Advances in Web Intelligence

First International Atlantic Web Intelligence Conference
AWIC 2003
Madrid, Spain, May 5-6, 2003
Proceedings
Preface

We are pleased to present the proceedings of the 2003 Atlantic Web Intelligence Conference, AWIC 2003. The conference was located in Madrid, Spain during May 5–6, 2003, organized locally by the Technical University of Madrid.

AWIC 2003 aimed to be the first of a series of conferences on Web Intelligence, to be celebrated annually, alternatively in Europe and America, starting in Madrid. It was born as an activity of the recently created WIC-Poland Research Centre and the WIC-Spain Research Centre, both belonging to the Web Intelligence Consortium (WIC) (http://wi-consortium.org). AWIC 2003 was supported with grants from the Spanish Ministry for Science and Technology and the European Network of Excellence in Knowledge Discovery, KDNet.

AWIC 2003 brought together scientists, engineers, computer users, and students to exchange and share their experiences, new ideas, and research results about all aspects (theory, applications, and tools) of artificial intelligence techniques applied to Web-based systems, and to discuss the practical challenges encountered and the solutions adopted. Almost 70 contributions were submitted. After a preliminary evaluation, 60 of these papers were accepted to the conference and were assigned at least two reviewers from the international program committee. Out of this 60, 33 were conditionally accepted, and 32 of them were finally accepted after the conditions set by the reviewers had been met, which resulted in an acceptance ratio of 45%.

AWIC 2003 had a general session with 9 papers, and the others were organized into special sessions on areas such as:

– Web security,
– Semantic Web,
– Web authoring and design,
– Web information retrieval,
– Agents for the web, and
– User behavior

During AWIC 2003 we were pleased to have some important invited speakers. Prof. Lotfi Zadeh, who does not need any introduction, was the honorary chair and gave a talk about the transition of search engines to question-answering systems. This book contains what could be considered an extended abstract of his actual lecture. Prof. Ning Zhong, President of WIC, who can be considered as one of the fathers of the term “Web Intelligence,” gave a talk about the past, present, and future of the area. Finally, Prof. Ryszard Tadeusiewicz spoke about the application of artificial intelligence techniques for information retrieval.

We thank all the authors for contributing to what in our minds was a more than interesting program in Web Intelligence. We also want to thank the remarkable effort made by Coro Perez, secretary of the conference, and the rest of the local committee. We hope to see all of them, and many others at AWIC 2004.

May 2003

E. Menasalvas
J. Segovia
P.S. Szczepaniak
AWIC 2003 Organization

Honorary Chair
Lotfi A. Zadeh (University of California, USA)

Conference Co-chairs
Javier Segovia (Universidad Politécnica de Madrid, Spain)
Piotr Szczepaniak (Technical University of Lodz, Poland)

Steering Committee
Ning Zhong (Maebashi Institute of Technology, Japan)
Enric Trillas (Universidad Politécnica de Madrid, Spain)
Jiming Liu (Hong Kong Baptist University)
Janusz Kacprzyk (Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland)

Organizing Committee
Ernestina Menasalvas (Universidad Politécnica de Madrid, Spain)
Jesús Favela (Centro de Investigación Científica y de Educación Superior de Ensenada, México)

Local Committee
José M. Peña (Universidad Politécnica de Madrid, Spain)
Oscar Marbán (Universidad Carlos III de Madrid, Spain)
Ángel Lucas (Universidad Politécnica de Madrid, Spain)
Sonia Frutos (Universidad Politécnica de Madrid, Spain)
Javier Crespo (Universidad Carlos III de Madrid, Spain)
Eugenio Santos (Universidad Politécnica de Madrid, Spain)

Organization of the Special Track on Uncertainty, Hypermedia and the Web: Search, Personalization and Usability
José Angel Olivas (Universidad de Castilla-La Mancha, Spain)
Miguel-Angel Sicilia (Universidad Carlos III, Spain)
AWIC 2003 Program Committee

S.S. Anand, UK
R. Baeza-Yates, Northern Ireland, UK
T. Bao Ho, Japan
B. Berendt, Germany
A. Büchner, Northern Ireland, UK
J. Favela, México
U. Fayyad, USA
W. Gaul, Germany
J. Ghosh, USA
O. Guenther, Germany
M. Hadjimichael, USA
E. Hochsztain, Uruguay
P. Isasi, Spain
L. Joyanes Aguilar, Spain
J. Kacprzyk, Poland
S. Kaski, Finland
C. A. Knoblock, USA
D. H. Kraft, USA
R. Kruse, Germany
T. Y. Lin, USA
J. Liu, Hong Kong
L. Mason, USA
Y. Manolopoulos, Greece
S. Millán, Colombia
B. Mobasher, USA
C. Montes, Spain
O. Nasraoui, USA
M. Piattini, Spain
P. Quaresma, Portugal
R. Rastogi, USA
V. Robles, Spain
M. Rudnicki, Poland
J.W. Shavlik, USA
A. Skowron, Poland
M. Spiliopoulou, Germany
R. Tadeusiewicz, Poland
M.A. Vila, Spain
A. Wasilewska, USA
X. Wu, USA
R. Yager, USA
Y. Yao, Canada
P.S. Yu, USA
S. Zadrozny, Poland
M. Zaki, USA
W. Ziarko, Canada
Z. Zheng, USA
N. Zhong, Japan
# Table of Contents

## Invited Speakers

Toward Web Intelligence ................................................................. 1  
*Ning Zhong*

From Search Engines to Question-Answering Systems – The Need for New Tools ........................................................................... 15  
*Lotfi A. Zadeh*

Artificial Intelligence Techniques in Retrieval of Visual Data Semantic Information .......................................................... 18  
*Ryszard Tadeusiewicz, Marek R. Ogiela*

## General Session

A Framework to Integrate Business Goals in Web Usage Mining .............. 28  
*E. Hochsztain, S. Millán, B. Pardo, J.M. Peña, E. Menasalvas*

Clustered Organized Conceptual Queries in the Internet Using Fuzzy Interrelations ........................................................................ 37  
*Pablo J. Garcés, José A. Olivas, Francisco P. Romero*

Collaborative Filtering Using Interval Estimation Naïve Bayes .................. 46  
*V. Robles, P. Larrañaga, J.M. Peña, O. Marbán, J. Crespo, M.S. Pérez*

DJ-Boids: Flocks of DJ’s to Program Internet Multichannel Radio Stations ......................................................................................... 54  
*Jesus Ibanez, Antonio Gomez-Skarmeta, Josep Blat*

Evaluating the Informative Quality of Web Sites by Fuzzy Computing with Words .................................................................................. 62  
*Enrique Herrera-Viedma, Eduardo Peis, María Dolores Olvera, Juan Carlos Herrera, Yusef Hassan Montero*

Mining Association Rules Using Fuzzy Inference on Web Data .................. 73  
*Mariluz Martínez, Gelver Vargas, Andrés Dorado, Marta Millán*

The “Who Is It?” Problem Application for Customizable Web Sites ........... 83  
*Pierre-Emmanuel Jouve, Nicolas Nicoloyannis*
<table>
<thead>
<tr>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Using Case-Based Reasoning to Improve Information Retrieval in Knowledge Management Systems</td>
<td>94</td>
</tr>
<tr>
<td>Norbert Gronau, Frank Laskowski</td>
<td></td>
</tr>
<tr>
<td>Web Page Classification: A Soft Computing Approach</td>
<td>103</td>
</tr>
<tr>
<td>Angela Ribeiro, Víctor Fresno, María C. García-Alegre, Domingo Guinea</td>
<td></td>
</tr>
<tr>
<td>Special Session: Web Security</td>
<td></td>
</tr>
<tr>
<td>Content-Based Methodology for Anomaly Detection on the Web</td>
<td>113</td>
</tr>
<tr>
<td>Mark Last, Bracha Shapira, Yuval Elovici, Omer Zaafrany, Abraham Kandel</td>
<td></td>
</tr>
<tr>
<td>Designing Reliable Web Security Systems Using Rule-Based Systems</td>
<td>124</td>
</tr>
<tr>
<td>Grzegorz J. Nalepa, Antoni Ligęza</td>
<td></td>
</tr>
<tr>
<td>Secure Intelligent Agents Based on Formal Description Techniques</td>
<td>134</td>
</tr>
<tr>
<td>L. Mengual and C. de la Puente</td>
<td></td>
</tr>
<tr>
<td>Special Session: Semantic Web</td>
<td></td>
</tr>
<tr>
<td>A Natural Language Interface for Information Retrieval on Semantic Web Documents</td>
<td>142</td>
</tr>
<tr>
<td>Paulo Quaresma, Irene Pimenta Rodrigues</td>
<td></td>
</tr>
<tr>
<td>Conceptual User Tracking</td>
<td>155</td>
</tr>
<tr>
<td>Daniel Oberle, Bettina Berendt, Andreas Hotho, Jorge Gonzalez</td>
<td></td>
</tr>
<tr>
<td>Coping with Web Knowledge</td>
<td>165</td>
</tr>
<tr>
<td>J.L. Arjona, R. Corchuelo, J. Peña, D. Ruiz</td>
<td></td>
</tr>
<tr>
<td>Formalization of Web Design Patterns Using Ontologies</td>
<td>179</td>
</tr>
<tr>
<td>Susana Montero, Paloma Díaz, Ignacio Aedo</td>
<td></td>
</tr>
<tr>
<td>Special Session: Web Authoring and Design</td>
<td></td>
</tr>
<tr>
<td>A Multi-agent System for Web Document Authoring</td>
<td>189</td>
</tr>
<tr>
<td>M. Pérez-Coutiño, A. López-López, M. Montes-y-Gómez, L. Villaseñor-Pineda</td>
<td></td>
</tr>
</tbody>
</table>
On the Vague Modelling of Web Page Characteristics Regarding Usability .......................................................... 199
  Elena García Barrio canal, Miguel A. Sicilia Urbán, J. Antonio Gutiérrez

Semantic Interoperability Based on Ontology Mapping in Distributed Collaborative Design Environment ...................... 208
  Jingyi Zhang, Shensheng Zhang, Chun Chen, Bo Wang

The Use of Data Mining Techniques to Classify Styles and Types of Web-Based Information: Generation of Style Sheets Using Data Mining ...... 218
  Kathryn E.B. Thornton, Mark Carrington

Special Session: Web Information Retrieval

Building Topic Profiles Based on Expert Profile Aggregation ....................... 230
  Miguel Delgado, María J. Martín-Bautista, Daniel Sánchez, José M. Serrano, María-Amparo Vila

Carrot² and Language Properties in Web Search Results Clustering ................. 240
  Jerzy Stefanowski, Dawid Weiss

Practical Evaluation of Textual Fuzzy Similarity as a Tool for Information Retrieval .................................................... 250
  Piotr S. Szczepaniak, Marcin Gil

Rough – Fuzzy Reasoning for Customized Text Information Retrieval ............. 258
  Shailendra Singh, P. Dhanalakshmi, Lipika Dey

Special Session: Agents for the Web

A Machine Learning Based Evaluation of a Negotiation between Agents Involving Fuzzy Counter-Offers ............................ 268
  Javier Carbo, Agapito Ledezma

Autonomous Agents to Support Interoperability and Physical Integration in Pervasive Environments ................................. 278
  Marcela Rodríguez, Jesus Favela

Feature Selection Algorithms to Improve Documents’ Classification Performance ............................................................ 288
  Pedro A.C. Sousa, João Paulo Pimentão, Bruno René D. Santos, Fernando Moura-Pires

ROSA – Multi-agent System for Web Services Personalization ......................... 297
  Przemysław Kazienko, Maciej Kiewra
Supporting Software Maintenance in Web Repositories through a Multi-agent System ................................................................. 307
  Aurora Vizcaíno, Jesús Favela, Mario Piattini, Félix García

Special Session: User Behavior

A Hybrid Method for Patterns Mining and Outliers Detection in the Web Usage Log ................................................................. 318
  Mikhail Petrovskiy

Analysis of User’s Behaviour in Very Large Business Application Systems with Methods of the Web Usage Mining – A Case Study on SAP® R/3® ........................................ 329
  Jorge Marx Gómez, Gamal Kassem, Claus Rautenstrauch, Markus Melato

Automation of the Deep Web with User Defined Behaviours ............... 339
  Vicente Luque Centeno, Carlos Delgado Kloos, Peter T. Breuer,
  Luis Sánchez Fernández, Ma. Eugenia Gonzalo Cabellos,
  Juan Antonio Herráiz Pérez

Author Index ................................................................. 349