This volume includes extended and revised versions of the papers presented at the DELOS Conference. We believe that it should be of interest to a broad audience potentially interested in the digital library research area. It has been structured into 10 sections, corresponding to the different sessions in which the

conference was structured, which in turn correspond to the major areas of research where DELOS has focussed its attention recently.

October 2007

Costantino Thanos
Francesca Borri
Leonardo Candela

Organization

General Chair

Costantino Thanos (Chair) Istituto di Scienza e Tecnologie dell'Informazione
"A. Faedo" - Consiglio Nazionale delle Ricerche -
Pisa, Italy

Scientific Committee

Vittore Casarosa Istituto di Scienza e Tecnologie dell'Informazione
"A. Faedo" - Consiglio Nazionale delle Ricerche -
Pisa, Italy

Tiziana Catarci Dipartimento di Informatica e Sistemistica
"A. Ruberti" - Università di Roma "La Sapienza"
- Rome, Italy

Stavros Christodoulakis Laboratory of Distributed Multimedia Information
System - Technical University of Crete - Crete,
Greece

Alberto del Bimbo Dipartimento di Sistemi e Informatica - Facoltà di
Ingegneria University of degli Studi di Firenze -
Florence, Italy

Norbert Fuhr Department of Computational and Cognitive
Sciences - Faculty of Engineering Sciences of the
University of Duisburg-Essen - Duisburg, Germany

Yannis Ioannidis Department of Informatics - National and
Kapodistrian University of Athens - Athens,
Greece

Bruno Le Dantec The European Research Consortium for Informat-
ics and Mathematics - Sophia Antipolis, France

Liz Lyon UKOLN - University of Bath - Bath, UK

Seamus Ross Humanities Advanced Technology and Information
Institute - University of Glasgow - Glasgow, UK

Hans-Jörg Schek Database & Information Systems Group -
Universität Konstanz - Konstanz, Germany

Organizing Committee

Francesca Borri Istituto di Scienza e Tecnologie dell'Informazione
"A. Faedo" - Consiglio Nazionale delle Ricerche -
Pisa, Italy

Alessandro Launaro Istituto di Scienza e Tecnologie dell'Informazione
"A. Faedo" - Consiglio Nazionale delle Ricerche -
Pisa, Italy

Table of Contents

Similarity Search

MESSIF: Metric Similarity Search Implementation Framework	1
<i>Michal Batko, David Novak, and Pavel Zezula</i>	
Image Indexing and Retrieval Using Visual Terms and Text-Like Weighting	11
<i>Giuseppe Amato, Pasquale Savino, and Vanessa Magionami</i>	

Architectures

A Reference Architecture for Digital Library Systems: Principles and Applications	22
<i>Leonardo Candela, Donatella Castelli, and Pasquale Pagano</i>	
DelosDLMS - The Integrated DELOS Digital Library Management System	36
<i>Maristella Agosti, Stefano Berretti, Gert Brettlecker, Alberto Del Bimbo, Nicola Ferro, Norbert Fuhr, Daniel Keim, Claus-Peter Klas, Thomas Lidy, Diego Milano, Moira Norrie, Paola Ranaldi, Andreas Rauber, Hans-Jörg Schek, Tobias Schreck, Heiko Schuldt, Beat Signer, and Michael Springmann</i>	
ISIS and OSIRIS: A Process-Based Digital Library Application on Top of a Distributed Process Support Middleware	46
<i>Gert Brettlecker, Diego Milano, Paola Ranaldi, Hans-Jörg Schek, Heiko Schuldt, and Michael Springmann</i>	
An Architecture for Sharing Metadata Among Geographically Distributed Archives	56
<i>Maristella Agosti, Nicola Ferro, and Gianmaria Silvello</i>	
Integration of Reliable Sensor Data Stream Management into Digital Libraries	66
<i>Gert Brettlecker, Heiko Schuldt, Peter Fischer, and Hans-Jörg Schek</i>	

Personalization

Content-Based Recommendation Services for Personalized Digital Libraries	77
<i>G. Semeraro, P. Basile, M. de Gemmis, and P. Lops</i>	

Integrated Authoring, Annotation, Retrieval, Adaptation, Personalization, and Delivery for Multimedia	87
<i>Horst Eidenberger, Susanne Boll, Stavros Christodoulakis, Doris Divotkey, Klaus Leopold, Alessandro Martin, Andrea Perego, Ansgar Scherp, and Chrisa Tsinaraki</i>	
Gathering and Mining Information from Web Log Files	104
<i>Maristella Agosti and Giorgio Maria Di Nunzio</i>	

Interoperability

Modelling Intellectual Processes: The FRBR - CRM Harmonization . . .	114
<i>Martin Doerr and Patrick LeBoeuf</i>	
XS2OWL: A Formal Model and a System for Enabling XML Schema Applications to Interoperate with OWL-DL Domain Knowledge and Semantic Web Tools	124
<i>Chrisa Tsinaraki and Stavros Christodoulakis</i>	
A Framework and an Architecture for Supporting Interoperability Between Digital Libraries and eLearning Applications	137
<i>Polyxeni Arapi, Nektarios Moutzidis, Manolis Mylonakis, and Stavros Christodoulakis</i>	

Evaluation

An Experimental Framework for Interactive Information Retrieval and Digital Libraries Evaluation	147
<i>Claus-Peter Klas, Sascha Kriewel, and Norbert Fuhr</i>	
The Importance of Scientific Data Curation for Evaluation Campaigns	157
<i>Maristella Agosti, Giorgio Maria Di Nunzio, and Nicola Ferro</i>	
An Approach for the Construction of an Experimental Test Collection to Evaluate Search Systems that Exploit Annotations	167
<i>Maristella Agosti, Tullio Coppotelli, Nicola Ferro, and Luca Pretto</i>	
Evaluation and Requirements Elicitation of a DL Annotation System for Collaborative Information Sharing	177
<i>Preben Hansen, Annelise Mark Pejtersen, and Hanne Albrechtsen</i>	
INEX 2002 - 2006: Understanding XML Retrieval Evaluation	187
<i>Mounia Lalmas and Anastasios Tombros</i>	

Miscellaneous

Task-Centred Information Management	197
<i>Tiziana Catarci, Alan Dix, Akrivi Katifori, Giorgios Lepouras, and Antonella Poggi</i>	
Viewing Collections as Abstractions	207
<i>Carlo Meghini and Nicolas Spyratos</i>	
Adding Multilingual Information Access to the European Library	218
<i>Martin Braschler and Nicola Ferro</i>	
The OntoNL Framework for Natural Language Interface Generation and a Domain-Specific Application	228
<i>Anastasia Karanastasi, Alexandros Zotos, and Stavros Christodoulakis</i>	

Preservation

Evaluating Preservation Strategies for Electronic Theses and Dissertations	238
<i>Stephan Strodl, Christoph Becker, Robert Neumayer, Andreas Rauber, Eleonora Nicchiarelli Bettelli, Max Kaiser, Hans Hofman, Heike Neuroth, Stefan Strathmann, Franca Debole, and Giuseppe Amato</i>	
Searching for Ground Truth: A Stepping Stone in Automating Genre Classification	248
<i>Yunhyong Kim and Seamus Ross</i>	

Video Data Management

Video Transcoding and Streaming for Mobile Applications	262
<i>Giovanni Gualdi, Andrea Prati, and Rita Cucchiara</i>	
Prototypes Selection with Context Based Intra-class Clustering for Video Annotation with Mpeg7 Features	268
<i>Costantino Grana, Roberto Vezzani, and Rita Cucchiara</i>	
Automatic, Context-of-Capture-Based Categorization, Structure Detection and Segmentation of News Telecasts	278
<i>Arne Jacobs, George T. Ioannidis, Stavros Christodoulakis, Nektarios Moumoutzis, Stratos Georgoulakis, and Yiannis Papachristoudis</i>	

3D Objects

Description, Matching and Retrieval by Content of 3D Objects	288
<i>S. Berretti, A. Del Bimbo, and P. Pala</i>	
3D-Mesh Models: View-Based Indexing and Structural Analysis	298
<i>Mohamed Daoudi, Tarik Filali Ansary, Julien Tierny, and Jean-Philippe Vandeborre</i>	
Similarity-Based Retrieval with MPEG-7 3D Descriptors: Performance Evaluation on the Princeton Shape Benchmark	308
<i>Costantino Grana, Matteo Davolio, and Rita Cucchiara</i>	

Peer to Peer

Application of the Peer-to-Peer Paradigm in Digital Libraries	318
<i>Stratis D. Viglas, Theodore Dalamagas, Vassilis Christophides, Timos Sellis, and Aggeliki Dimitriou</i>	
Efficient Search and Approximate Information Filtering in a Distributed Peer-to-Peer Environment of Digital Libraries	328
<i>Christian Zimmer, Christos Tryfonopoulos, and Gerhard Weikum</i>	
Management of and Access to Virtual Electronic Health Records	338
<i>M. Springmann, L. Bischofs, P.M. Fischer, H.-J. Schek, H. Schuldt, U. Steffens, and R. Vogl</i>	

Author Index	349
------------------------	-----