Proceedings of the Dateso 2011 Workshop

Databases, Texts
Specifications, and Objects

2011

http://www.ceur-ws.org/Vol-706/

Supported by


April 20 – 22, 2011
Písek
Workshop Partner

DATESO 2011 workshop proceedings were published in cooperation with:

SoSIREČR

Social Network of Computer Scientists in the Regions of the Czech Republic
The project aims to promote communication and establish cooperation between the Czech IT communities in academic and corporate sectors. The main output will be a Web portal that will build a social network of computer scientists, which will differ from the usual social and professional networks as Facebook or LinkedIn by providing unique tools to search for IT specialists, to create teams (for joint projects, tenders), and in general to allow cooperation. The aim of the project activities is to increase the competitiveness of the Czech Republic in the field of computer science, as well as improvement of the status of IT in the Czech Republic and its contribution to society.

Personal interconnection of educational and research activities and the cooperation of academic and applied research are required for high-quality tertiary education and its fruitful cooperation with industry and Government. From the social network will benefit students and graduates, scientific personnel, schools, and businesses. It will be possible to search here for a job or an interesting project, skilled workers or whole teams, even the schools can compare their learning plans with the situation on the labor market. Beta-version of the portal will be available from April 2011, but now it is possible to test partial functionality (scientific profile) on: http://www.sosirecr.cz/forms/sciprof/.

Another component of the project is to support personal meeting of the IT experts in the regions (both among themselves and with companies), the sharing of experience and the identification of needs. These are the conferences, regional seminars and workshops, aimed at improving communication between the academic, corporate and public sector. One of the following activities are for example Conversations with computer scientists (Hovory s informatiky), organised in the framework of this project by Institute of Computer Science, Academy of Sciences of the CR.

On the SoSIREČR project cooperate:

- Faculty of Mathematics and Physics, Charles University in Prague
- Institute of Computer Science, Academy of Sciences of the Czech Republic
- Czech Technical University in Prague
- The Faculty of Informatics and Statistics, University of Economics, Prague
- Higher Vocational School and Technical High School, Šumperk

Project (no CZ.07/2.4.00/12.0039) is running in period 01.11.2009 - 31.10.2012.

For more information please visit the sosirecr web pages.
Steering Committee

Václav Snášel VŠB-Technical University of Ostrava, Ostrava
Jaroslav Pokorný Charles University, Prague
Karel Richta Czech Technical University, Prague

Program Committee

Václav Snášel (chair) VŠB-Technical University of Ostrava, Ostrava
Jaroslav Pokorný Charles University, Prague
Karel Richta Czech Technical University, Prague
Vojtěch Svátek University of Economics, Prague
Peter Vojtáš Charles University, Prague
Dušan Húsek Inst. of Computer Science, Academy of Sciences, Prague
Michal Krátký VŠB-Technical University of Ostrava, Ostrava
Tomáš Skopál Charles University, Prague
Pavel Moravec VŠB-Technical University of Ostrava, Ostrava
Irena Mlynková Charles University, Prague
Michal Valenta Czech Technical University, Prague
Pavel Loupal Czech Technical University, Prague
Martin Nečaský Charles University, Prague
Jiří Dvorský VŠB-Technical University of Ostrava, Ostrava
Radim Bača VŠB-Technical University of Ostrava, Ostrava

Organizing Committee

Pavel Moravec VŠB-Technical University of Ostrava, Ostrava
Yveta Geletičová VŠB-Technical University of Ostrava, Ostrava
Barbora Klečková VŠB-Technical University of Ostrava, Ostrava
DATESO 2011, the international workshop on current trends on Databases, Information Retrieval, Algebraic Specification and Object Oriented Programming, was held on April 20 – 22, 2011 in Otav Arena hotel, Písek.

The 11th year was organized by Department of Computer Science VŠB-Technical University Ostrava, Department of Software Engineering MFF UK Praha, Department of Computer Science and Engineering FEL ČVUT Praha, and the Working group on Computer Science and Society of Czech Society for Cybernetics and Informatics. The DATESO workshops aim for strengthening connections between these various areas of Computer science. The proceedings of DATESO 2011 are also available at DATESO Web site: http://www.cs.vsb.cz/dateso/2011/ and CEUR Workshop Proceeding site: http://www.ceur-ws.org/Vol-706/ (ISSN 1613-0073). The Program Committee selected 22 papers (12 full papers, 4 special session full papers and 6 posters) from 33 submissions, based on three independent reviews.

The workshop program also included 2 invited lectures: Database Trends and Directions: Current Challenges and Opportunities by George Feuerlicht, and Simple Mathematical Models in Biometric Image Analysis by Khalid Saeed.

We wish to express our sincere thanks to all the authors who submitted papers, the members of the Program Committee, who reviewed them on the basis of originality, technical quality, and presentation. We are also thankful to the Organizing Committee and Amphora Research Group (ARG, http://www.cs.vsb.cz/arg/) for preparation of workshop and its proceedings as well as the Czech Society for Cybernetics and Informatics, SoSiReČR project and MIR Labs for their support of publishing this issue.

April, 2011 V. Snášel, J. Pokorný, K. Richta (Eds.)
# Table of Contents

## Full Papers

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top-k Search in Product Catalogues</td>
<td>1</td>
</tr>
<tr>
<td>Martin Šumák, Peter Gurský</td>
<td></td>
</tr>
<tr>
<td>XML Schema Integration with Reusable Schema Parts</td>
<td>13</td>
</tr>
<tr>
<td>Jakub Klímek, Jakub Malý, Martin Nečaský</td>
<td></td>
</tr>
<tr>
<td>Transformation of Binary Relationships with Particular Multiplicity</td>
<td>25</td>
</tr>
<tr>
<td>Zdeněk Rybola, Karel Richta</td>
<td></td>
</tr>
<tr>
<td>Towards e-Government Interoperability Framework</td>
<td>39</td>
</tr>
<tr>
<td>Jiří Feuerlicht, David Cunek</td>
<td></td>
</tr>
<tr>
<td>XML Document Versioning and Revalidation</td>
<td>49</td>
</tr>
<tr>
<td>Jakub Malý, Jakub Klímek, Irena Mlýnková, Martin Nečaský</td>
<td></td>
</tr>
<tr>
<td>XML Document Correction and XQuery Analysis with Analyzer</td>
<td>61</td>
</tr>
<tr>
<td>Jakub Stárka, Martin Svoboda, Jiří Schejbal, Irena Mlýnková, David</td>
<td></td>
</tr>
<tr>
<td>Bednárek</td>
<td></td>
</tr>
<tr>
<td>Supporting Language Interoperability by Dynamically Switched Behaviors</td>
<td>73</td>
</tr>
<tr>
<td>Jan Kurš, Jan Vraný, Alexandre Bergel</td>
<td></td>
</tr>
<tr>
<td>Task Scheduling in Data Stream Processing</td>
<td>85</td>
</tr>
<tr>
<td>Zbyněk Falt, Jakub Yaghob</td>
<td></td>
</tr>
<tr>
<td>Exploration of Semantic Spaces Obtained from Czech Corpora</td>
<td>97</td>
</tr>
<tr>
<td>Lubomír Krčmář, Miloslav Konopík, Karel Ježek</td>
<td></td>
</tr>
<tr>
<td>Using SVM and Clustering Algorithms in IDS Systems</td>
<td>108</td>
</tr>
<tr>
<td>Peter Scherer, Martin Vicher, Pavla Dráždilová, Jan Martinovič, Jiří</td>
<td></td>
</tr>
<tr>
<td>Dvorský, Václav Snášel</td>
<td></td>
</tr>
<tr>
<td>Combined Method for Effective Clustering based on Parallel SOM and</td>
<td>120</td>
</tr>
<tr>
<td>Spectral Clustering</td>
<td></td>
</tr>
<tr>
<td>Lukáš Vojáček, Jan Martinovič, Kateřina Slaninová, Pavla Dráždilová,</td>
<td></td>
</tr>
<tr>
<td>Jiří Dvorský</td>
<td></td>
</tr>
</tbody>
</table>

## Posters

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analysis of the DBLP Publication Classification Using Concept Lattices</td>
<td>132</td>
</tr>
<tr>
<td>Saleh Alwahaishi, Jan Martinovič, and Václav Snášel, Miloš Kudělka</td>
<td></td>
</tr>
</tbody>
</table>
Automatic Keyphrase Extraction based on NLP and Statistical Methods. 140
Martin Dostal, Karel Ježek

Flexible Cache for Database Management Systems ....................... 146
Radim Bača, David Bednář

Modeling of Lexical Relations Between Topics Retrieved from DBLP Journals ................................................................. 153
Lukáš Hlaváček, Michal Výmola

Overview and Comparison of Plagiarism Detection Tools .................. 161
Asim M. El Tahir Ali, Hussam M. Dahwa Abdulla, Václav Snášel

Précis Index Implementation for Efficient Fulltext Data Mining .......... 173
Michal Kopecký, Martin Čermák

Special Session: Data Processing for Geocomputation

Fuzzy Approach to Landslide Susceptibility Zonation ..................... 181
Miloš Marjanović, Jan Caha

Fractal Dimension Calculation for CORINE Land-Cover Evaluation in GIS – A Case Study ........................................................... 196
Vít Pászto, Lukáš Marek, Pavel Tuček

Determining Ecotones by Decision Support Systems ...................... 206
Vilém Pechanec, Jan Brus, Jan Caha

Orthophoto Map Feature Extraction Based on Neural Networks ........ 216
Zdeněk Horák, Miloš Kudělka, Václav Snášel, Vít Vojenílek

Invited Papers

Content-based Retrieval of Compressed Images .............................. 226
Gerald Schaefer

Simple Mathematical Models in Biometric Image Analysis .............. 227
Dr. Khalid Saeed, DSc, PhD, MSc, BSc Engg

Best Paper

Revision of Relational Joins for Multi-Core and Many-Core Architectures 229
Martin Kruliš, Jakub Yaghob

Author Index ................................................................. 241