PREFACE

Special section on
Exploring Complex and Big Data

In its over 20 years of presence in the research space, the conference on *Advances in Databases and Information Systems (ADBIS)* has been accompanied by workshops. Its 21st edition was held in Cyprus on 24–27 September 2017, with four workshops affiliated. This special section includes four best papers (substantially extended) from the workshops and two external papers. They cover a wide spectrum of approaches and techniques to data storing, modelling, and processing.

The first paper, entitled *Introducing narratives in Europeana: A case study* (by C. Meghini, V. Bartalesi, D. Metilli, and F. Benedetti), focuses on enriching digital libraries with networks of semantic data, called narratives. The purpose of using narratives is to offer the user a broader context for his/her search. The authors discuss possible extensions to a digital library data model to represent narratives, followed by an implementation of prototype software. They study the applicability of the proposed solution in the Europeana digital library—the largest European digital library in the cultural heritage domain.

The second paper, entitled *Ontology-based access to temporal data with Ontop: A framework proposal* (by E. Güzel Kalayci, S. Brandt, D. Calvanese, V. Ryzhikov, G. Xiao, and M. Zakharyaschev), proposes to use ontologies for providing additional semantics of data in order to support advanced data analysis. As a basis, the concept of ontology-based data access (OBDA) is used, which serves as a conceptual view of the data enriched with an ontology. The OBDA is extended to handle temporal data, i.e., those evolving over time. To this end, the authors propose (i) extended mapping languages to extract data about temporal events, (ii) classical ontology and rule languages for static data, (iii) a temporal rule language, (iv) a SPARQL-based query language for retrieving temporal data, and (v) a prototype system architecture.

Temporal data analysis is also addressed in the external paper entitled *Modeling and querying facts with period timestamps in data warehouses* (by G. Mahlknecht, A. Dignös, and N. Kozmina). The authors focus on representing data with an associated valid time period in a data warehouse context. To this end, they propose three distinct models called an instant model, a period model, and a period* model. The authors study the impact of these models on four classes of analytical queries, focusing their investigation on query construction and performance.

The fourth paper, entitled *Fusion of clinical data: A case study to predict the type of treatment of bone fractures* (by A. Haq, S. Wilk, and A. Abelló), proposes a solution for integrating multiple types of medical data for their analysis. To this end, the authors apply six different data fusion models: one based on a combination of data (COD) and five based on the combination of interpretation (COI). The proposed fusion models are experimentally evaluated in a use case of predicting the type of treatment for patients with bone fractures. The obtained results show that the constructed decision models based on COI are more accurate than those based on COD.

The fifth paper, entitled *Parallelizing user-defined functions in the ETL workflow using orchestration style sheets* (by S.M.F. Ali, J. Mey, and M. Thiele), addresses the problem of optimizing data integration workflows composed of user-defined functions. The authors propose to apply parallelization by means of the so-called orchestration style sheets (OSSs). OSSs allow separating UDF implementation from a particular parallelization framework and a parallel execution environment. Then, based on parallel code templates, configuration files, and OSSs, an executable user code can be generated for a given parallelization framework and execution environment. The proposed technique was tested in an EC2 cloud, proving its applicability.

The sixth paper, entitled *Exploiting multi-core and many-core parallelism for subspace clustering* (by A. Datta, A. Kaur, T. Lauer, and S. Chabbouh), discusses the applicability of multi-core processors for discovering clusters in high dimensional data spaces. In particular, the authors focus on the Subscale clustering algorithm and show that it can be parallelized and executed in a multi-core hardware, including GPUs.

The guest editors of this special section express special thanks to the journal Editor-in-Chief, Prof. Józef Korbicz, for accepting this special section as well as for excellent co-operation and assistance. Also, the support for R. Wrembel from the National Science Center in Poland under the grant no. 2015/19/B/ST6/02637 is kindly acknowledged.

Johann Gamper

Faculty of Computer Science
Free University of Bozen-Bolzano
Italy

Robert Wrembel

Faculty of Computing
Poznan University of Technology
Poland

March 2019

Johann Gamper is a full professor of databases at the Faculty of Computer Science of the Free University of Bozen-Bolzano. He received his MSc degree from TU Vienna in 1989 and a PhD degree from RWTH Aachen in 1996, both in computer science. His main research interests are in data intensive systems, with a focus on database technologies for processing and analysing huge amounts of data, and a particular emphasis on time-referenced and time-series data. Many of his research activities are embedded in projects with local and international companies and institutions from various sectors, such as agriculture, medicine, public administration and tourism. Johann Gamper is an author of 120+ publications, many of which have been published in the most prestigious outlets of the database community. He is a member of the ACM and IEEE and regularly serves as a reviewer for technical journals, as well as a PC member and an organizer of conferences. More details are available at `http://www.inf.unibz.it/~gamper/`.

Robert Wrembel is an associate professor at the Faculty of Computing of the Poznan University of Technology, Poland. His main research is in data warehouse systems. He has been involved in 7 research projects in the area of databases and data warehouses as well as 7 industrial projects in the field of information technologies. He has designed and delivered numerous courses in the area of programming languages, database administration, and information systems design for multiple companies and institutions in Poland, including Oracle, Microsoft, IBM, BAE Systems. He has been a visiting professor at Université Lyon 2 (Lyon, France) and Loyola University (New Orleans, USA), an invited lecturer at Universidad de Costa Rica (San Jose, Costa Rica), and has completed a post-graduate programme in entrepreneurship and innovation at Stanford University. He has also been an intern at the BI company Targit (Tampa, USA), a prizewinner of the IBM Faculty Award for highly competitive research, a regular editorial board member of the *Data & Knowledge Engineering* journal, and a member of the IFIP Working Group 2.6 (Database). More details are available at `http://www.cs.put.poznan.pl/rwrembel`.