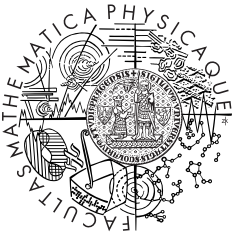
Charles University in Prague, MFF, Department of Software Engineering
Czech Technical University in Prague, FEE, Dept. of Computer Science & Eng.
VŠB–TU Ostrava, FEECS, Department of Computer Science
Czech Society for Cybernetics and Informatics

Proceedings of the Dateso 2007 Workshop

# DATESO 2007

Databases, Texts, Specifications, and Objects

http://www.cs.vsb.cz/dateso/2007/
http://www.ceur-ws.org/Vol-235/



April 18 – 20, 2007
Desná – Černá Říčka

DATESO 2007
© J. Pokorný, V. Snášel, K. Richta, editors

This work is subject to copyright. All rights reserved. Reproduction or publication of this material, even partial, is allowed only with the editors' permission.

Technical editor:
Pavel Moravec, pavel.moravec@vsb.cz
VŠB – Technical University of Ostrava
Faculty of Electrical Engineering and Computer Science
Department of Computer Science

| | |
|---|---|
| Page count: | 147 |
| Publication: | 196$^{\text{th}}$ |
| Impression: | 150 |
| Edition: | 1$^{st}$ |
| First published: | 2007 |

This proceedings was typeset by PDFLaTeX.
Cover design by Pavel Moravec (pavel.moravec@vsb.cz) and Tomáš Skopal.
Printed and bound in Ostrava, Czech Republic by TiskServis Jiří Pustina.

Published by MATFYZPRESS publishing house of Faculty of Mathematics and Physics Charles University, Ke Karlovu 3, 121 16 Praha 2, Czech Republic as its 196$^{\text{th}}$ publication.

# Preface

DATESO 2007, the international workshop on current trends on Databases, Information Retrieval, Algebraic Specification and Object Oriented Programming, was held on April 18 – 20, 2007 in Desná – Černá Říčka. This was the $7^{\text{th}}$ annual workshop organized by VŠB-Technical University Ostrava, Department of Computer Science, FEL ČVUT Praha, Department of Computer Science and Engineering and MFF UK Praha, Department of Software Engineering. The DATESO aims for strengthening the connection between this various areas of informatics. The proceedings of DATESO 2007 are also available at DATESO Web site `http://www.cs.vsb.cz/dateso/2007/` `http://www.ceur-ws.org/Vol-235/`.

The Program Committee selected 13 papers (11 full papers and 2 posters) from 23 submissions, based on two independent reviews.

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. Special thanks belong to Czech Society for Cybernetics and Informatics for its support of publishing this issue.

March, 2007 J. Pokorný, V. Snášel, K. Richta (Eds.)

## Program Committee

| | |
|---|---|
| Jaroslav Pokorný (chair) | Charles University, Prague |
| Václav Snášel | VŠB-Technical University of Ostrava, Ostrava |
| Karel Richta | Czech Technical University, Prague |
| Vojtěch Svátek | University of Economics, Prague |
| Peter Vojtáš | Charles University, Prague |
| Tomáš Skopal | Charles University, Prague |

## Organizing Committee

| | |
|---|---|
| Pavel Moravec | VŠB-Technical University of Ostrava |
| David Hoksza | Charles University, Prague |

# Table of Contents

## Full Papers

Syllable-Based Burrows-Wheeler Transform ..... 1
*Jan Lánský, Katsiaryna Chernik, Zuzana Vlčková*

Updating Typed XML Documents Using a Functional Data Model ..... 11
*Pavel Loupal*

Genetic Algorithms in Syllable-Based Text Compression ..... 21
*Tomáš Kuthan, Jan Lánský*

Using XSEM for Modeling XML Interfaces of Services in SOA ..... 35
*Martin Nečaský*

A Modular XQuery Implementation ..... 47
*Jan Vraný, Jan Žák*

A Content-Oriented Data Model for Semistructured Data ..... 55
*Tomáš Novotný*

Index-Based Approach to Similarity Search in Protein and Nucleotide Databases ..... 67
*David Hoksza, Tomáš Skopal*

Using BMH Algorithm to Solve Subset of XPath Queries ..... 81
*David Toth*

Shape Extraction Framework for Similarity Search in Image Databases .. 89
*Jan Klíma, Tomáš Skopal*

Inductive Models of User Preferences for Semantic Web ..... 103
*Alan Eckhardt*

Improvement of Text Compression Parameters Using Cluster Analysis ... 115
*Jiří Dvorský, Jan Martinovič*

## Posters

Work with Knowledge on the Internet – Local Search ..... 127
*Antonín Pavlíček, Josef Muknšnábl*

The Use of Ontologies in Wrapper Induction ..... 132
*Marek Nekvasil*

**Author Index** ..... 136