

Communications in Computer and Information Science


1177

Commenced Publication in 2007

Founding and Former Series Editors:

Phoebe Chen, Alfredo Cuzzocrea, Xiaoyong Du, Orhun Kara, Ting Liu,
Krishna M. Sivalingam, Dominik Ślęzak, Takashi Washio, Xiaokang Yang,
and Junsong Yuan

Editorial Board Members

Simone Diniz Junqueira Barbosa 

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

Joaquim Filipe 

Polytechnic Institute of Setúbal, Setúbal, Portugal

Ashish Ghosh

Indian Statistical Institute, Kolkata, India

Igor Kotenko 

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

Lizhu Zhou

Tsinghua University, Beijing, China

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


Michelangelo Ceci · Stefano Ferilli ·
Antonella Poggi (Eds.)

Digital Libraries: The Era of Big Data and Data Science

16th Italian Research Conference
on Digital Libraries, IRCDL 2020
Bari, Italy, January 30–31, 2020
Proceedings

Editors

Michelangelo Ceci 
University of Bari
Bari, Italy

Stefano Ferilli 
University of Bari
Bari, Italy

Antonella Poggi
Sapienza University of Rome
Rome, Italy

ISSN 1865-0929 ISSN 1865-0937 (electronic)
Communications in Computer and Information Science
ISBN 978-3-030-39904-7 ISBN 978-3-030-39905-4 (eBook)
<https://doi.org/10.1007/978-3-030-39905-4>

© Springer Nature Switzerland AG 2020, corrected publication 2020

The chapter “Identifying, Classifying and Searching Graphic Symbols in the NOTAE System” is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>). For further details see license information in the chapter.

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, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The Italian Research Conference on Digital Libraries (IRCDL) is an annual forum for the Italian research community to discuss the research topics pertaining to digital libraries and related technical, practical, and social issues both on the computer science and the humanities side. Since 2005, it has served as a key meeting for the Italian digital library community. During these years, IRCDL has touched upon many of the facets of digital library, adapting its topics of interest to the evolution of the research in this domain and of the whole process of scholarly communication.

The theme of the 2020 edition of the conference was “Big Data and Data Science.” This theme was motivated by the characteristics of data resources of many modern digital libraries, requiring existing library services to use methods that are able to face the main issues of big data: volume, variety, and velocity. This also paves the way to new forms of data analysis from big digital libraries that go far beyond the classical techniques of data analysis from digital libraries adopted in the last decade.

This volume contains 18 selected contributions accepted for presentation at the 16th Italian Research Conference on Digital Libraries (IRCDL 2020), which was held at the Department of Computer Science of the University of Bari Aldo Moro (Bari, Italy) during January 30–31, 2020. These contributions touched a rich range of topics: information retrieval, big data and data science in digital libraries, digital libraries for cultural heritage, and open science. The papers were selected after a rigorous evaluation of a Program Committee composed of 19 members, including representatives of the most active Italian research groups on digital libraries.

We are proud to present a rich scientific program, including high-profile keynotes and many technical presentations.

To conclude, we would like to express our gratitude to all our colleagues for submitting papers to IRCDL 2020 and for helping to finalize this volume in due time. We would like to thank those institutions and individuals who made the conference and this volume possible. In particular, we would like to thank the Program Committee members, the Steering Committee members, the Advisory Board members, and the Organizing Committee. We would also like to thank Artificial Brain S.r.l. for the helpful conference secretariat and the Department of Computer Science of the University of Bari Aldo Moro for supporting and hosting the event.

January 2020

Michelangelo Ceci
Stefano Ferilli
Antonella Poggi

Organization

Program Committee Chairs

Michelangelo Ceci	University of Bari Aldo Moro, Italy
Stefano Ferilli	University of Bari Aldo Moro, Italy
Antonella Poggi	Sapienza University of Rome, Italy

Advisory Board

Tiziana Catarci	Sapienza University of Rome, Italy
Floriana Esposito	University of Bari Aldo Moro, Italy
Costantino Thanos	ISTI-CNR, Italy

Steering Committee

Maristella Agosti	University of Padua, Italy
Alberto Del Bimbo	University of Florence, Italy
Stefano Ferilli	University of Bari Aldo Moro, Italy
Costantino Grana	University of Modena and Reggio Emilia, Italy
Antonella Poggi	Sapienza University of Rome, Italy
Carlo Tasso	University of Udine, Italy

Program Committee

Lorenzo Baraldi	University of Modena and Reggio Emilia, Italy
Nicola Barbuti	University of Bari Aldo Moro, Italy
Valentina Bartalesi	ISTI-CNR, Italy
George Bruseker	FORTH, Greece
Leonardo Candela	ISTI-CNR, Italy
Fabio Ciotti	Tor Vergata University of Rome, Italy
Marilena Daquino	University of Bologna, Italy
Giorgio Maria Di Nunzio	University of Padua, Italy
Achille Felicetti	PIN S.c.R.L., Italy
Nicola Ferro	University of Padua, Italy
Elena Giglia	University of Turin, Italy
Fabio Leuzzi	Polizia di Stato, Italy
Claudio Lucchese	Ca' Foscari University of Venice, Italy
Paolo Manghi	ISTI-CNR, Italy
Nicola Orio	University of Padua, Italy
Eloy Rodrigues	University of Minho, Portugal
Marco Schaerf	Sapienza University of Rome, Italy
Giuseppe Serra	University of Udine, Italy

Luigi Siciliano

Gianmaria Silvello

Anna Maria Tammaro

Fabio Vitali

University of Bolzen-Bolzano, Italy

University of Padua, Italy

Bibelot, Italy

University of Bologna, Italy

Invited Talks

Context-Aware Distributed Mobile Search

Fabio Crestani

Faculty of Informatics, Università della Svizzera Italiana

Abstract. Distributed Information Retrieval (DIR) has been around for a while, yet in this talk I will present an approach to using uncooperative DIR in a mobile environment. We called this Unified Mobile Search (UMS). I will present how this concept extends that of DIR by considering fully the context in which the search is carried out and present our first approach to target application selection for mobile devices. I will then expand on this idea toward a conversational assistant for context-aware distributed mobile search, providing a view on what we should expect coming to our mobile phones very soon.

Short Bio. Fabio Crestani is a full professor at the Faculty of Informatics of the Università' della Svizzera Italiana (USI) in Lugano, Switzerland, since 2007. Previously he was a professor at the University of Strathclyde in Glasgow, UK. He holds a degree in Statistics from the University of Padua (Italy) and a MSc and PhD in Computing Science from the University of Glasgow (UK). Prof. Crestani is an internationally recognized researcher in Information Retrieval, Text Mining, and Digital Libraries. In these areas he has published over two hundred refereed papers on both theoretical and experimental investigations. He has also served in the Organizing and Program Committees of several conferences and the editorial boards of several journals.

Aiuto – ho troppi dati! Digital Library Users and Their Challenges With Big Data

Elke Greifeneder

Institut für Bibliotheks- und Informationswissenschaft,
Humboldt-Universität zu Berlin

Abstract. Digital libraries have many users: the repository holder, the repositories' day to day managers, the users who provide big data for storage and use in the digital library, the researchers as end-user who benefit from large data sets, the industry firm who seeks large, free amounts of data for commercial reuse, and many others. Yet, they all have something in common: the initial enthusiasm for the simple availability of a magic thing called “big data” often changes to an attitude of despair, frustration, and defense within months of working with big data. This presentation uses a personal approach and explains the various needs and practices of those user groups, and depicts their pain points when it comes to operating with big data.

Short Bio. Prof. Elke Greifeneder is a full professor at the Berlin School of Library and Information Science at Humboldt-Universität zu Berlin (Germany). She is currently leading the research group on information behavior and focuses on methodological questions, in particular on empirical research completed in digital environments. She was tenured in Berlin in 2019 and was previously an assistant professor at the University of Copenhagen (Denmark). She has been very active in the field as reviewer and Program Committee member, as well as program chair of the iConference, ASIST chair for International Relations, IFLA section member on library services to people with special needs, and has been the co-editor of Library Hi Tech for 10 years.

Contents

Information Retrieval

Reproducibility of the Neural Vector Space Model via Docker	3
<i>Nicola Ferro, Stefano Marchesin, Alberto Purpura, and Gianmaria Silvello</i>	
Towards a Decision Support Framework for Forensic Analysis of Dynamic Signatures.	9
<i>Daniela Mazzolini, Patrizia Pavan, Giuseppe Pirlo, and Gennaro Vessio</i>	
An Information Visualization Tool for the Interactive Component-Based Evaluation of Search Engines	15
<i>Giacomo Rocco and Gianmaria Silvello</i>	
3D Average Common Submatrix Measure	26
<i>Federica Franco, Alessia Amelio, and Sergio Greco</i>	

Big Data and Data Science in DL

Lost in Translation: Can We Talk About Big Data Fairly?	35
<i>Matilde Fontanin and Paola Castellucci</i>	
An Ontology and Knowledge Graph Infrastructure for Digital Library Knowledge Representation	47
<i>Stefano Ferilli and Domenico Redavid</i>	
Text-to-Image Synthesis Based on Machine Generated Captions	62
<i>Marco Menardi, Alex Falcon, Saida S. Mohamed, Lorenzo Seidenari, Giuseppe Serra, Alberto Del Bimbo, and Carlo Tasso</i>	
A Streamlined Pipeline to Enable the Semantic Exploration of a Bookstore	75
<i>Miguel Ceriani, Eleonora Bernasconi, and Massimo Mecella</i>	
Re-implementing and Extending Relation Network for R-CBIR	82
<i>Nicola Messina, Giuseppe Amato, and Fabrizio Falchi</i>	
Actual Researcher Contribution (ARC) Versus the Perceived Contribution to the Scientific Body of Knowledge.	93
<i>Mohanad Halaweh</i>	

Cultural Heritage

Towards a Tool for Visual Link Retrieval and Knowledge
Discovery in Painting Datasets 105
Giovanna Castellano and Gennaro Vessio

Identifying, Classifying and Searching Graphic Symbols
in the NOTAE System. 111
*Maria Boccuzzi, Tiziana Catarci, Luca Deodati, Andrea Fantoli,
Antonella Ghignoli, Francesco Leotta, Massimo Mecella, Anna Monte,
and Nina Sietis*

TindArt, an Experiment on User Profiling for Museum Applications 123
Daniel Zilio, Nicola Orio, and Camilla Toniolo

Recognition of Concordances for Indexing in Digital Libraries 135
*Simone Marinai, Samuele Capobianco, Zahra Ziran, Andrea Giuntini,
and Pierluigi Mansueto*

Open Science

RepOSGate: Open Science Gateways for Institutional Repositories 151
Michele Artini, Leonardo Candela, Paolo Manghi, and Silvia Giannini

Training Data Stewards in Italy: Reflection on the FAIR
RDM Summer School 163
Anna Maria Tammaro and Stefano Caselli

Creating Digital Cultural Heritage with Open Data:
From FAIR to FAIR⁵ Principles 173
Nicola Barbuti

Nanocitation: Complete and Interoperable Citations of Nanopublications 182
Erika Fabris, Tobias Kuhn, and Gianmaria Silvello

Correction to: Identifying, Classifying and Searching Graphic Symbols
in the NOTAE System. C1
*Maria Boccuzzi, Tiziana Catarci, Luca Deodati, Andrea Fantoli,
Antonella Ghignoli, Francesco Leotta, Massimo Mecella,
Anna Monte, and Nina Sietis*

Author Index 189