

Communications in Computer and Information Science

612

Commenced Publication in 2007

Founding and Former Series Editors:

Alfredo Cuzzocrea, Dominik Ślęzak, and Xiaokang Yang

Editorial Board

Simone Diniz Junqueira Barbosa

*Pontifical Catholic University of Rio de Janeiro (PUC-Rio),
Rio de Janeiro, Brazil*

Phoebe Chen

La Trobe University, Melbourne, Australia

Xiaoyong Du

Renmin University of China, Beijing, China

Joaquim Filipe

Polytechnic Institute of Setúbal, Setúbal, Portugal

Orhun Kara

TÜBİTAK BİLGEM and Middle East Technical University, Ankara, Turkey

Igor Kotenko

*St. Petersburg Institute for Informatics and Automation of the Russian
Academy of Sciences, St. Petersburg, Russia*

Ting Liu

Harbin Institute of Technology (HIT), Harbin, China

Krishna M. Sivalingam

Indian Institute of Technology Madras, Chennai, India

Takashi Washio

Osaka University, Osaka, Japan

More information about this series at <http://www.springer.com/series/7899>

Diego Calvanese · Dario De Nart
Carlo Tasso (Eds.)

Digital Libraries on the Move

11th Italian Research Conference
on Digital Libraries, IRCDL 2015
Bolzano, Italy, January 29–30, 2015
Revised Selected Papers

Editors

Diego Calvanese
Fac. di Scienze e Tecnologie Informatiche
Libera Univ. di Bolzano
Bolzano
Italy

Dario De Nart
Dip. di Scienze Matematiche, Informatiche
e Fisiche
Università degli Studi di Udine
Udine
Italy

Carlo Tasso
Dip. di Scienze Matematiche, Informatiche
e Fisiche
Università degli Studi di Udine
Udine
Italy

ISSN 1865-0929

ISSN 1865-0937 (electronic)

Communications in Computer and Information Science

ISBN 978-3-319-41937-4

ISBN 978-3-319-41938-1 (eBook)

DOI 10.1007/978-3-319-41938-1

Library of Congress Control Number: 2016943797

© Springer International Publishing Switzerland 2016

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

This Springer imprint is published by Springer Nature

The registered company is Springer International Publishing AG Switzerland

Preface

Since 2005, the Italian Research Conference on Digital Libraries (IRCDL) has provided a great opportunity for Italian researchers in the field of digital libraries to present and discuss their current research activities and to envision together further developments. The 2015 edition of IRCDL continued the traditional emphasis on the multidisciplinary nature of the research on digital libraries, which has been characterizing the conference over the years.

Digital libraries are “on the move” – they experience today a significant, almost revolutionary development, which can be seen in various ways in their numerous multidisciplinary aspects, ranging from computer science to humanities in the broader sense, including research areas such as archival and library information sciences, information management systems, semantic technologies, information retrieval, new knowledge environments, and new organizational/business models. This is a tremendous challenge for the digital libraries field and there is a strong need to improve the cooperation between the many communities that share common objectives. Moreover, deep changes are affecting the world of scientific communication, where new non-linear but object-centric products are progressively taking shape, and new problem areas emerge, requiring for the integration, publication, and preservation of research data.

The aim of the IRCDL Conference series has always been to provide attendees with the opportunity to explore new ideas, techniques, and tools, and to exchange experiences also from on-going projects. The IRCDL series was launched and initially sponsored by DELOS, an EU FP6 Network of Excellence on digital libraries, together with the Department of Information Engineering of the University of Padua. Over the years, IRCDL has become a self-sustainable event that is supported by the Italian digital libraries community.

The 2015 edition of IRCDL was organized with the patronage and the sponsorship of the Faculty of Computer Science and the Library of the Free University of Bozen-Bolzano.

This volume contains the reports on the invited presentations and revised versions of a selection of the accepted papers presented at the conference. After the presentation, the papers were resubmitted by the authors in a revised version that included the suggestions received during the presentation at the conference. The peer-review process was performed under the responsibility of the Scientific Committee of IRCDL 2015.

The topics included in the papers cover many aspects of digital libraries, which were organized under the following five categories:

- Semantic Modelling
- Projects
- Models and Applications
- Content Analysis
- Digital Libraries Infrastructures.

The program of IRCDL 2015 also included two invited events: Dr. Antonio Gulli from Elsevier presented a very interesting talk on “Smart Data for the Publishing Industry,” and Prof. Maristella Agosti from the University of Padua was invited to chair a Special Session devoted to the first 10 years of the IRCDL Conference series.

We would like to thank those institutions and individuals who made the conference and this volume possible. More specifically, we thank the Program Committee members and the additional reviewers, the honorary chair, the Steering Committee members, the authors, the local Organizing Committee, the Faculty of Computer Science, and the University Library of the Free University of Bozen-Bolzano.

April 2016

Diego Calvanese
Dario De Nart
Carlo Tasso

Organization

IRCDL 2015 was organized by the Faculty of Computer Science and the University Library of the Free University of Bozen-Bolzano.

Honorary Chair

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

Program Chairs

Diego Calvanese	Free University of Bozen-Bolzano, Italy
Carlo Tasso	University of Udine, Italy

IRCDL Steering Committee

Maristella Agosti	University of Padua, Italy
Tiziana Catarci	Sapienza University of Rome, Italy
Alberto Del Bimbo	University of Florence, Italy
Floriana Esposito	University of Bari, Italy
Carlo Tasso	University of Udine, Italy
Costantino Thanos	ISTI CNR, Pisa, Italy

Program Committee

Stefano Berretti	University of Florence, Italy
Marco Bertini	University of Florence, Italy
Maria Teresa Biagetti	Sapienza University of Rome, Italy
Rossella Caffo	ICCU Rome, Italy
Vittore Casarosa	ISTI CNR, Pisa, Italy
Michelangelo Ceci	University of Bari, Italy
Fabio Ciotti	University of Rome 2, Italy
Rita Cucchiara	University of Modena and Reggio Emilia, Italy
Nicola Di Mauro	University of Bari, Italy
Stefano Ferilli	University of Bari, Italy
Nicola Ferro	University of Padua, Italy
Costantino Grana	University of Modena and Reggio Emilia, Italy
Maria Guercio	Sapienza University of Rome, Italy
Donato Malerba	University of Padua, Italy
Paolo Manghi	ISTI CNR, Pisa, Italy
Simone Marinai	University of Florence, Italy
Carlo Meghini	ISTI CNR, Pisa, Italy
Maurizio Messina	Biblioteca Nazionale Marciana, Venice, Italy
Antonella Poggi	Sapienza University of Rome, Italy

Marco Schaerf	Sapienza University of Rome, Italy
Gianmaria Silvello	University of Padua, Italy
Anna Maria Tammaro	University of Parma, Italy
Francesca Tomasi	University of Bologna, Italy
Paul G. Weston	University of Pavia, Italy

Organizing Committee

Diego Calvanese	Free University of Bozen-Bolzano, Italy
Luigi Siciliano	University Library, Free University of Bozen-Bolzano, Italy
Hannes Hell	Press and Event Management, Free University of Bozen-Bolzano, Italy

Smart Data for the Publishing Industry (Invited Talk)

Antonio Gulli

Elsevier, Amsterdam, Netherlands
a.gulli@elsevier.com

Abstract. Smart Data is a central element for the Publishing Industry. A review of the rich data published by Elsevier via 9 Research platforms will be presented. Particular emphasis will be devoted to our investments in Knowledge graphs, Information Extraction, and Entity annotation. After this initial introduction, we will focus our attention on three particular applications built on top of our rich data. Authors' disambiguation is an open research problem where the goal is to match different identities attributed to the same Author during his/her life. Article fingerprinting is a rich set of information extraction techniques used for classifying articles and for building analytical tools leveraged by research institutions. News Trend detection is another rich set of matching tools used for tracking how institutions and researchers are mentioned by public and social media.

Keywords: Data publishing • Publishing industry • Information extraction

Contents

Invited Talk

10 Years of IRCDL: 2005–2014 (<i>Invited Paper</i>)	3
<i>Maristella Agosti</i>	

Semantic Modeling

Description Logics for Documentation	11
<i>Carlo Meghini</i>	
Towards a Semantic Web Enabled Representation of DL Foundational Models: The Quality Domain Example	24
<i>Nicola Ferro and Gianmaria Silvello</i>	
A Semantic Model for Content Description in the Sapienza Digital Library . . .	36
<i>Angela Di Iorio and Marco Schaerf</i>	
Text Encoding Initiative Semantic Modeling. A Conceptual Workflow Proposal	48
<i>Fabio Ciotti, Marilena Daquino, and Francesca Tomasi</i>	
Structured Descriptions of Roles, Activities, and Procedures in the Roman Constitution	61
<i>Yoonmi Chu and Robert B. Allen</i>	

Projects

The AstroBID: Preserving and Sharing the Italian Astronomical Heritage. . . .	71
<i>Mauro Gargano, Antonella Gasperini, Emilia Olostro Cirella, Riccardo Smareglia, and Valeria Zanini</i>	
The EAGLE Europeana Network of Ancient Greek and Latin Epigraphy: A Technical Perspective.	75
<i>Andrea Mannocci, Vittore Casarosa, Paolo Manghi, and Franco Zoppi</i>	
The TRAME Project – Text and Manuscript Transmission of the Middle Ages in Europe.	79
<i>Emiliano Degl’Innocenti, Alfredo Cosco, Fabrizio Butini, Roberta Giacomi, and Vinicio Serafini</i>	

The PREFORMA Project: Federating Memory Institutions for Better Compliance of Preservation Formats	86
<i>Linda Cappellato, Nicola Ferro, Antonella Fresa, Magnus Geber, Börje Justrell, Bert Lemmens, Claudio Prandoni, and Gianmaria Silvello</i>	

Models and Applications

Keep, Change or Delete? Setting up a Low Resource OCR Post-correction Framework for a Digitized Old Finnish Newspaper Collection	95
<i>Kimmo Kettunen</i>	
Collaborative Information Seeking with Ant Colony Ranking in Real-Time . . .	104
<i>Tommaso Turchi, Alessio Malizia, Paola Castellucci, and Kai Olsen</i>	
Modeling the Concept of Movie in a Software Architecture for Film-Induced Tourism	116
<i>Giulia Lavarone, Nicola Orio, Farah Polato, and Sandro Savino</i>	

Content Analysis

Unsupervised Author Identification and Characterization	129
<i>Stefano Ferilli, Domenico Redavid, and Floriana Esposito</i>	
A Content-Based Approach to Social Network Analysis: A Case Study on Research Communities	142
<i>Dario De Nart, Dante Degl'Innocenti, Marco Basaldella, Maristella Agosti, and Carlo Tasso</i>	
Analysis and Re-Use of Videos in Educational Digital Libraries with Automatic Scene Detection	155
<i>Lorenzo Baraldi, Costantino Grana, and Rita Cucchiara</i>	

Digital Library Infrastructures

An Interoperability Infrastructure for Digital Identifiers in e-Science	167
<i>Barbara Bazzanella and Paolo Bouquet</i>	
Finding a Needle in a Haystack: The BEIC Digital Library in Search of Its Space on the Web: A Case Study	179
<i>Chiara Consonni and Paul Gabriele Weston</i>	

Author Index	191
-------------------------------	-----