

Model and Data Engineering, MEDI 2016 Special Issue

Preface

At the 6th edition of the International Conference on Model & Data Engineering (MEDI 2016) that took place in Aguadulce, Almera, Spain, on September 21-23, 2016, participants from around the world met to exchange information in the area of data and model engineering, to provide a forum for the dissemination of research accomplishments, and to promote the interaction and collaboration between the models and data research communities. This international scientific event, initiated by researchers from Euro-Mediterranean countries, aims also at promoting the creation of north-south scientific networks, projects and faculty/student exchanges. MEDI 2016 follows the success of the Óbidos (Portugal, 2011), Poitiers (France, 2012) and Armantea (Italy, 2013), Larnaca (Cyprus 2014), Island of Rhodes (Greece 2015) editions.

MEDI 2016 received 62 submissions covering both model and data engineering activities. These papers range on a wide spectrum covering fundamental contributions, applications and tool developments and improvements. Each paper was reviewed by at least three reviewers and the programme committee accepted 18 long papers and 10 short papers leading to an attractive scientific programme.

This special issue of the *Fundamenta Informaticae* journal includes the best papers selected from the MEDI 2016 conference which covered the topics of the journal. The authors of these papers have extended their original work presented at MEDI 2016 conference with significant minimum 30% additional material added.

We invited 8 regular papers from MEDI 2016 to our special issue of *Fundamenta Informaticae*. After the second round of reviews, we finally accepted 5 papers that reflect the current advances and development of model and data engineering. We congratulate the authors of these 5 papers and also thank all authors who submitted articles to MEDI 2016. The topics of the accepted papers cover eleven main themes: social network, data intensive flows, opinion analysis, quality modeling, data warehouse design, Linked Open Data.

The first paper, titled *Gender-Based Analysis of Intra-Institutional Research Productivity and Collaboration*, by Miloš Savić, Mirjana Ivanović, Miloš Radovanović and Bojana Dimić Surla, addresses a crucial issue related the measurement of gender differences in research and teaching productivity. They proposed GERBER (GEndeR Based Evaluation of Researchers), a methodology and accompanying tool for performing gender-based analysis of CRIS (Current Research Information Systems) data. It is an interesting state of art on Scientometric Analysis of Gender in Research and a presentation of the CRIS UNS that represents the information system for storing and managing data about scientific research activity at the University of Novi Sad (UNS), Serbia. This work is a part of the

project no. OI174023, Intelligent techniques and their integration into wide-spectrum decision support, with additional support obtained in cooperation with the Slovenian Research Agency through bilateral project no. 451-03-3095/2014-09/43, Culture sensitive aspects in data technologies. GERBER enables the extraction of co-authorship networks, computation of various author metrics, and statistical comparison of male and female researchers. Functionality of GERBER is demonstrated on data extracted from the CRIS of the University of Novi Sad. A plan to integrate GERBER into CRIS UNS in order to facilitate continuous gender-based researcher evaluation is widely discussed.

The second paper titled *Intermediate Results Materialization Selection and Format for Data-Intensive Flows*, by Rana Faisal Munir, Sergi Nadal, Oscar Romero, Alberto Abello, Petar Jovanovic, Maik Thiele and Wolfgang Lehner tackles two challenging problems related to data intensive flows (DIF): (1) how to decide on which intermediate data to materialize taking into account multiple and conflicting criteria, and (2) how to choose among the various alternatives regarding storage. ETL and data Ingestion processes can be modeled by DIF. A motivating example is given which illustrates the identified problems and hints of their resolution. From theoretical perspective, the paper proposes a reduction of the problem of intermediate results materialization in DIF to the traditional materialized view selection problem, widely studied in the context of data warehouse and known to be NP-hard. As a consequence, a set of heuristic rules that selects an appropriate format for an intermediate result is nicely presented. The experimental results show that our approach provides 40% better average speedup with respect to the current state-of-the-art, as well as an improvement on disk access time of 18% as compared to fixed format solutions.

The third paper, *Impact of Credibility on Opinion Analysis in Social Media*, by Fatima Zohra Ennaji, Lobna Azaza, Zakaria Maamar, Abdelaziz El Fazziki, Marinette Savonnet, Mohamed Sadgal, Eric Leclercq, Idir Amine Amarouche and Djamel Benslimane, discusses how to extract and consolidate customers' opinions so that enterprises understand their needs and concerns well. To this end, the authors use hierarchical agglomerative clustering to filter opinions that could be related to virtual and malicious users and, also, use a credibility model to measure to what extent an expressed opinion is true. During opinion analysis, user multi-identity and limited credibility could undermine this analysis. A set of experiments demonstrating the technical feasibility of the work is discussed in the paper.

The fourth paper titled, *Quality-Aware Architectural Model Transformations in Adaptive Mashups User Interfaces*, by Javier Criado, Silverio Martinez-Fernandez, David Ameller, Luis Iribarne, Nicolás Padilla and Andreas Jedlitschka, presents a generic quality-aware transformation process to support the adaptation of software architectures. The proposed solution is applied in ENIA (ENvironmental Information Agent, <http://acg.ual.es/enia>) software through developing a quality model for the adaptation of mashup user interfaces. In addition, the proposed model is also used in evaluation of other transformation approaches. The quality model has been built from the application of the ISO/IEC 25010 standard and the Quamoco (Open Quality Model and Tool Support for Quality Modelling and Evaluation, <http://www.quamoco.de/>) approach (that offers modularisation and an editor for creating, adjusting, and managing quality models for various application domains in contrast to fixed, predefined models) by applying the Goal Question Metric (GQM) strategy. The proposed solution is validated through four scenarios to demonstrate how it can be applied to maximize the scope, the attractiveness, the customization or the cataloging, among other possible examples.

The fifth paper titled *A Unified Approach to Multisource Data Analyses*, by Franck Ravat and Jiefu Song revisits the problem of designing data warehouses in the era of Linked Open Data (LOD). The integration of the LOD in the data warehouse construction may increase its final value. As a

consequence, it proposes a conceptual multidimensional model aiming to provide a uniform vision of warehoused data and the LOD. The paper also describes an implementation framework compatible with the proposed model Unified Cubes. The framework is composed of two modules: schema and instance modules. The schema module manages the overall structure of a "Unified Cube", while the instance module serves as an integrated repository of equivalent data instances from different sources. The framework automatically generates queries in one or multiple sources based on the measures and the attributes chosen by decision-makers and unifies the results into one final result. The problem studied in this paper opens several perspectives for the data warehouse community covering different phases of their lifecycle: requirements, conceptual, logical, ETL, physical and exploitation phases.

We hope readers will find the content of this special issue interesting and will inspire them to look further into the challenges that are still ahead before designing advanced database applications and complex systems and software. We would like to thank all the authors who submitted their papers to this special issue. In addition, we are grateful for the support of various reviews that ensured the high quality of this special issue. Last but not least, we would like to thank Professor Damian Niwinski, Editor-In-Chief of *Fundamental Informaticae Journal* for accepting our proposal of a special issue focused on models and data engineering and promoting MEDI conference, and for assisting us whenever required. The complete International Program Committee of this Special Issue is listed next.

International Program Committee:

- Cuccuru Arnaud, CEA LIST, Paris, France
- Moh Boughanem, IRIT University Paul Sabatier Toulouse
- Anastasios Gounaris, Aristotle University of Thessaloniki
- Antonia Gogoglou, Aristotle University of Thessaloniki, Greece
- Nantia Iakovidou, Aristotle University of Thessaloniki, Greece
- George Kakarontzas, Aristotle University of Thessaloniki
- Selma Khouri, ESI, Algiers, Algeria
- Amel Mammam, TELECOM SudParis, Paris, France
- Samir Ouchani, CESI, France
- Neeraj Singh, INPT-ENSEEIH/IRIT, University of Toulouse, France
- Spyros Sioutas, Ionian University, Corfu, Greece
- Panos Vassiliadis, University of Ioannina, Greece

The Special Issue Editors:

Ladjel Bellatreche

LIAS/ISAE-ENSMA - Poitiers University
bellatreche@ensma.fr

Zbigniew W. Ras

University of North Carolina at Charlotte and
Polish-Japanese Academy of Information Technology
ras@uncc.edu