

Commenced Publication in 1973

Founding and Former Series Editors:

Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board

David Hutchison

Lancaster University, Lancaster, UK

Takeo Kanade

Carnegie Mellon University, Pittsburgh, PA, USA

Josef Kittler

University of Surrey, Guildford, UK

Jon M. Kleinberg

Cornell University, Ithaca, NY, USA

Friedemann Mattern

ETH Zurich, Zürich, Switzerland

John C. Mitchell

Stanford University, Stanford, CA, USA

Moni Naor

Weizmann Institute of Science, Rehovot, Israel

C. Pandu Rangan

Indian Institute of Technology, Madras, India

Bernhard Steffen

TU Dortmund University, Dortmund, Germany

Demetri Terzopoulos

University of California, Los Angeles, CA, USA

Doug Tygar

University of California, Berkeley, CA, USA

Gerhard Weikum

Max Planck Institute for Informatics, Saarbrücken, Germany

More information about this series at <http://www.springer.com/series/7409>

Norbert Fuhr · László Kovács
Thomas Risse · Wolfgang Nejdl (Eds.)

Research and Advanced Technology for Digital Libraries

20th International Conference on Theory and Practice
of Digital Libraries, TPD L 2016
Hannover, Germany, September 5–9, 2016
Proceedings

Editors

Norbert Fuhr
Universität Duisburg-Essen
Duisburg
Germany

László Kovács
Hungarian Academy of Sciences
Budapest
Hungary

Thomas Risse
Leibniz Universität Hannover
Hannover
Germany

Wolfgang Nejdl
Leibniz Universität Hannover
Hannover
Germany

ISSN 0302-9743 ISSN 1611-3349 (electronic)
Lecture Notes in Computer Science
ISBN 978-3-319-43996-9 ISBN 978-3-319-43997-6 (eBook)
DOI 10.1007/978-3-319-43997-6

Library of Congress Control Number: 2016947200

LNCS Sublibrary: SL3 – Information Systems and Applications, incl. Internet/Web, and HCI

© Springer International Publishing Switzerland 2016

Chapter 27 is distributed 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, 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

These proceedings contain the reviewed papers presented at the 20th International Conference on Theory and Practice of Digital Libraries (TPDL), which was held in Hannover, Germany, during September 5–9, 2016. The L3S Research Center and the German National Library of Science and Technology (TIB) organized the conference.

The TPDL conference constitutes a leading scientific forum on digital libraries that brings together researchers, developers, content providers, and users in the field of digital libraries.

The 20th conference in the series of ECDL/TPDL conferences is a milestone in the professional activity of the digital library community in Europe. Digital library research and technology are becoming mainstream. Information creation, acquisition, access, processing, distribution, evaluation, and preservation are no longer possible without the methods and technologies developed in our field.

Our community in Europe is rapidly changing. Great “pioneers” of the domain are now retiring, and newcomers are creating productive research atmosphere in which our community members meet with representatives of other communities. For the first time, the conference included two additional tracks on “Digital Humanities” and “e-Infrastructures,” opening our door to these neighboring communities.

This year’s lead topic at the conference was “Overcoming the Limits of Digital Archives” with a clear intention to absorb the problems of other domains in which techniques and technologies of digital archiving could be truly challenged.

From the 93 papers submitted to the conference, 46 long papers and 13 short ones belonged to the main conference track; also, nine papers were submitted for the Digital Humanities track, and five papers to the e-Infrastructures track. The Program Committee accepted 28 full papers and eight short papers. Furthermore, five posters were accepted.

In addition to these papers and posters, we also invited three keynote speakers: Jan Rybicki (Jagiellonian University, Kraków, Poland), Tony Veale (University College Dublin, Ireland), and David Bainbridge (University of Waikato, and Director of the New Zealand Digital Library).

The conference program also included a Doctoral Consortium, and three tutorials on “Introduction to Fedora 4,” “Building Digital Library Collections with Greenstone 3,” and “Text Mining Workflows for Indexing Archives with Automatically Extracted Semantic Metadata.” Finally, three workshops were organized in conjunction with the main conference: “Videos in Digital Libraries: What’s in it for Libraries, Publishers, and Scientists?” “NKOS – 15th European Networked Knowledge Organization Systems,” and the “First International Workshop on Reproducible Open Science (RepScience2016).”

The success of TPDL 2016 is a result of the collegial teamwork of many individuals, who worked tirelessly to make the conference a top research forum. We acknowledge and thank the Program Committee members and the Program Committee Chairs for

their extraordinary efforts in assembling this outstanding program: Gerhard Lauer and Milena Dobрева (Digital Humanities Track Chairs), Tobias Blanke and Laurent Romary (eInfrastructure Track Chairs), Stefan Rüger and Hannu Toivonen (Creativity and Multimedia Chairs), Heiko Schuldt and Peter Löwe (Workshops Chairs), Wolf-Tilo Balke (Panels Chair), Fabio Crestani and Adam Jatowt (Poster and Demonstration Chairs), Ingo Frommholz (Tutorial Chair), and Kjetil Nørvåg (Doctoral Consortium Chair). We would especially like to thank Uwe Rosemann, who served as the honorary chair.

We also express our deep appreciation of the outstanding work put in over many months by Nattiya Kanhabua (Publicity Chair), Gerhard Gossen (Publication Chair), and the local organization team comprising Marion Wicht, Nicole Petri, Katrin Hanebutt, and Miroslav Sheltev. Without their tireless efforts, this conference would not have been a success. We are also thankful to the many student volunteers from Leibniz University Hannover and L3S Research Center.

In addition, there are many other individuals whose contributions we warmly acknowledge. We benefited greatly from the sage advice provided by José Borbinha (TPDL Steering Committee Chair) and the fruitful discussions with the TPDL Steering Committee members.

We warmly acknowledge the financial support of our corporate sponsors: Elsevier B.V. (at the gold level) and Ex Libris (at the bronze level).

Finally, we thank all the authors, presenters, and participants of the conference. We hope that you enjoy the conference proceedings.

September 2016

Norbert Fuhr
László Kovács
Thomas Risse
Wolfgang Nejdl

Organization

Honorary Chair

Uwe Rosemann

TIB Hannover, Germany

General Co-chairs

Wolfgang Nejdl

Leibniz Universität Hannover and L3S Research
Center, Germany

Thomas Risse

Leibniz Universität Hannover and L3S Research
Center, Germany

Program Co-chairs

Norbert Fuhr

University of Duisburg-Essen, Germany

László Kovács

MTA SZTAKI, Hungary

TPDL-DH Track Co-chairs

Gerhard Lauer

University of Göttingen, Germany

Milena Dobрева

University of Malta, Malta

TPDL-eInfra Track Co-chairs

Tobias Blanke

King's College London, UK

Laurent Romary

Inria, France

TPDL-Creativity and Multimedia Track Co-chairs

Stefan Rüger

The Open University, UK

Hannu Toivonen

University of Helsinki, Finland

Workshops Chairs

Heiko Schuldt

University of Basel, Switzerland

Peter Löwe

TIB Hannover, Germany

Panels Chair

Wolf-Tilo Balke

TU Braunschweig, Germany

Poster and Demonstration Chairs

Fabio Crestani	University of Lugano, Switzerland
Adam Jatowt	Kyoto University, Japan

Tutorial Chair

Ingo Frommholz	University of Bedfordshire, UK
----------------	--------------------------------

Doctoral Consortium Chair

Kjetil Nørvåg	Norwegian University of Science and Technology, Norway
---------------	---

Publicity Chair

Nattiya Kanhabua	Aalborg University, Denmark
------------------	-----------------------------

Publication Chair

Gerhard Gossen	Leibniz Universität Hannover and L3S Research Center, Germany
----------------	--

Local Organizing Committee

Marion Wicht	L3S Research Center, Germany
Katrin Hanebutt	TIB Hannover, Germany
Nicole Petri	TIB Hannover, Germany
Miroslav Shaltev	L3S Research Center, Germany
Dimitar Mitev	L3S Research Center, Germany

Senior Program Committee

Maristella Agosti	University of Padua, Italy
Wolf-Tilo Balke	TU Braunschweig, Germany
Tobias Blanke	King's College London, UK
Jose Borbinha	INESC-ID, Instituto Superior Técnico, Universidade de Lisboa, Portugal
George Buchanan	City University London, UK
Donatella Castelli	CNR-ISTI, Italy
Stavros Christodoulakis	Technical University of Crete, Greece
Panos Constantopoulos	Athens University of Economics and Business, Greece
Fabio Crestani	Università della Svizzera Italiana (USI), Switzerland
Sally Jo Cunningham	Waikato University, New Zealand
Nicola Ferro	University of Padua, Italy
Richard Furuta	Texas A&M University, USA

C. Lee Giles	Pennsylvania State University, USA
Marcos Goncalves	Federal University of Minas Gerais, Brazil
Matthias Hemmje	FU Hagen, Germany
Andreas Henrich	University of Bamberg, Germany
Adam Jatowt	Kyoto University, Japan
Carl Lagoze	University of Michigan, USA
Ray Larson	University of California, USA
Peter Löwe	TIB Hannover, Germany
Clifford Lynch	CNI, USA
Carlo Meghini	CNR-ISTI, Italy
Wolfgang Nejdl	Leibniz Universität Hannover and L3S Research Center, Germany
Erich Neuhold	University of Vienna, Austria
Christos Papatheodorou	Ionian University, Greece
Andreas Rauber	Vienna University of Technology, Austria
Thomas Risse	Leibniz Universität Hannover, Germany
Laurent Romary	Inria, France
Seamus Ross	University of Toronto, Canada
Pertti Vakkari	University of Tampere, Finland
Herbert van de Sompel	Los Alamos National Laboratory, USA

Program Committee

Trond Aalberg	Norwegian University of Science and Technology, Norway
Robert B. Allen	Yonsei University, South Korea
Werner Bailer	Joanneum Research, Austria
David Bainbridge	University of Waikato, New Zealand
Thomas Baker	DCMI, Germany
Christoph Becker	University of Toronto, Canada
Jöran Beel	University of Konstanz, Germany
Maria Bielikova	Slovak University of Technology in Bratislava, Slovakia
Pável Calado	INESC-ID, Instituto Superior Técnico, Universidade de Lisboa, Portugal
Jose H. Canos	Universitat Politecnica de Valencia, Spain
Vittore Casarosa	ISTI-CNR, Italy
Lillian Cassel	Villanova University, USA
Theodore Dalamagas	Athena Research Center, Greece
Lois Delcambre	Portland State University, USA
Giorgio Maria Di Nunzio	University of Padua, Italy
Boris Dobrov	Research Computing Center of Moscow State University, Russia
Shyamala Doraisamy	University Putra Malaysia, Malaysia
Fabien Duchateau	Universite Claude Bernard Lyon 1 – LIRIS, France
Pinar Duygulu	Hacettepe University, Turkey

Pierluigi Feliciati	Università degli studi di Macerata, Italy
Niels Ole Finnemann	University of Copenhagen, Denmark
Schubert Foo	Nanyang Technological University, Singapore
Edward Fox	Virginia Polytechnic Institute and State University, USA
Nuno Freire	The European Library, Portugal
Ingo Frommholz	University of Bedfordshire, UK
Manolis Gergatsoulis	Ionian University, Greece
Pablo Gervas	Universidad Complutense de Madrid, Spain
Jane Greenberg	College of Computing and Informatics, USA
Cathal Gurrin	Dublin City University, Ireland
Martin Halvey	University of Strathclyde, UK
Allan Hanbury	TU Wien, Austria
Preben Hansen	Stockholm University, Sweden
Mark Hedges	King's College London, UK
Annika Hinze	University of Waikato, New Zealand
Frank Hopfgartner	University of Glasgow, UK
Jieh Hsiang	National Taiwan University, Taiwan, R.O.C
Jane Hunter	University of Queensland, Australia
Antoine Isaac	Europeana and VU University Amsterdam, The Netherlands
Jaap Kamps	University of Amsterdam, The Netherlands
Sarantos Kapidakis	Ionian University, Greece
Michael Khoo	Drexel University, USA
Claus-Peter Klas	GESIS – Leibniz Institute for Social Sciences, Germany
Martin Klein	University of California Los Angeles, USA
Stefanos Kollias	NTUA, Greece
Alexandros Koulouris	TEI of Athens, Greece
Mounia Lalmas	Yahoo Labs, UK
Ronald Larsen	University of Pittsburgh, USA
Séamus Lawless	Trinity College Dublin, Ireland
Hyowon Lee	Singapore University of Technology and Design, Singapore
Suzanne Little	Dublin City University, Ireland
Haiming Liu	University of Bedfordshire, UK
Penousal Machado	University of Coimbra, Portugal
Joao Magalhaes	Universidade Nova de Lisboa, Portugal
Richard Marciano	University of Maryland, USA
Bruno Martins	INESC-ID, Instituto Superior Técnico, Universidade de Lisboa, Portugal
Philipp Mayr	GESIS – Leibniz Institute for Social Sciences, Germany
Cezary Mazurek	IchB PAN – PCSS, Poland
Dana Mckay	Swinburne University of Technology Library, Australia
Andras Micsik	MTA SZTAKI, Hungary

Agnieszka Mykowiecka	Institute of Computer Science, Polish Academy of Sciences, Poland
Michael Nelson	Old Dominion University, USA
Heike Neuroth	Potsdam University of Applied Sciences, Germany
David Nichols	University of Waikato, New Zealand
Ragnar Nordlie	Oslo and Akershus University College, Norway
Moirra C. Norrie	ETH Zurich, Switzerland
Andreas Nürnberger	Otto-von-Guericke-Universität Magdeburg, Germany
Pasquale Pagano	CNR-ISTI, Italy
Raul Palma	Poznan Supercomputing and Networking Center (PSNC), Poland
Ioannis Papadakis	Ionian University, Greece
Rafael Perez y Perez	UAM Cujimalpa, Mexico
Maggy Pezeril	University of Montpellier 2, France
Dimitris Plexousakis	Foundation for Research and Technology Hellas (FORTH), Greece
Edie Rasmussen	University of British Columbia, Canada
Raivo Ruusalepp	National Library of Estonia, Estonia
Hany SalahEldeen	Microsoft, USA
Philipp Schaer	GESIS – Leibniz Institute for Social Sciences, Germany
Michalis Sfakakis	Ionian University, Greece
Gianmaria Silvello	University of Padua, Italy
Alan Smeaton	Dublin City University, Ireland
Shigeo Sugimoto	University of Tsukuba, Japan
Hussein Suleman	University of Cape Town, South Africa
Tamara Sumner	University of Colorado Boulder, USA
Atsuhiko Takasu	National Institute of Informatics, Japan
Giannis Tsakonas	University of Patras, Greece
Theodora Tsikrika	CERTH, Greece
Chrisa Tsinaraki	European Union – Joint Research Center (EU-JRC), Italy
Yannis Tzitzikas	University of Crete and FORTH-ICS, Greece
Dan Ventura	Brigham Young University, USA
Felisa Verdejo	Universidad Nacional de Educacion a Distancia, Spain
Stefanos Vrochidis	ITI-CERTH, Greece
Michele C. Weigle	Old Dominion University, USA
Marcin Werla	Poznan Supercomputing and Networking Center (PSNC), Poland
Geraint Wiggins	Queen Mary University of London, UK
Iris Xie	University of Wisconsin-Milwaukee, USA
Maja Zumer	University of Ljubljana, Slovenia

Keynotes

Pretty Things Done with (Electronic) Texts: Why We Need Full-Text Access

Jan Rybicki

Jagiellonian University, Kraków, Poland

Abstract. Stylometry, aka computational stylistics, is a field that has produced compelling visualizations of patterns of similarity and difference between texts, based on various quantitative, i.e. countable features. Even if these countable features are often very basic elements that have not been traditionally associated with “style” or “meaning” or “message” – more often than not, stylometrists work with frequencies of function words or of part-of-speech n-grams – the various statistical measures applied to them often yield results that “make sense” in terms of authorial attribution, chronology, genre, or gender – or simply from the point of view of traditional literary studies. It has now become quite simple to produce a “map,” or in fact a network analysis, of 1000 novels in English that shows a very clear progression from early (green) to modern (purple) writing:



Whether or not this is a simple effect of linguistic change (there are reasons to think it is NOT), the feasibility of such approaches – apart from a plethora of methodological problems – relies on stylometrists’ access to full texts. This is still a particularly unpleasant stumbling block: even when dealing with public-domain material, textual collections are dispersed or incompatible or unreliable or fragmentary (pick one or any of these), and stylometrists continue to struggle, often steering on the margins of reliability and of (copyright) laws. Perhaps the main reason is that they do not complain enough to the right people; digital librarians might be a good group to start with.

Jan Rybicki is Assistant Professor at the Institute of English Studies, Jagiellonian University, Kraków, Poland; he also taught at Rice University, Houston, TX and Kraków’s Pedagogical University. His interests include translation, comparative literature and humanities computing (especially stylometry and authorship attribution).

He has worked extensively (both traditionally and digitally) on Henryk Sienkiewicz and the reception of the Polish novelist's works into English, and on the reception of English literature in Poland. Rybicki is also an active literary translator, with more than twenty translated novels by authors such as Coupland, Fitzgerald, Golding, Gordimer, le Carré or Winterson.

Metaphors All the Way Down: The Many Practical Uses of Figurative Language Understanding

Tony Veale

University College Dublin, Dublin, Ireland

Abstract. More and more of the content we consume on social networking platforms such as Twitter is computer-generated. If this content were just another form of spam then these networks would truly become the 21st century's version of Borges' Library of Babel, a world in which the content that is actually worth consuming is ultimately lost in a sea of random, meaningless noise. Yet there are encouraging signs that it need not be so. On Twitter, savvy users follow machine generators of content – called Twitterbots – knowing them to be machines and valuing their outputs all the more for their oddity and artificiality. These bots are not designed to fool humans, but to engage them in language games that explore the relation of form to meaning and provocatively flit along the boundary between sense and nonsense. As there is a long history of human artists doing precisely this – ranging from Duchamp and Breton to Dali and Burroughs – bots are simply the next stage in the evolution of thought-provoking automated content creation. Bots that lack a semantic grounding often compensate with a statistical brio for superficial production that naturally yields outputs which can appear poetic and deeply metaphorical. In this presentation I shall consider how an AI bot can deliberately produce meaningful metaphors from a knowledge-base of everyday facts and beliefs. But we need not stop there: given an ability to produce metaphors on demand, a computational system can produce higher-level digital constructs such as stories and games that make the familiar seem strange and the strange seem oddly and meaningfully familiar. I will describe how metaphor and blending can play a pivotal role in these engaging new forms of digital content, to the extent that they truly do rest on metaphors all the way down.

Dr. Tony Veale is a Computer Scientist at University College Dublin, Ireland, where his research focuses on the computational modeling of creative linguistic phenomena, including metaphor, blending, simile, analogy and verbal irony. He leads the European Commission's PROSECCO network (PROSECCO-network.eu and @PROSECCO network), an international coordination action that aims to promote the scientific exploration of Computational Creativity. He is particularly interested in the generative creativity of metaphor, and builds generative models of metaphor, simile and blending which are made publically available as reusable Web services to promote the integration of figurative language processing capabilities in third-party applications. He is the author of the 2012 monograph on computational linguistic creativity titled

Exploding the Creativity Myth: The Computational Foundations of Linguistic Creativity from Bloomsbury press and co-author (with Ekaterina Shutova and Beata Beigman Klebanov) of *Metaphor: A Computational Perspective* from Morgan Claypool press. He is the creator of the metaphor-generating and story-telling Twitterbot @MetaphorMagnet and the founder of the educational Web-site RobotComix.com, which promotes the philosophy and practice of Computational Creativity to the general public.

Mozart's Laptop: Implications for Creativity in Multimedia Digital Libraries and Beyond

David Bainbridge

University of Waikato, and Director of the New Zealand Digital Library
Research Project, Hamilton, New Zealand

Abstract. If Mozart were alive today, what sorts of musical apps would such an innovative composer use on his laptop? In this keynote I will attempt to answer—at least in part—this question. We will metaphorically drop in on Wolfgang composing at home in the morning, at an orchestra rehearsal in the afternoon, and find him unwinding in the evening playing a spot of the new game Piano Hero which is (in my fictional narrative) all the rage in the Viennese coffee shops! From a pedagogical perspective, these three scenarios are chosen because they cover the main forms of digital music representation: audio, sheet music, and symbolic notation. In each case I will demonstrate software prototypes that combines digital music library and music information retrieval research to provide novel forms of access and management of musical digital content. I will then broaden the discussion and relate the work to other forms of media, and (going beyond this) contemplate whether the presented research fits the established definition of a digital library, or if it is perhaps time to repurpose traditional ideas about the structure and capabilities of digital libraries, or even revisit what we define as a digital library.

David Bainbridge is a Professor of Computer Science at the University of Waikato, and Director of the New Zealand Digital Library Research Project. He is an advocate of open source software, and an active coder on the Greenstone digital library project, and the spatial hypermedia system, Expeditee. His research interests include multimedia content analysis, and human computer interaction in addition to digital libraries. He has published widely in these areas, including the book *How to Build a Digital Library*, with colleagues Ian Witten and Dave Nichols, now into its second edition. David graduated with a Bachelor of Engineering in Computer Science from Edinburgh University, UK as the class medalist, and undertook his PhD in Computer Science at Canterbury University, New Zealand as a Commonwealth Scholar.

Contents

Digital Library Design

Realizing Inclusive Digital Library Environments: Opportunities and Challenges	3
<i>Wondwossen M. Beyene</i>	
A Maturity Model for Information Governance	15
<i>Diogo Proença, Ricardo Vieira, and José Borbinha</i>	
Usage-Driven Dublin Core Descriptor Selection: A Case Study Using the Dendro Platform for Research Dataset Description	27
<i>João Rocha da Silva, Cristina Ribeiro, and João Correia Lopes</i>	

Search

Retrieving and Ranking Similar Questions from Question-Answer Archives Using Topic Modelling and Topic Distribution Regression	41
<i>Pedro Chahuara, Thomas Lampert, and Pierre Gançarski</i>	
Survey on High-level Search Activities Based on the Stratagem Level in Digital Libraries	54
<i>Zeljko Carevic and Philipp Mayr</i>	
Content-Based Video Retrieval in Historical Collections of the German Broadcasting Archive	67
<i>Markus Mühling, Manja Meister, Nikolaus Korfhage, Jörg Wehling, Angelika Hörth, Ralph Ewerth, and Bernd Freisleben</i>	

User Aspects

Profile-Based Selection of Expert Groups	81
<i>Georgios A. Sfyris, Nikolaos Fragkos, and Christos Doulkeridis</i>	
Tracking and Re-finding Printed Material Using a Personal Digital Library. . . .	94
<i>Annika Hinze and Amay Dighe</i>	
ArchiveWeb: Collaboratively Extending and Exploring Web Archive Collections	107
<i>Zeon Trevor Fernando, Ivana Marenzi, Wolfgang Nejdl, and Rishita Kalyani</i>	

Web Archives

Web Archive Profiling Through Fulltext Search	121
<i>Sawood Alam, Michael L. Nelson, Herbert Van de Sompel, and David S.H. Rosenthal</i>	
Comparing Topic Coverage in Breadth-First and Depth-First Crawls Using Anchor Texts	133
<i>Thaer Samar, Myriam C. Traub, Jacco van Ossenbruggen, and Arjen P. de Vries</i>	
How to Search the Internet Archive Without Indexing It	147
<i>Nattiya Kanhabua, Philipp Kemkes, Wolfgang Nejdl, Tu Ngoc Nguyen, Feliipe Reis, and Nam Khanh Tran</i>	

Semantics

BIB-R: A Benchmark for the Interpretation of Bibliographic Records	163
<i>Joffrey Decourselle, Fabien Duchateau, Trond Aalberg, Naimdjon Takhirov, and Nicolas Lumineau</i>	
Querying the Web of Data with SPARQL-LD	175
<i>Pavlos Fafalios, Thanos Yannakis, and Yannis Tzitzikas</i>	
A Scalable Approach to Incrementally Building Knowledge Graphs	188
<i>Gleb Gawriljuk, Andreas Harth, Craig A. Knoblock, and Pedro Szekely</i>	

Multimedia and Time Aspects

Sub-document Timestamping: A Study on the Content Creation Dynamics of Web Documents	203
<i>Yue Zhao and Claudia Hauff</i>	
Archiving Software Surrogates on the Web for Future Reference	215
<i>Helge Holzmann, Wolfram Sperber, and Mila Runnwerth</i>	
From Water Music to ‘Underwater Music’: Multimedia Soundtrack Retrieval with Social Mass Media Resources	227
<i>Cynthia C.S. Liem</i>	

Digital Library Evaluation

The “Nomenclature of Multidimensionality” in the Digital Libraries Evaluation Domain	241
<i>Leonidas Papachristopoulos, Giannis Tsakonas, Michalis Sfakakis, Nikos Kleidis, and Christos Papatheodorou</i>	

Dissecting a Scholar Popularity Ranking into Different Knowledge Areas . . .	253
<i>Gabriel Pacheco, Pablo Figueira, Jussara M. Almeida, and Marcos A. Gonçalves</i>	

Exploring Comparative Evaluation of Semantic Enrichment Tools for Cultural Heritage Metadata	266
<i>Hugo Manguinhas, Nuno Freire, Antoine Isaac, Juliane Stiller, Valentine Charles, Aitor Soroa, Rainer Simon, and Vladimir Alexiev</i>	

Digital Humanities

Data Integration for the Arts and Humanities: A Language Theoretical Concept.	281
<i>Tobias Gradl and Andreas Henrich</i>	

The Challenge of Creating Geo-Location Markup for Digital Books	294
<i>Annika Hinze, David Bainbridge, and Sally Jo Cunningham</i>	

Exploring Large Digital Libraries by Multimodal Criteria.	307
<i>David Zellhöfer</i>	

Person-Centric Mining of Historical Newspaper Collections	320
<i>Mariona Coll Ardanuy, Jürgen Knauth, Andrei Beliankou, Maarten van den Bos, and Caroline Sporleder</i>	

e-Infrastructures

Usability in Digital Humanities - Evaluating User Interfaces, Infrastructural Components and the Use of Mobile Devices During Research Process.	335
<i>Natasa Bulatovic, Timo Gnadt, Matteo Romanello, Juliane Stiller, and Klaus Thoden</i>	

CERN Analysis Preservation: A Novel Digital Library Service to Enable Reusable and Reproducible Research.	347
<i>Xiaoli Chen, Sünje Dallmeier-Tiessen, Anxhela Dani, Robin Dasler, Javier Delgado Fernández, Pamfilos Fokianos, Patricia Herterich, and Tibor Šimko</i>	

DataQ: A Data Flow Quality Monitoring System for Aggregative Data Infrastructures	357
<i>Andrea Mannocci and Paolo Manghi</i>	

Short Papers

Scientific Social Publications for Digital Libraries	373
<i>Fidan Limani, Atif Latif, and Klaus Tochtermann</i>	

Formal Representation of Socio-Legal Roles and Functions for the Description of History.	379
<i>Yoonmi Chu and Robert B. Allen</i>	
Evaluating Co-authorship Networks in Author Name Disambiguation for Common Names	386
<i>Fakhri Momeni and Philipp Mayr</i>	
Ten Months of Digital Reading: An Exploratory Log Study	392
<i>Pavel Braslavski, Vivien Petras, Valery Likhoshesterov, and Maria Gäde</i>	
A Case Study of Summarizing and Normalizing the Properties of DBpedia Building Instances.	398
<i>Michail Agathos, Eleftherios Kalogeros, and Sarantos Kapidakis</i>	
What Happens When the Untrained Search for Training Information	405
<i>Jorgina Paihama and Hussein Suleman</i>	
InterPlanetary Wayback: Peer-To-Peer Permanence of Web Archives.	411
<i>Mat Kelly, Sawood Alam, Michael L. Nelson, and Michele C. Weigle</i>	
Exploring Metadata Providers Reliability and Update Behavior.	417
<i>Sarantos Kapidakis</i>	

Posters and Demos

TIB AV-Portal: Integrating Automatically Generated Video Annotations into the Web of Data.	429
<i>Jörg Waiteonis, Margret Plank, and Harald Sack</i>	
Supporting Web Surfers in Finding Related Material in Digital Library Repositories	434
<i>Jörg Schlötterer, Christin Seifert, and Michael Granitzer</i>	
Do Ambiguous Words Improve Probing for Federated Search?	438
<i>Günter Urak, Hermann Ziak, and Roman Kern</i>	
Automatic Recognition and Disambiguation of Library of Congress Subject Headings	442
<i>Rosa Tsegaye Aga, Christian Wartena, and Michael Franke-Maier</i>	
The Problem of Categorizing Conferences in Computer Science	447
<i>Suhendry Effendy and Roland H.C. Yap</i>	
Germania Sacra Online – The Research Portal of Clerics and Religious Institutions Before 1810.	451
<i>Bärbel Kröger and Christian Popp</i>	

Open Digital Forms.	454
<i>Hiep Le, Thomas Rebele, and Fabian Suchanek</i>	
A Text Mining Framework for Accelerating the Semantic Curation of Literature	459
<i>Riza Batista-Navarro, Jennifer Hammock, William Ulate, and Sophia Ananiadou</i>	
Erratum to: CERN Analysis Preservation: A Novel Digital Library Service to Enable Reusable and Reproducible Research.	E1
<i>Xiaoli Chen, Sünje Dallmeier-Tiessen, Anxhela Dani, Robin Dasler, Javier Delgado Fernández, Pamfilos Fokianos, Patricia Herterich, and Tibor Šimko</i>	
Tutorials	
Introduction to Fedora 4	465
<i>David Wilcox</i>	
Building Digital Library Collections with Greenstone 3 Tutorial	468
<i>David Bainbridge</i>	
Text Mining Workflows for Indexing Archives with Automatically Extracted Semantic Metadata	471
<i>Riza Batista-Navarro, Axel J. Soto, William Ulate, and Sophia Ananiadou</i>	
Author Index	475