

Lecture Notes in Computer Science

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

2163

Springer

Berlin

Heidelberg

New York

Barcelona

Hong Kong

London

Milan

Paris

Tokyo

Panos Constantopoulos
Ingeborg T. Sølvsberg (Eds.)

Research and Advanced Technology for Digital Libraries

5th European Conference, ECDL 2001
Darmstadt, Germany, September 4-9, 2001
Proceedings



Springer

Series Editors

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

Volume Editors

Panos Constantopoulos
University of Crete, Department of Computer Science
Leof. Knossou, P.O. Box 1470, 71409 Heraklion, Greece, and
Foundation for Research and Technology - Hellas
Institute of Computer Science
Vassilika Vouton, P.O. Box 1385, 71110 Heraklion, Greece
E-mail: panos@ics.forth.gr

Ingeborg T. Sølvsberg
The Norwegian University of Science and Technology
Department of Computer and Information Science
7491 Trondheim, Norway
E-mail: ingeborg.solvberg@idi.ntnu.no

Cataloging-in-Publication Data applied for

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Research and advanced technology for digital libraries : 5th European
conference ; proceedings / ECDL 2001, Darmstadt, Germany, September 4 - 9,
2001. Panos Constantopoulos ; Ingeborg T. Sølvsberg (ed.). - Berlin ; Heidelberg ;
New York ; Barcelona ; Hong Kong ; London ; Milan ; Paris ; Tokyo : Springer, 2001
(Lecture notes in computer science ; Vol. 2163)
ISBN 3-540-42537-3

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

ISSN 0302-9743

ISBN 3-540-42537-3 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

<http://www.springer.de>

© Springer-Verlag Berlin Heidelberg 2001
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Boller Mediendesign
Printed on acid-free paper SPIN 10840321 06/3142 5 4 3 2 1 0

Statement of the Chairpersons

ECDL 2001 was the 5th conference in the series of European Conferences on Research and Advanced Technology for Digital Libraries. Following previous events in Pisa (1997), Heraklion (1998), Paris (1999), and Lisbon (2000), this year ECDL was held in Darmstadt, Germany, and was co-organized by the Technical University of Darmstadt, GMD-IPSI, and Die Deutsche Bibliothek (the German National Library), Frankfurt. The overall objective of the ECDL series is to foster the formation of a multi-disciplinary scientific community in Europe specific to digital library research, development, and deployment. The ECDL series is supported by the DELOS Network of Excellence in Digital Libraries.

ECDL 2001 was seen as a good opportunity to review the impact that DLs have had on science, technology, and society in general. The call for papers therefore placed emphasis on applications of DL systems and on their integration into practical work. Looking to the future from the same perspective, we feel it is timely for the DL community to discuss such issues as the possible convergence of libraries, museums, and archives in their capacity as carriers of collective memories, which will yield large digital information spaces, and will encourage future co-operation in information spaces. This process seems, important as it is, to be only at its beginning. The Programme Committee, comprised of 36 members from 14 countries, selected 39 papers, 22 for regular and 17 for short presentations, out of a total of 79 submissions coming from 21 countries and 4 continents, in the following areas:

- User modelling
- Digitisation
- Interpretation and annotation of documents
- Knowledge management
- Data and metadata models
- Integration in user communities
- Information retrieval and filtering
- Multimedia digital libraries
- Multilinguality.

Besides the technical papers, the programme also featured three invited talks by Mike Keller, Stanford University, Eric Miller, W3C, and Dr. Michael Türkay, Senckenberg Museum, Frankfurt, three panel discussions, and a poster and demonstration session. Moreover, tutorials and workshops conveniently arranged immediately before and after the conference extended the range of subjects and opportunities for exchange of knowledge and opinions offered by the conference.

July 2001

P. Constantopoulos
E.J. Neuhold
E. Niggemann
I. Sølvsberg

Organisation

General Chairs

Erich J. Neuhold, GMD-IPSI, Darmstadt
Elisabeth Niggemann, Die Deutsche Bibliothek Frankfurt a. M.

Programme Chairs

Ingeborg Torvik Sølvsberg, Norwegian University of Science and Technology,
Trondheim
Panos Constantopoulos, FORTH, Crete

Demo & Poster Chairs

Dan Atkins, University of Michigan
Karsten Wendland, Darmstadt University of Technology, Darmstadt

Tutorial Chairs

Norbert Fuhr, University of Dortmund
Richard Furuta, Texas A&M University, USA

Workshop Chair

Carol Peters, IEI-CNR, Pisa

Organisation Chairs

Rudi Schmiede, Darmstadt University of Technology, Darmstadt
Matthias Hemmje, GMD-IPSI, Darmstadt

Publicity Chairs

Thomas Baker, GMD, Sankt Augustin
Nick Belkin, Rutgers University, USA

Programme Committee

T. Aalberg, NTNU, Norway
D. E. Atkins, University of Michigan, USA
Th. Baker, GMD, Germany
N. Belkin, Rutgers University, USA
J. Borbinha, INESC, Portugal
A. Brüggemann-Klein, Technical University of Munich, Germany
P. Constantopoulos, FORTH, Greece, (co-chair)
L. Dempsey, King's College London, UK
M. Doerr, FORTH, Greece
M. Dovey, Oxford Univ., UK
E. Fox, Virginia Tech, USA
N. Fuhr, University of Dortmund, Germany
R. Furuta, Texas A&M University, USA
H. Haddouti, Bavarian Research Center, Germany
P. Hansen, SICS, Sweden
L. Hardman, CWI, The Netherlands
M. Hedstrom, University of Michigan, USA
M. Hemmje, GMD, Germany
L. Hill, University of California at Santa Barbara, USA
O. Husby, BIBSYS, Norway
L. Kalinichenko, Russian Academy of Sciences, Russia
T. Koch, Lund University, Sweden
L. Kovacs, SZTAKI, Hungary
S. Krause, Germanisches Nationalmuseum, Germany
C. Lagoze, Cornell Univ., USA
P. Miller, UKOLN, UK
A. Paepcke, Stanford University, USA
C. Peters, IEI-CNR, Italy
C.J. van Rijsbergen, University of Glasgow, UK
L. Rold, National Museum, Denmark
Ch. Rusbridge, University of Warwick, UK
D. Soergel, University of Maryland, USA
I. Sølvsberg, NTNU, Norway, (co-chair)
S. Sugimoto, ULIS, Japan
C. Thanos, IEI-CNR, Italy
U. Thiel, GMD, Germany
A.-M. Vercoustre, CSIRO, Australia
S. Weibel, OCLC, USA

Steering Committee

Costantino Thanos, Chair, DELOS director, CNR-IEI, Italy
Yannis Yoannidis, DELOS Research Forum leader, University of Athens, Greece
Alan Smeaton, University of Dublin, Ireland
Hans-Jörg Scheck, ETHZ, Switzerland
Serge Abiteboul, ECDL 1999 Programme Chair, INRIA, France
Jose Borbinha, ECDL 2000 Programme Chair, BN-IST-INESC, Portugal
Erich Neuhold, ECDL 2001 General Chair, GMD-IPSI, Germany
Ingeborg Sølvsberg, ECDL 2001 Programme Chair, NTNU, Norway
Panos Constantopoulos, ECDL 2001 Programme Chair, FORTH, Crete
Nick Belkin, JCDL representative, Rutgers University, USA

Additional Reviewers

Andre Everts
Anna Tonazzini
Claudio Gennaro
Donna Bergmark
Fernando Adrian Das Neves
Ingo Frommholz
Giuseppe Amato
Marcos André Gonçalves
Jörg Cassens
Marcello L'Abbate
Mingfang Wu
Naomi Dushay

Table of Contents

User Modelling

Evaluating Electronic Textbooks: A Methodology	1
<i>Ruth Wilson, Monica Landoni</i>	
Search Behavior in a Research-Oriented Digital Library	13
<i>Malika Mahoui, Sally Jo Cunningham</i>	
A Combined Phrase and Thesaurus Browser for Large Document Collections	25
<i>Gordon W. Paynter, Ian H. Witten</i>	
Customizable Retrieval Functions Based on User Tasks in the Cultural Heritage Domain	37
<i>Holger Brocks, Ulrich Thiel, Adelheit Stein, Andrea Dirsch-Weigand</i>	

Digitisation, Interpretation, and Annotation of Documents

Digital Watermark	49
<i>Hamid Reza Mehrabi</i>	
Document Classification and Interpretation through the Inference of Logic-Based Models	59
<i>Giovanni Semeraro, Stefano Ferilli, Nicola Fanizzi, Floriana Esposito</i>	
The Cervantes Project: Steps to a Customizable and Interlinked On-Line Electronic Variorum Edition Supporting Scholarship	71
<i>Richard Furuta, Siddarth S. Kalasapur, Rajiv Kochumman, Eduardo Urbina, Ricardo Vivancos-Pérez</i>	
Fusion Approaches for Mappings between Heterogeneous Ontologies	83
<i>Thomas Mandl, Christa Womser-Hacker</i>	
Enhancing Digital Library Documents by A Posteriori Cross Linking Using XSLT	95
<i>Michael G. Bauer, Günther Specht</i>	
Using Copy-Detection and Text Comparison Algorithms for Cross-Referencing Multiple Editions of Literary Works	103
<i>Arkady Zaslavsky, Alejandro Bia, Krisztian Monostori</i>	

Knowledge Management I

- An Architecture for Automatic Reference Linking 115
Donna Bergmark, Carl Lagoze
- Disambiguating Geographic Names in a Historical Digital Library 127
David A. Smith, Gregory Crane
- Greenstone: A Platform for Distributed Digital Library Applications 137
David Bainbridge, George Buchanan, John McPherson, Steve Jones, Abdelaziz Mahoui, Ian H. Witten
- Linking Information with Distributed Objects 149
Trond Aalberg

Data and Metadata Models

- Metadata for Digital Preservation: A Review of Recent Developments 161
Michael Day
- MARIAN: Flexible Interoperability for Federated Digital Libraries 173
Marcos André Gonçalves, Robert K. France, Edward A. Fox
- Digital Libraries: A Generic Classification and Evaluation Scheme 187
Norbert Fuhr, Preben Hansen, Michael Mabe, Andras Micsik, Ingeborg Sølberg
- A Deposit for Digital Collections 200
Norman Noronha, João P. Campos, Daniel Gomes, Mário J. Silva, José Borbinha

Integration in User Communities

- Digital Libraries in a Clinical Setting: Friend or Foe? 213
Anne Adams, Ann Blandford
- Interactive, Domain-Independent Identification and Summarization of Topically Related News Articles 225
Dragomir R. Radev, Sasha Blair-Goldensohn, Zhu Zhang, Revathi Sundara Raghavan
- Digital Work Environment (DWE): Using Tasks to Organize Digital Resources 239
Narayanan Meyyappan, Suliman Al-Hawamdeh, Schubert Foo
- Learning Spaces in Digital Libraries 251
Anita S. Coleman, Terence R. Smith, Olha A. Buchel, Richard E. Mayer

Ethnography, Evaluation, and Design as Integrated Strategies: A Case Study from WES	263
<i>Michael Khoo</i>	

Dynamic Models of Expert Groups to Recommend Web Documents	275
<i>DaeEun Kim, Sea Woo Kim</i>	

Information Retrieval and Filtering

Enhancing Information Retrieval in Federated Bibliographic Data Sources Using Author Network Based Stratagems	287
<i>Peter Mutschke</i>	

Architecture for Event-Based Retrieval from Data Streams in Digital Libraries	300
<i>Mohamed Kholief, Stewart Shen, Kurt Maly</i>	

The Effects of the Relevance-Based Superimposition Model in Cross-Language Information Retrieval	312
<i>Teruhito Kanazawa, Akiko Aizawa, Atsuhiko Takasu, Jun Adachi</i>	

An On-Line Document Clustering Method Based on Forgetting Factors ...	325
<i>Yoshiharu Ishikawa, Yibing Chen, Hiroyuki Kitagawa</i>	

Knowledge Management II

Towards a Theory of Information Preservation	340
<i>James Cheney, Carl Lagoze, Peter Botticelli</i>	

C-Merge: A Tool for Policy-Based Merging of Resource Classifications ...	352
<i>Florian Matthes, Claudia Niederée, Ulrike Steffens</i>	

Truth in the Digital Library: From Ontological to Hermeneutical Systems .	366
<i>Aurélien Bénel, Elöd Egyed-Zsigmond, Yannick Prié, Sylvie Calabretto, Alain Mille, Andréa Iacovella, Jean-Marie Pinon</i>	

Multimedia Digital Libraries

XSL-based Content Management for Multi-presentation Digital Museum Exhibitions	378
<i>Jen-Shin Hong, Bai-Hsuen Chen, Jieh Hsiang</i>	

Iterative Design and Evaluation of a Geographic Digital Library for University Students: A Case Study of the Alexandria Digital Earth Prototype (ADEPT)	390
<i>Christine L. Borgman, Gregory H. Leazer, Anne J. Gilliland-Swetland, Rich Gazan</i>	

Automatically Analyzing and Organizing Music Archives	402
<i>Andreas Rauber, Markus Frühwirth</i>	

Building and Indexing a Distributed Multimedia Presentation Archive
Using SMIL 415
Jane Hunter, Suzanne Little

Multilinguality

Digitization, Coded Character Sets, and Optical Character Recognition for
Multi-script Information Resources: The Case of the *Letopis' Zhurnal'nykh*
Statei 429
George Andrew Spencer

Document Clustering and Language Models for System-Mediated
Information Access 438
Gheorghe Muresan, David J. Harper

Research and Development of Digital Libraries in China: Major Issues and
Trends 450
Guohui Li, Michael Bailou Huang

Panel 1

What's Holding Up the Development of Georeferenced DLs? 458
Michael Freeston, Linda L. Hill

Panel 2

Open Archive Initiative, Publishers and Scientific Societies: Future of
Publishing – Next Generation Publishing Models 459
Elisabeth Niggemann, Matthias Hemmje

Panel 3

Digital Library Programs: Current Status and Future Plans 460
Erich J. Neuhold

Author Index 461