Lecture Notes in Computer Science 1923

Edited by G. Goos, J. Hartmanis and J. van Leeuwen

Springer
Berlin
Heidelberg
New York
Barcelona
Hong Kong
London
Milan
Paris
Singapore
Tokyo

José Borbinha Thomas Baker (Eds.)

Research and Advanced Technology for Digital Libraries

4th European Conference, ECDL 2000
Lisbon, Portugal, September 18-20, 2000
Proceedings

Springer

Series Editors

Gerhard Goos, Karlsruhe University, Germany
Juris Hartmanis, Cornell University, NY, USA
Jan van Leeuwen, Utrecht University, The Netherlands

Volume Editors

José Borbinha
National Library of Portugal
Campo Grande, 83, 1749-081 Lisboa, Portugal
E-mail: Jose.Borbinha@bn.pt

Thomas Baker
GMD Library, Schloss Birlinghoven
53754 Sankt Augustin, Germany
E-mail: Thomas.Baker@gmd.de

Cataloging-in-Publication Data applied for

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Research and advanced technology for digital libraries : 4th European conference ; proceedings / ECDL 2000, Lisbon, Portugal, September 18 - 20, 2000. José Borbinha ; Thomas Baker (ed.). - Berlin ; Heidelberg ; New York ; Barcelona ; Hong Kong ; London ; Milan ; Paris ; Singapore ; Tokyo : Springer, 2000
(Lecture notes in computer science ; Vol. 1923)
ISBN 3-540-41023-6

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

ISSN 0302-9743
ISBN 3-540-41023-6 Springer-Verlag 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-Verlag. Violations are liable for prosecution under the German Copyright Law.

Springer-Verlag Berlin Heidelberg New York
a member of BertelsmannSpringer Science+Business Media GmbH
© Springer-Verlag Berlin Heidelberg 2000
Printed in Germany

Typesetting: Camera-ready by author, data conversion by PTP-Berlin, Stefan Sossna
Printed on acid-free paper SPIN: 10722840 06/3142 5 4 3 2 1 0

Preface

ECDL2000, the Fourth European Conference on Research and Advanced Technology for Digital Libraries, is being held this year in Lisbon, Portugal, following previous events in Pisa (1997), Heraklion (1998), and Paris (1999). One major goal of the ECDL conference series has been to draw information professionals, stakeholders, and user communities from both the research world and from industry into a discussion of the alternative technologies, policies, and scenarios for global digital libraries. The success of previous conferences makes them a hard act to follow.

The field of digital libraries draws on a truly diverse set of scientific and technical disciplines. In the past three years, moreover, global cooperation on research and development has emerged as an urgent priority, particularly in the new European Framework Programme and in the Digital Library Initiative in the United States.

Because of this diversity, the field is perhaps still struggling for an identity. But this struggle for identity is itself a source of energy and creativity. Participants in this field feel themselves to be part of a special community, with special people. Each of us may claim expertise on a narrow issue, with specific projects, but the choices we make and the methods we use in local solutions can have unforeseen impacts within a growing universe of interconnected resources. This shared commitment to global interconnectedness gives the field a certain spirit, and this spirit in turn gives it a sense of coherence in the absence of a well-defined manifesto or established set of core themes. It also makes events such as ECDL2000 so very enjoyable!

This year we have tried to reflect this richness and diversity in the local organisation, in the Programme Committee, and in the final conference programme. The local organisation was handled by the National Library of Portugal; IST, the engineering institute of the Lisbon Technical University; and INESC, a research institution, with the additional support of DELOS, the European Network of Excellence for Digital Libraries. The Programme Committee includes members from 17 countries (eleven European Union countries plus Russia, USA, Australia, Japan, Morocco, and Brazil). The committee reflects a broad range of expertise from computer scientists and engineers to librarians, archivists, and managers.

Developing the conference programme has been a challenge. This year, 28 research papers address a cross-section of established topics in the field, while a new category of "short papers" provides an opportunity to explore emerging issues and re-evaluate the boundaries of our field. Soliciting short papers was an experiment, entailing some risk, but we felt this could help open the door to broader participation and perspectives. We are happy to see this interesting program complemented by some very high-quality panels, tutorials, keynote speakers, and workshops.

September 2000

José Borbinha
Thomas Baker

Organisation

ECDL2000 was organised in Lisbon, Portugal, by the National Library of Portugal, INESC (Instituto de Engenharia de Sistemas e Computadores), and IST (Instituto Superior Técnico – Universidade Técnica de Lisbon), and with the assistance of the DELOS Network of Excellence (supported by the Information Society Program of the European Commission).

General chair

José Tribolet, IST / INESC, Portugal

Programme and local chair

José Borbinha, BN / IST / INESC, Portugal

Workshops co-chairs

Thomas Baker, GMD, Germany
Alberto Silva, IST / INESC, Portugal

Tutorials co-chairs

Preben Hansen, SICS, Sweden
Eloy Rodrigues, Universidade do Minho, Portugal

Programme Committee

Philipe Agre, University of California, USA
Daniel E. Atkins, University of Michigan, USA
Joaquina Barrulas, INETI, Portugal
Rudolf Bayer, Technical University of Munich, Germany
Eurico Carrapatoso, University of Porto / INESC, Portugal
Rui Casteleiro, MIND S.A., Portugal
Panos Constantopoulos, FORTH, Greece
Ana M. R. Correia, UNL, Portugal
Bruce Croft, University of Massachusetts, USA
Murilo Bastos Cunha, University of Brasília, Brasil
Edward Fox, Virginia Tech, USA
Michael Freeston, University of Aberdeen, UK / University of California, USA
Norbert Fuhr, Dortmund University, Germany

Hachim Haddouti, debis Systemhaus Industry, Germany
Lynda Hardman, CWI, Netherlands
Cecília Henriques, AN / TT, National Archives, Portugal
Renato Iannella, IPR Systems Pty Ltd, Australia
Gareth Jones, University of Exeter, UK
Joaquim Jorge, IST / INESC, Portugal
Leonid Kalinichenko, Russian Academy of Sciences, Russia
Judith Klavans, Columbia University, USA
Carl Lagoze, Cornell University, USA
Alain Michard, INRIA, France
Marc Nanard, University of Montpellier, France
Erich J. Neuhold, GMD, Germany
Christos Nikolau, University of Crete / FORTH, Greece
Arlindo Oliveira, IST / INESC, Portugal
Mike Papazoglou, Tilburg University, Netherlands Carol Peters, CNR, Italy
Alexander Plemnek, St. Petersburg State Technical University / Soros Foundation, Russia
Mogens Sandfær, DTV, Denmark
Leonel Santos, University of Minho, Portugal
João Sequeira, BAD / RTP, Portugal
Mário Silva, University of Lisbon, Portugal
Miguel Silva, IST / INESC, Portugal
Alan Smeaton, Dublin City University, Ireland
Ingeborg Solvberg, University of Trondheim, Norway
Shigeo Sugimoto, University of Library and Information Science, Japan
Constantino Thanos, CNR, Italy
Isabel Trancoso, IST / INESC, Portugal
Anne-Marie Vercoustre, INRIA, France
Paula Viana, ISEP / INESC, Portugal
Howard Wactlar, Carnegie Mellon University, USA
Stuart Weibel, OCLC, USA
Jian Yang, Tilburg University, The Netherlands

Reviewers

Philipe Agre, University of California, USA
Daniel E Atkins, University of Michigan, USA
Rudolf Bayer, Technical University of Munich, Germany
Eurico Carrapatoso, University of Porto / INESC, Portugal
Rui Casteleiro, MIND S.A., Portugal
Panos Constantopoulos, FORTH, Greece
Ana M. R. Correia, UNL, Portugal
Bruce Croft, University of Massachusetts, USA
Murilo Bastos Cunha, University of Brasília, Brasil
Nicola Fanizzi, University of Bari, Italy
Stefano Ferilli, University of Bari, Italy

Edward Fox, Virginia Tech, USA
Nuno Freire, INESC / IST, Portugal
Michael Freeston, University of Aberdeen, UK / University of California, USA
Norbert Fuhr, Dortmund University, Germany
Hachim Haddouti, debis Systemhaus Industry, Germany
Cecília Henriques, AN / TT, National Archives, Portugal
Renato Iannella, IPR Systems Pty Ltd, Australia
Gareth Jones, University of Exeter, UK
Joaquim Jorge, IST / INESC, Portugal
Leonid Kalinichenko, Russian Academy of Sciences, Russia
Carl Lagoze, Cornell University, USA
Alain Michard, INRIA, France
Marc Nanard, University of Montpellier, France
Erich J. Neuhold, GMD, Germany
Christos Nikolau, University of Crete / FORTH, Greece
Arlindo Oliveira, IST / INESC, Portugal
Mike Papazoglou, Tilburg University, Netherlands
Carol Peters, CNR, Italy
Alexander Plemnek, St. Petersburg State Technical University / Soros Foundation, Russia
Mogens Sandfær, DTV, Denmark
Leonel Santos, University of Minho, Portugal
Giovanni Semeraro, University of Bari, Italy
João Sequeira, BAD / RTP, Portugal
Mário Silva, University of Lisbon, Portugal
Miguel Silva, IST / INESC, Portugal
Alan Smeaton, Dublin City University, Ireland
Ingeborg Solvberg, University of Trondheim, Norway
Shigeo Sugimoto, University of Library and Information Science, Japan
Constantino Thanos, CNR, Italy
Isabel Trancoso, IST / INESC, Portugal
Anne-Marie Vercoustre, INRIA, France
Paula Viana, ISEP / INESC, Portugal
Howard Wactlar, Carnegie Mellon University, USA
Jian Yang, Tilburg University, The Netherlands

Local Organising Committee

José Borbinha, BN / IST / INESC, Portugal
Eulália Carvalho, BN, Portugal
Fernando Cardoso, BN, Portugal
João Leal, BN, Portugal
João Maria Lopes, BN, Portugal
Rita Alves Pereira, BN, Portugal
Nuno Freire, INESC/IST, Portugal
Hugo Amorim, IST, Portugal

Table of Contents

Research Papers

Optical Recognition

Automatic Feature Extraction and Recognition for Digital Access of Books of the Renaissance 1
F. Muge, I. Granado, M. Mengucci, P. Pina, V. Ramos, N. Sirakov, J. R. Caldas Pinto, A. Marcolino, Mário Ramalho, P. Vieira, and A. Maia do Amaral

Content Based Indexing and Retrieval in a Digital Library of Arabic Scripts and Calligraphy 14
Suliman Al-Hawamdeh and Gul N. Khan

Ancient Music Recovery for Digital Libraries 24
J. Caldas Pinto, P. Vieira, M. Ramalho, M. Mengucci, P. Pina, and F. Muge

Probabilistic Automaton Model for Fuzzy English-Text Retrieval 35
Manabu Ohta, Atsuhiro Takasu, and Jun Adachi

Information Retrieval

Associative and Spatial Relationships in Thesaurus-Based Retrieval 45
Harith Alani, Christopher Jones, and Douglas Tudhope

Experiments on the Use of Feature Selection and Negative Evidence in Automated Text Categorization 59
Luigi Galavotti, Fabrizio Sebastiani, and Maria Simi

The Benefits of Displaying Additional Internal Document Information on Textual Database Search Result Lists 69
Offer Drori

Interactive-Time Similarity Search for Large Image Collections Using Parallel VA-Files 83
Roger Weber, Klemens Böhm, and Hans-J. Schek

Metadata

Dublin Core Metadata for Electronic Journals 93
Ann Apps and Ross MacIntyre

An Event-Aware Model for Metadata Interoperability 103
Carl Lagoze, Jane Hunter, and Dan Brickley

QUEST – Querying Specialized Collections on the Web 117
Martin Heß, Christian Mönch, and Oswald Drobnik

Personal Data in a Large Digital Library 127
José Manuel Barrueco Cruz, Markus J.R. Klink, and Thomas Krichel

Frameworks

Implementing a Reliable Digital Object Archive 128
Brian Cooper, Arturo Crespo, and Hector Garcia-Molina

Policy-Carrying, Policy-Enforcing Digital Objects 144
Sandra Payette and Carl Lagoze

INDIGO – An Approach to Infrastructures for Digital Libraries 158
Christian Mönch

Scalable Digital Libraries Based on NCSTRL/Dienst 168
Kurt Maly, Mohammad Zubair, Hesham Anan, Dun Tan, and Yunchuan Zhang

Multimedia

OMNIS/2: A Multimedia Meta System for Existing Digital Libraries 180
Günther Specht and Michael G. Bauer

Modeling Archival Repositories for Digital Libraries 190
Arturo Crespo and Hector Garcia-Molina

Implementation and Analysis of Several Keyframe-Based Browsing Interfaces to Digital Video 206
Hyowon Lee, Alan F. Smeaton, Catherine Berrut, Noel Murphy, Seán Marlow, and Noel E. O'Connor

Functional and Intentional Limitations of Interactivity on Content Indexing Topics: Possible Uses of Automatic Classification and Contents Extraction Systems, in Order to Create Digital Libraries Databases 219
Florent Pasquier

Users

Interaction Profiling in Digital Libraries through Learning Tools 229
G. Semeraro, F. Esposito, N. Fanizzi, and S. Ferilli

DEBORA: Developing an Interface to Support Collaboration in a Digital Library 239
David M. Nichols, Duncan Pemberton, Salah Dalhoumi, Omar Larouk, Claire Belisle, and Michael B. Twidale

Children as Design Partners and Testers for a Children's Digital Library . . 249
Yin Leng Theng, Norliza Mohd Nasir, Harold Thimbleby, George Buchanan, Matt Jones, David Bainbridge, and Noel Cassidy

Evaluating a User-Model Based Personalisation Architecture for Digital News Services 259
Alberto Díaz Esteban, Pablo Gervás Gómez-Navarro, and Antonio García Jiménez

Papers Complementing Invited Talks

Aging Links 269
Claudia Niederée, Ulrike Steffens, Joachim W. Schmidt, and Florian Matthes

Core Elements of Digital Gazetteers: Placenames, Categories, and Footprints 280
Linda L. Hill

The Application of an Event-Aware Metadata Model to an Online Oral History Archive 291
Jane Hunter and Darren James

From the Visual Book to the WEB Book: The Importance of Good Design 305
M. Landoni, R. Wilson, and F. Gibb

Short-Papers

Multimedia

Topic Detection in Read Documents 315
Rui Amaral and Isabel Trancoso

Map Segmentation by Colour Cube Genetic K-Mean Clustering 319
Victorino Ramos and Fernando Muge

Spoken Query Processing for Information Access in Digital Libraries 324
Fabio Crestani

A Metadata Model for Historical Documentary Films 328
Giuseppe Amato, Donatella Castelli, and Serena Pisani

Image Description and Retrieval Using MPEG-7 Shape Descriptors 332
Carla Zibreira and Fernando Pereira

A Large Scale Component-Based Multi-media Digital Library System - Development Experience and User Evaluation 336
Hiroshi Mukaiyama

Users in Digital Libraries

Personalised Delivery of News Articles from Multiple Sources 340
Gareth J. F. Jones, David J. Quested, and Katherine E. Thomson

Building a Digital Library of Web News 344
Nuno Maria and Mário J. Silva

Automatically Detecting and Organizing Documents into Topic Hierarchies: A Neural-Network Based Approach to Bookshelf Creation and Arrangement 348
Andreas Rauber, Michael Dittenbach, and Dieter Merkl

Daffodil: Distributed Agents for User-Friendly Access of Digital Libraries . 352
Norbert Gövert, Norbert Fuhr, and Claus-Peter Klas

An Adaptive Systems Approach to the Implementation and Evaluation of Digital Library Recommendation Systems 356
Johan Bollen and Luis M. Rocha

Are End-Users Satisfied by Using Digital Libraries? 360
Mounir A. Khalil

Information Retrieval

CAP7: Searching and Browsing in Distributed Document Collections 364
Norbert Fuhr, Kai Großjohann, and Stefan Kokkelink

Representing Context-Dependent Information Using *Multidimensional XML* 368
Yannis Stavrakas, Manolis Gergatsoulis, and Theodoros Mitakos

AQUA (Advanced Query User Interface Architecture) 372
László Kovács, András Micsik, Balázs Pataki, and István Zsámboki

Fusion of Overlapped Result Sets 376
Joaquim Macedo, António Costa, and Vasco Freitas

ActiveXML: Compound Documents for Integration of Heterogeneous Data Sources 380
João P. Campos and Mário J. Silva

newsWORKS©, the Complete Solution for Digital Press Clippings and Press Reviews: Capture of Information in an Intelligent Way 385
Begoña Aguilera Caballero and Richard Lehner

Internet Cataloguing

Effects of Cognitive and Problem Solving Style on Internet Search Tool . . . 389
Tek Yong Lim and Enya Kong Tang

Follow the Fox to Renardus: An Academic Subject Gateway Service for Europe 395
Lesly Huxley

CORC: Helping Libraries Take a Leading Role in the Digital Age 399
Kay Covert

Automatic Web Rating: Filtering Obscene Content on the Web 403
K. V. Chandrinos, Ion Androutsopoulos, G. Paliouras, and C. D. Spyropoulos

The Bibliographic Management of Web Documents in Digital and Hybrid Libraries 407
Wallace C. Koehler, Jr.

Technical Collections

The Economic Impact of an Electronic Journal Collection on an Academic Library 413
Carol Hansen Montgomery and John A. Bielec

A Comparative Transaction Log Analysis of Two Computing Collections . . 418
Malika Mahoui and Sally Jo Cunningham

ERAM - Digitisation of Classical Mathematical Publications 424
Hans Becker and Bernd Wegner

The Electronic Library in EMIS - European Mathematical Information Service 428
Bernd Wegner

Model for an Electronic Access to the Algerian Scientific Literature: Short Description 432
Bakelli Yahia

Cases 1

A Digital Library of Native American Images 437
Elaine Peterson

China Digital Library Initiative and Development 441
Michael B. Huang and Guohui Li

Appropriation of Legal Information: Evaluation of Data Bases for Researchers 445
Céline Hembise

Publishing 30 Years of the Legislation of Brazil's São Paulo State in CD-ROM and Internet 449
Paulo Leme, Dilson da Costa, Ricardo Baccarelli, Maurício Barbosa, Andréa Bolanho, Ana Reis, Rose Bicudo, Eduardo Barbosa, Márcio Nunes, Innocêncio Pereira Filho, Guilherme Plonski, and Sérgio Kobayashi

Electronic Dissemination of Statistical Information at Local Level: A Cooperative Project between a University Library and Other Public Institutions 452
Eugenio Pelizzari

Cases 2

Building Archaeological Photograph Library 456
Rei S. Atarashi, Masakazu Imai, Hideki Sunahara, Kunihiro Chihara, and Tadashi Katata

EULER - A DC-Based Integrated Access to Library Catalogues and Other Mathematics Information in the Web 461
Bernd Wegner

Decomate: Unified Access to Globally Distributed Libraries 467
Thomas Place and Jeroen Hoppenbrouwers

MADILIS, the Microsoft Access-Based Digital Library System 471
Scott Herrington and Philip Konomos

Leveraging Electronic Content: Electronic Linking Initiatives at Arizona State University 475
Dennis Brunning

Cases 3

Asian Film Connection: Developing a Scholarly Multilingual Digital Library - A Case Study 481
Marianne Afifi

Conceptual Model of Children's Electronic Textbook 485
Norshuhada Shiratuddin and Monica Landoni

An Information Food Chain for Advanced Applications on the WWW 490
Stefan Decker, Jan Jannink, Sergey Melnik, Prasenjit Mitra, Steffen Staab, Rudi Studer, and Gio Wiederhold

An Architecture for a Multi Criteria Exploration of a Documents Set 494
Patricia Dzeakou and Jean-Claude Derniame

An Open Digital Library Ordering System 498
Sarantos Kapidakis and Kostas Zorbadelos

Special Workshop

Special NKOS Workshop on Networked Knowledge Organization Systems . 502
Martin Doerr, Traugott Koch, Douglas Tudhope, and Repke de Vries

Implementing Electronic Journals in the Library and Making them Available to the End-User: An Integrated Approach 506
Gerrit Alewaeters, Serge Gilen, Paul Nieuwenhuysen, Stefaan Renard, and Marc Verpoorten

Author Index 511