

Lecture Notes in Computer Science

1590

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

Springer

Berlin

Heidelberg

New York

Barcelona

Hong Kong

London

Milan

Paris

Singapore

Tokyo

Paolo Atzeni Alberto Mendelzon
Giansalvatore Mecca (Eds.)

The World Wide Web and Databases

International Workshop WebDB'98
Valencia, Spain, March 27- 28, 1998
Selected Papers



Springer

Series Editors

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

Volume Editors

Paolo Atzeni
Università di Roma Tre
Dipartimento di Informatica e Automazione
via della Vasca Navale, 79, I-00146 Roma, Italy
E-mail: atzeni@dia.uniroma3.it

Alberto Mendelzon
University of Toronto, Department of Computer Science
6 King's College Road, Toronto, Ontario, Canada M5S 3H5
E-mail: mendel@cs.toronto.edu

Giansalvatore Mecca
Università della Basilicata
via della Tecnica, 3 85100 Potenza, Italy
E-mail: mecca@dia.uniroma3.it

Cataloging-in-Publication Data applied for

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

The world wide web and databases : selected papers / Workshop WebDB '98, Valencia, Spain, March 27 - 28, 1998. Paolo Atzeni ... (ed.). - Berlin ; Heidelberg ; New York ; Barcelona ; Hong Kong ; London ; Milan ; Paris ; Santa Clara ; Tokyo : Springer, 1999

(Lecture notes in computer science ; Vol. 1590)

ISBN 3-540-65890-4

CR Subject Classification (1998): H.5, H.4.3, H.2, H.3, C.2.4

ISSN 0302-9743

ISBN 3-540-65890-4 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 1999
Printed in Germany

Typesetting: Camera-ready by author

SPIN: 10704656

06/3142 - 5 4 3 2 1 0

Printed on acid-free paper

Foreword

This volume is based on the contributions to the International Workshop on the Web and Databases (WebDB'98), held in Valencia, Spain, March 27 and 28, 1998, in conjunction with the Sixth International Conference on Extending Database Technology (EDBT'98).

In response to the workshop call for papers, 37 manuscripts were submitted to the program committee. The review process was conducted entirely by e-mail. While the quality of submissions was generally high, only 16 papers could be accepted for presentation within the limited time allowed by the workshop schedule. Authors of workshop papers were invited to submit extended versions of their papers for publication in these post-workshop proceedings. The 13 papers appearing in this volume were submitted and selected after a second round of reviews.

We would like to thank the program committee of WebDB'98, all those who submitted their work, all additional reviewers, and the conference officials of EBDT'98 for their invaluable support. Special thanks go to Paolo Merialdo, who actively participated in the organization of the workshop.

February 1999

Paolo Atzeni, Alberto Mendelzon and Gianni Mecca
WebDB'98 Post-Workshop Proceedings Editors

Workshop Organization

Workshop Co-Chairs

Paolo Atzeni	(Università di Roma Tre, Italy)
Alberto Mendelzon	(University of Toronto, Canada)

Program Committee

Paolo Atzeni	(Università di Roma Tre, Italy)
Sophie Cluet	(INRIA, France)
Jon Kleinberg	(Cornell University, USA)
Alon Levy	(University of Washington, USA)
Udi Manber	(University of Arizona, USA)
Giansalvatore Mecca	(Università della Basilicata, Italy)
Alberto Mendelzon	(University of Toronto, Canada)
Eric Neuhold	(GMD-IPSI, Germany)
Oded Shmueli	(Technion, Israel)

Organizing Committee

Paolo Atzeni	(Università di Roma Tre, Italy)
Giansalvatore Mecca	(Università della Basilicata, Italy)
Alberto Mendelzon	(University of Toronto, Canada)
Paolo Merialdo	(Università di Roma Tre, Italy)

EDBT Workshop Coordinator

Oscar Pastor	(Universitat Politècnica de València)
--------------	---------------------------------------

Sponsoring Institutions

Universitat Politècnica de València
The EDBT Foundation
Università di Roma Tre

Table of Contents

Internet Programming: Tools and Applications

A Unified Algorithm for Cache Replacement and Consistency in Web Proxy Servers	1
<i>J. Shim, P. Scheuermann, R. Vingralek</i>	
Transactional Services for the Internet	14
<i>D. Billard</i>	
On the Unification of Persistent Programming and the World Wide Web ..	34
<i>R. Connor, K. Sibson, P. Manghi</i>	

Integration and Access to Web Data

Interactive Query and Search in Semistructured Databases	63
<i>R. Goldman, J. Widom</i>	
Bringing Database Functionality to the WWW	64
<i>D. Konopnicki, O. Shmueli</i>	
Fixpoint Calculus for Querying Semistructured Data	78
<i>N. Bidoit, M. Ykhlef</i>	

Hypertext Views on Databases

Incremental Maintenance of Hypertext Views	98
<i>G. Sindoni</i>	
Using YAT to Build a Web Server	118
<i>J. Siméon, S. Cluet</i>	
Languages and Tools to Specify Hypertext Views on Databases	136
<i>G. Falquet, J. Guyot, L. Nerima</i>	

Searching and Mining the Web

WebSuite – A Tool Suite for Harnessing Web Data	152
<i>C. Beeri, G. Elber, T. Milo, Y. Sagiv, O. Shmueli, N. Tishby, Y. Kogan, D. Konopnicki, P. Mogilevski, N. Slonim</i>	
Extracting Patterns and Relations from the World Wide Web	184
<i>S. Brin</i>	
WUM: A Tool for Web Utilization Analysis	185
<i>M. Spiliopoulou, L. C. Faulstich</i>	

Finding Near-Replicas of Documents on the Web 204
 N. Shivakumar, H. Garcia-Molina

Author Index 213