

Commenced Publication in 1973

Founding and Former Series Editors:

Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board

David Hutchison

Lancaster University, 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, Switzerland

John C. Mitchell

Stanford University, CA, USA

Moni Naor

Weizmann Institute of Science, Rehovot, Israel

Oscar Nierstrasz

University of Bern, Switzerland

C. Pandu Rangan

Indian Institute of Technology, Madras, India

Bernhard Steffen

University of Dortmund, Germany

Madhu Sudan

Massachusetts Institute of Technology, MA, USA

Demetri Terzopoulos

University of California, Los Angeles, CA, USA

Doug Tygar

University of California, Berkeley, CA, USA

Moshe Y. Vardi

Rice University, Houston, TX, USA

Gerhard Weikum

Max-Planck Institute of Computer Science, Saarbruecken, Germany

László Kovács Norbert Fuhr
Carlo Meghini (Eds.)

Research and Advanced Technology for Digital Libraries

11th European Conference, ECDL 2007
Budapest, Hungary, September 16-21, 2007
Proceedings

 Springer

Volume Editors

László Kovács
Hungarian Academy of Sciences
Computer and Automation Research Institute
Department of Distributed Systems
H-1111 Budapest, XI. Lágymányosi u. 11., Hungary
E-mail: laszlo.kovacs@sztaki.hu

Norbert Fuhr
University of Duisburg-Essen
Information Systems Department of Computational and Cognitive Sciences
47048 Duisburg, Germany
E-mail: norbert.fuhr@uni-due.de

Carlo Meghini
Consiglio Nazionale delle Ricerche
Istituto di Scienza e Tecnologie della Informazione
56124 Pisa, Italy
E-mail: meghini@isti.cnr.it

Library of Congress Control Number: 2007934833

CR Subject Classification (1998): H.3.7, H.2, H.3, H.4.3, H.5, J.7, J.1, I.7

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

ISSN 0302-9743
ISBN-10 3-540-74850-4 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-74850-2 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: 12121076 06/3180 5 4 3 2 1 0

Preface

We are proud to present the proceedings of the 11th European Conference on Research and Advanced Technology for Digital Libraries (ECDL 2007) which, following Pisa (1997), Heraklion (1998), Paris (1999), Lisbon (2000), Darmstadt (2001), Rome (2002), Trondheim (2003), Bath (2004), Vienna (2005) and Alicante (2006), took place on September 16-21, 2007 in Budapest, Hungary. Over the last 11 years, ECDL has created a strong interdisciplinary community of researchers and practitioners in the field of digital libraries, and has formed a substantial body of scholarly publications contained in the conference proceedings.

ECDL 2007 featured separate calls for paper and poster submissions, resulting in 119 full papers and 34 posters being submitted to the conference. All papers were subject to an in-depth peer-review process; three reviews per submission were produced by a Program Committee of 69 members from 27 countries. In total 36 of 119 full paper submissions were accepted at the Program Committee meeting for presentation at the conference and publication in the proceedings with Springer, resulting in an acceptance rate of 30%. Also, 24 poster/demo submissions and another 15 papers from the full paper submissions were accepted for poster presentation and publication in the proceedings volume.

ECDL 2007 was devoted to discussions about hot issues and applications and primarily provided a forum to reinforce the collaboration of researchers and practitioners. The main conference consisted of 12 technical sessions and a poster /demo session on the following topics: Ontologies, Digital Libraries and the Web, Models, Multimedia and Multilingual Digital Libraries, Grid and Peer-to-Peer, Preservation, User Interfaces, Document Linking, Information Retrieval, Personal Information Management, New DL Applications and User Studies.

The conference featured two panels, which addressed timely and important topics, namely, experiences of DL projects in synergy with the European Commission's initiatives in the panel "On the Move Towards the European Digital Library: BRICKS, TEL, MICHAEL and DELOS converging experiences" chaired by Massimo Bertoncini and the special challenges of Digital Library research and development in the host region of the conference in the panel "Digital Libraries in Central and Eastern Europe: Infrastructure Challenges for the New Europe" chaired by Christine Borgman.

The keynote talk by Seamus Ross (Humanities Computing and Information Management, University of Glasgow) addressed the questions of digital preservation, while the keynote talk by Arne Solvberg (Dept. of Computer and Information Science, Norwegian University of Science and Technology) focused on the challenges of WiFi-Trondheim – an experiment in providing Broadband Everywhere for All.

The preceding tutorials provided further in-depth looks at areas of current interest, including “Thesauri and Ontologies in DLs by Dagobert Soegrel, Introduction to DLs” by Ed Fox, “Approaches for Large Scale Digital Library Infrastructures” by Thomas Risse, and “Building DLs On-Demand by Sharing Content, Services and Computing Resources” by Donatella Castelli.

The workshops, held in conjunction with ECDL2007, covered wide areas of interest: CLEF 2007 – Cross-Language Evaluation Forum, Workshop on “Foundations of Digital Libraries”; LADL 2007 – Cross-Media and Personalized Learning Applications on Top of Digital Libraries, Curriculum Development in Digital Libraries: An Exercise in Developing Lesson Plans , Towards a European Repository Ecology: Conceptualizing Interactions Between Networks of Repositories and Services, NKOS -Networked Knowledge Organization Systems and Services, and Libraries in the Digital Age: What If...?

We would like to take the opportunity to thank everybody who made this conference possible, all the conference participants and presenters, who provided an exciting full-week program of high technical quality. We greatly appreciate the contribution of the Program Committee members, who did an outstanding reviewing job under tight time constraints; and we are grateful to all Chairs and members of the Organization Committee, who worked hard to make the best out of the Conference.

Finally, we would also like to thank the conference sponsors and cooperating agencies: the Computer and Automation Research Institute of the Hungarian Academy of Sciences (MTA SZTAKI), the DELOS Network of Excellence on Digital Libraries, and the Hungarian Tourism Office.

September 2007

László Kovács
Norbert Fuhr
Carlo Meghini

Organization

Organization Committee

General Chair

László Kovács

Department of Distributed Systems, Computer and Automation Research Institute, Hungarian Academy of Sciences, Hungary

Program Co-chairs

Norbert Fuhr

Information Systems Faculty of Engineering Sciences, University of Duisburg-Essen, Germany

Carlo Meghini

Consiglio Nazionale delle Ricerche Istituto di Scienza e Tecnologie dell'Informazione, Italy

Workshops Chairs

Maristella Agosti

University of Padua, Italy

Birte Christensen-Dalsgaard

State and University Library, Denmark

Poster and Demo Chairs

Ulrike Steffens

OFFIS, Germany

José Borbinha

DEI/IST/UTL and INESC-ID, Portugal

Tutorials Chair

Rudi Schmiede

Darmstadt University of Technology, Germany

Publicity and Exhibit Chairs

Yuzuru Tanaka for Asia

Meme Media Laboratory, Hokkaido University, Japan

Jane Hunter for Australia

School of ITEE, Australia

Hussein Suleman for Africa

University of Cape Town, South Africa

Panel Chairs

Seamus Ross

University of Glasgow, UK

Edward Fox

Virginia Tech / Dept. of Computer Science, USA

Doctoral Consortium Chairs

Tiziana Catarci

University of Rome 1, Italy

Nicolas Spyratos

Université de Paris-Sud, France

Local Arrangements Chair

Gusztáv Hencsey

Computer and Automation Research Institute,
Hungarian Academy of Sciences, Hungary

Program Committee

Hanne Albrechtsen

Institute of Knowledge Sharing, Denmark

Margherita Antona

FORTH, Greece

Tom Baker

State and University Library, Germany

Nicholas Belkin

Rutgers University, USA

Maria Bieliková

Slovak University of Technology in Bratislava,
Slovakia

George Buchanan

University of Wales, Swansea

Gerhard Budin

University of Vienna, Austria

Tiziana Catarci

University of Rome 1, Italy

José H. Canós Cerda

Universidad Politecnica de Valencia, Spain

Hsinchun Chen

University of Arizona, Tucson, USA

Anita S.Coleman

University of Arizona, USA

Gregory Crane

Tufts University, USA

Sally Jo Cunningham

University of Waikato, New Zealand

Mário J. Gaspar da Silva

Universidade de Lisboa, Portugal

Pablo de la Fuente

University of Valladolid, Spain

Susanne Dobratz

Humboldt University, Germany

Boris V.Dobrov

Moscow State University, Russia

Jacques Ducloy

CNRS-INIST, France

Lim Ee-Peng

Nanyang Technological University, Singapore

Floriana Esposito

University of Bari, Italy

Schubert Foo

Nanyang Technological University, Singapore

Edward Fox

Virginia Tech, USA

Richard Furtuta

Texas A & M University, USA

Stefan Gradmann

University of Hamburg, Computing Center,
Germany

Allan Hanbury

Vienna University of Technology, Austria

Donna Harman

NIST, USA

Djoerd Hiemstra

Twente University, The Netherlands

Jen-Shin Hong

Department of Computer Science, National
ChiNan University, Taiwan

Leonid Kalinichenko

Russian Academy of Sciences, Russia

Sarantos Kapidakis

Ionian University, Greece

Claus-Peter Klas

University of Duisburg, Germany

Traugott Koch

Max Planck Digital Library, Germany

Harald Krottmaier

Graz University of Technology, Austria

Carl Lagoze

Cornell University, USA

Mounia Lalmas

Queen Mary University of London, UK

Ronald Larsen

University of Pittsburgh, USA

Ray Larson	University of California, Berkeley, USA
Clifford Lynch	Coalition for Networked Information, USA
Antonio Polo Márquez	University of Extremadura, Spain
Catherine C. Marshall	Microsoft Corporation, Redmond, WA, USA
András Micsik	MTA SZTAKI, Hungary
Reagan Moore	SDSC, USA
Marc Nanard	University of Montpellier, France
Liddy Nevile	La Trobe University, Australia
Fernando López Ostenero	UNED, Spain
Anselmo Peñañ	UNED, Spain
Dimitris Plexousakis	FORTH, Greece
Andy Powell	Eduserv Foundation, UK
Hansen Preben	SICS, Sweden
Andreas Rauber	University of Technology, Vienna, Austria
Thomas Risse	Fraunhofer IPSI, Germany
Laurent Romary	Laboratoire Loria CNRS, France
Lloyd Rutledge	CWI, The Netherlands
J. Alfredo Sánchez	Universidad de las Americas Puebla, Mexico
Heiko Schuldt	University of Basel, Switzerland
Timos Sellis	National Technical University of Athens, Greece
Dagobert Soergel	University of Maryland, USA
Ingeborg Solvberg	Norwegian University of Technology and Science, Norway
Jela Steinerova	Comenius University in Bratislava, Slovakia
Shigeo Sugimoto	Graduate School of Library, Information and Media Studies, University of Tsukuba, Japan
Tamara Sumner	University of Colorado, Boulder, USA
Jesús Tramullas	University of Zaragoza, Spain
Omar Valdiviezo	Universidad de las Americas Puebla, Mexico
Herbert Van de Sompel	Los Alamos National Laboratory, USA
Ian Witten	University of Waikato, New Zealand

Table of Contents

Ontologies

The Semantic GrowBag Algorithm: Automatically Deriving Categorization Systems	1
<i>Jörg Diederich and Wolf-Tilo Balke</i>	
Ontology-Based Question Answering for Digital Libraries	14
<i>Stephan Bloehdorn, Philipp Cimiano, Alistair Duke, Peter Haase, Jörg Heizmann, Ian Thurlow, and Johanna Völker</i>	
Formalizing the Get-Specific Document Classification Algorithm	26
<i>Fausto Giunchiglia, Ilya Zaihrayeu, and Uladzimir Kharkevich</i>	

Digital Libraries and the Web

Trustworthiness Analysis of Web Search Results	38
<i>Satoshi Nakamura, Shinji Konishi, Adam Jatowt, Hiroaki Ohshima, Hiroyuki Kondo, Taro Tezuka, Satoshi Oyama, and Katsumi Tanaka</i>	
Improved Publication Scores for Online Digital Libraries Via Research Pyramids	50
<i>Sulieaman Bani-Ahmad and Gultekin Ozsoyoglu</i>	
Key Element-Context Model: An Approach to Efficient Web Metadata Maintenance	63
<i>Ba-Quy Vuong, Ee-Peng Lim, Aixin Sun, Chew-Hung Chang, Kalyani Chatterjea, Dion Hoe-Lian Goh, Yin-Leng Theng, and Jun Zhang</i>	

Models

A Cooperative-Relational Approach to Digital Libraries	75
<i>Alessio Malizia, Paolo Bottoni, Stefano Levialdi, and Francisco Astorga-Paliza</i>	
Mind the (Intelligibility) Gap	87
<i>Yannis Tzitzikas and Giorgos Flouris</i>	
Using XML Logical Structure to Retrieve (Multimedia) Objects	100
<i>Zhigang Kong and Mounia Lalmas</i>	

Multimedia and Multilingual DLs

Lyrics-Based Audio Retrieval and Multimodal Navigation in Music Collections 112
Meinard Müller, Frank Kurth, David Damm, Christian Fremerey, and Michael Clausen

Automatic Identification of Music Works Through Audio Matching..... 124
Riccardo Miotto and Nicola Orio

Roadmap for MultiLingual Information Access in the European Library 136
Maristella Agosti, Martin Braschler, Nicola Ferro, Carol Peters, and Sjoerd Siebinga

Grid and Peer-to-Peer

MinervaDL: An Architecture for Information Retrieval and Filtering in Distributed Digital Libraries 148
Christian Zimmer, Christos Tryfonopoulos, and Gerhard Weikum

A Grid-Based Infrastructure for Distributed Retrieval 161
Fabio Simeoni, Leonardo Candela, George Kakaletris, Mads Sibeko, Pasquale Pagano, Giorgos Papanikos, Paul Polydoros, Yannis Ioannidis, Dagfinn Aarvaag, and Fabio Crestani

VIRGIL – Providing Institutional Access to a Repository of Access Grid Sessions 174
Ron Chernich, Jane Hunter, and Alex Davies

Preservation

Opening Schrödingers Library: Semi-automatic QA Reduces Uncertainty in Object Transformation 186
Lars R. Clausen

Texts, Illustrations, and Physical Objects: The Case of Ancient Shipbuilding Treatises 198
Carlos Monroy, Richard Furuta, and Filipe Castro

Trustworthy Digital Long-Term Repositories: The Nestor Approach in the Context of International Developments 210
Susanne Dobratz and Astrid Schoger

User Interfaces

Providing Context-Sensitive Access to the Earth Observation Product Library 223
Stephan Kiemle and Burkhard Freitag

T-Scroll: Visualizing Trends in a Time-Series of Documents for Interactive User Exploration	235
<i>Yoshiharu Ishikawa and Mikine Hasegawa</i>	

Thesaurus-Based Feedback to Support Mixed Search and Browsing Environments	247
<i>Edgar Meij and Maarten de Rijke</i>	

Document Linking

Named Entity Identification and Cyberinfrastructure	259
<i>Alison Babeu, David Bamman, Gregory Crane, Robert Kummer, and Gabriel Weaver</i>	

Finding Related Papers in Literature Digital Libraries	271
<i>Nattakarn Ratprasartporn and Gultekin Ozsoyoglu</i>	

Extending Semantic Matching Towards Digital Library Contexts	285
<i>László Kovács and András Micsik</i>	

Information Retrieval

Towards a Unified Approach Based on Affinity Graph to Various Multi-document Summarizations	297
<i>Xiaojun Wan and Jianqiao Xiao</i>	

Large-Scale Clustering and Complete Facet and Tag Calculation	309
<i>Bolette Ammitzbøll Madsen</i>	

Annotation-Based Document Retrieval with Probabilistic Logics	321
<i>Ingo Frommholz</i>	

Personal Information Management

Evaluation of Visual Aid Suite for Desktop Searching	333
<i>Schubert Foo and Douglas Hendry</i>	

Personal Environment Management	345
<i>Anna Zacchi and Frank Shipman</i>	

Empirical Evaluation of Semi-automated XML Annotation of Text Documents with the GoldenGATE Editor	357
<i>Guido Sautter, Klemens Böhm, Frank Padberg, and Walter Tichy</i>	

New DL Applications

Exploring Digital Libraries with Document Image Retrieval	368
<i>Simone Marinai, Emanuele Marino, and Giovanni Soda</i>	

Know Thy Sensor: Trust, Data Quality, and Data Integrity in Scientific Digital Libraries 380
Jillian C. Wallis, Christine L. Borgman, Matthew S. Mayernik, Alberto Pepe, Nithya Ramanathan, and Mark Hansen

Digital Libraries Without Databases: The Bleek and Lloyd Collection 392
Hussein Suleman

User Studies

A Study of Citations in Users' Online Personal Collections 404
Nishikant Kapoor, John T. Butler, Sean M. McNee, Gary C. Fouty, James A. Stemper, and Joseph A. Konstan

Investigating Document Triage on Paper and Electronic Media 416
George Buchanan and Fernando Loizides

Motivating and Supporting User Interaction with Recommender Systems 428
Andreas W. Neumann

Panels

On the Move Towards the European Digital Library: BRICKS, TEL, MICHAEL and DELOS Converging Experiences 440
Massimo Bertoincini

Digital Libraries in Central and Eastern Europe: Infrastructure Challenges for the New Europe 442
Christine L. Borgman, Tatjana Aparac-Jelušić, Sonja Pigac Ljubi, Zinaida Manžuch, György Sebestyén, and András Gábor

Posters and Demos

Electronic Work: Building Dynamic Services over Logical Structure Using Aqueducts for XML Processing 445
Miguel A. Martínez-Prieto, Pablo de la Fuente, Jesús Vegas, and Joaquín Adiego

A Model of Uncertainty for Near-Duplicates in Document Reference Networks 449
Claudia Hess and Michel de Rougemont

Assessing Quality Dynamics in Unsupervised Metadata Extraction for Digital Libraries 454
Alexander Ivanyukovich, Maurizio Marchese, and Patrick Reuther

Bibliographical Meta Search Engine for the Retrieval of Scientific Articles	458
<i>Artur Gajek, Stefan Klink, Patrick Reuther, Bernd Walter, and Alexander Weber</i>	
In-Browser Digital Library Services	462
<i>Hussein Suleman</i>	
Evaluating Digital Libraries with 5SQual	466
<i>Bárbara L. Moreira, Marcos A. Gonçalves, Alberto H.F. Laender, and Edward A. Fox</i>	
Reducing Costs for Digitising Early Music with Dynamic Adaptation . . .	471
<i>Laurent Pugin, John Ashley Burgoyne, and Ichiro Fujinaga</i>	
Supporting Information Management in Digital Libraries with Map-Based Interfaces	475
<i>Rudolf Mayer, Angela Roiger, and Andreas Rauber</i>	
Policy Decision Tree for Academic Digital Collections	481
<i>Alexandros Koulouris and Sarantos Kapidakis</i>	
Personalized Faceted Browsing for Digital Libraries	485
<i>Michal Tvarožek and Mária Bielíková</i>	
The Use of Metadata in Visual Interfaces to Digital Libraries	489
<i>Ali Shiri</i>	
Location and Format Independent Distributed Annotations for Collaborative Research	495
<i>Fabio Corubolo, Paul B. Watry, and John Harrison</i>	
NSDL MatDL: Adding Context to Bridge Materials e-Research and e-Education	499
<i>Laura Bartolo, Cathy Lowe, Dean Krafft, and Robert Tandy</i>	
A Framework for the Generation of Transformation Templates	501
<i>Manuel Llavador and José H. Canós</i>	
MultiMatch – Multilingual/Multimedia Access to Cultural Heritage	505
<i>Giuseppe Amato, Juan Cigarrán, Julio Gonzalo, Carol Peters, and Pasquale Savino</i>	
The Future of Large-Scale Evaluation Campaigns for Information Retrieval in Europe	509
<i>Maristella Agosti, Giorgio Maria Di Nunzio, Nicola Ferro, Donna Harman, and Carol Peters</i>	
Digital 101: Public Exhibition System of the National Digital Archives Program, Taiwan	513
<i>Ku-Lun Huang and Hsiang-An Wang</i>	

aScience: A Thematic Network on Speech and Tactile Accessibility to Scientific Digital Resources	515
<i>Cristian Bernareggi and Gian Carlo Dalto</i>	
PROBADO – A Generic Repository Integration Framework	518
<i>Harald Krottmaier, Frank Kurth, Thorsten Steenweg, Hans-Jürgen Appelrath, and Dieter Fellner</i>	
VCenter: A Digital Video Broadcast System of NDAP Taiwan	522
<i>Hsiang-An Wang, Chih-Yi Chiu, and Yu-Zheng Wang</i>	
Retrieving Tsunami Digital Library by Use of Mobile Phones	525
<i>Sayaka Imai, Yoshinari Kanamori, and Nobuo Shuto</i>	
Using Watermarks and Offline DRM to Protect Digital Images in DIAS	529
<i>Hsin-Yu Chen, Hsiang-An Wang, and Chin-Lung Lin</i>	
CIDOC CRM in Action – Experiences and Challenges	532
<i>Philipp Nussbaumer and Bernhard Haslhofer</i>	
The Legal Environment of Digital Curation – A Question of Balance for the Digital Librarian	534
<i>Mags McGinley</i>	
Demonstration: Bringing Lives to Light: Browsing and Searching Biographical Information with a Metadata Infrastructure	539
<i>Ray R. Larson</i>	
Repository Junction and Beyond at the EDINA (UK) National Data Centre	543
<i>Robin Rice, Peter Burnhill, Christine Rees, and Anne Robertson</i>	
A Scalable Data Management Tool to Support Epidemiological Modeling of Large Urban Regions	546
<i>Christopher L. Barrett, Keith Bisset, Stephen Eubank, Edward A. Fox, Yi Ma, Madhav Marathe, and Xiaoyu Zhang</i>	
Living Memory Annotation Tool – Image Annotations for Digital Libraries	549
<i>Wolfgang Jochum, Max Kaiser, Karin Schellner, and Franz Wirl</i>	
A User-Centred Approach to Metadata Design	551
<i>Emma Tonkin</i>	
A Historic Documentation Repository for Specialized and Public Access	555
<i>Cristina Ribeiro, Gabriel David, and Catalin Calistru</i>	

Finding It on Google, Finding It on del.icio.us.	559
<i>Jacek Gwizdka and Michael Cole</i>	
DIGMAP – Discovering Our Past World with Digitised Maps.....	563
<i>José Borbinha, Gilberto Pedrosa, Diogo Reis, João Luzio, Bruno Martins, João Gil, and Nuno Freire</i>	
Specification and Generation of Digital Libraries into DSpace Using the 5S Framework	567
<i>Douglas Gorton, Weiguo Fan, and Edward A. Fox</i>	
EOD - European Network of Libraries for eBooks on Demand	570
<i>Zoltán Mező, Sonja Svoljšak, and Silvia Gstrein</i>	
Semantics and Pragmatics of Preference Queries in Digital Libraries	573
<i>El hadji Mamadou Nguer</i>	
Applications for Digital Libraries in Language Learning and the Professional Development of Teachers.....	579
<i>Alannah Fitzgerald</i>	
Author Index	583