

# **Lecture Notes in Business Information Processing**

**253**

## Series Editors

Wil van der Aalst

*Eindhoven Technical University, Eindhoven, The Netherlands*

John Mylopoulos

*University of Trento, Povo, Italy*

Michael Rosemann

*Queensland University of Technology, Brisbane, QLD, Australia*

Michael J. Shaw

*University of Illinois, Urbana-Champaign, IL, USA*

Clemens Szyperski

*Microsoft Research, Redmond, WA, USA*

More information about this series at <http://www.springer.com/series/7911>

Esteban Zimányi · Alberto Abelló (Eds.)

# Business Intelligence

5th European Summer School, eBISS 2015  
Barcelona, Spain, July 5–10, 2015  
Tutorial Lectures

*Editors*

Esteban Zimányi  
Department of Computer and Decision  
Engineering  
Universite Libre de Bruxelles  
Brussels  
Belgium

Alberto Abelló  
Database Technologies and Information  
Management Group  
Universitat Politècnica de Catalunya  
Barcelona  
Spain

ISSN 1865-1348                      ISSN 1865-1356 (electronic)  
Lecture Notes in Business Information Processing  
ISBN 978-3-319-39242-4              ISBN 978-3-319-39243-1 (eBook)  
DOI 10.1007/978-3-319-39243-1

Library of Congress Control Number: 2016939568

© Springer International Publishing Switzerland 2016

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

This Springer imprint is published by Springer Nature  
The registered company is Springer International Publishing AG Switzerland

# Preface

The 5th European Business Intelligence Summer School (eBISS 2015) took place in Barcelona, Spain, in July 2015. Tutorials were given by renowned experts and covered several recent topics in business intelligence. This volume contains the lecture notes of the summer school.

The first paper reports on the state of the art (including research and tools) in relational database schema evolution (with special emphasis on data warehouse evolution) and how this impacts the surrounding applications. It includes a discussion on query-rewriting techniques to adapt database client software to the changes.

The second paper compares RDF Data Cube Vocabulary (QB) with QB4OLAP, an extension to support OLAP analysis. The former is an initiative of W3C to publish multi-dimensional data on the Web in such a way that it can be linked to related data sets and concepts. The later extends it to explicit aggregation functions, hierarchies, and descriptive attributes. Such extension facilitates the definition of a user-friendly query language (QL) that allows OLAP users not familiar with SW concepts or languages to retrieve data without any knowledge of RDF or SPARQL (QL is automatically translated into SPARQL queries). Examples of usage of QB4OLAP on Eurostat data sets are provided.

Next, the third paper presents a concrete application domain for linked open data analysis, which is social information. It presents a proposal for modeling social (mainly textual) data, so that it can be analyzed with OLAP tools. The complexity of the problem justifies the proposal of a software architecture and methodology for the management of this kind of project. The feasibility of the proposed approach is analyzed in the context of two specific projects: one in the subject area of Italian politics, and another in the subject area of a large consumer goods company.

The fourth paper explores the possibility of including linked open data in analytical tasks. Traditional data warehousing has relied on internal data to enable decision making. Nevertheless, more recent big data trends have moved the focus to external data. Success stories based on the use of data coming from social networks are well known, but we can also benefit from publicly available semantically annotated data, which is growing fast mainly but not exclusively with governmental support. In any case, the integration of external data presents new challenges in terms of lack of structure, high heterogeneity, and poor quality.

The last paper discusses the feasibility and importance of deriving key performance indicator (KPI) calculations (i.e., aggregate queries) from their informal specifications. Since the majority of KPIs are process-oriented, process models (i.e., Petri nets) are used. Thus, seven different patterns are identified, which relate query elements to process models tasks.

In addition to the lectures corresponding to the papers described here, eBISS 2015 had two other lectures directly related to industry:

- Toni Cebrián from Enerbyte, Spain: “Time Series DBs and Streaming Algorithms”
- Wilinton Tenorio and Eduard Gil from ClearPeaks, Spain: “Life at ClearPeaks, An Overview of the Most Relevant Projects”

These lectures are not included in this volume.

In this edition, eBISS joined forces with the Erasmus Mundus IT4BI-DC consortium and hosted its doctoral colloquium aiming at community building and promoting a corporate spirit among PhD candidates, advisors, and researchers of different organizations. The corresponding session, organized in two parallel tracks, included eight presentations, as follows:

- Waqas Ahmed, Pakistan: “Modeling Data Warehouses with Multiversion and Temporal Functionality”
- Nurefsan Gur, Turkey: “Business Intelligence over Linked Open Spatio-Temporal Data”
- Dilshod Ibragimov, Uzbekistan: “OLAP over Distributed RDF Sources”
- Azadeh Nasiri, Iran: “Requirements Engineering for Big Data Predictive Analytics”
- Bijay Neupane, Nepal: “Intelligence Detection and Prediction of Energy at the Device Level”
- Kasun Parera, Sri Lanka: “Model-Based Database Systems”
- Vasileios Theodorou, Greece: “Automating User-Centered Design of Data-Intensive Processes”
- Jovan Varga, Serbia: “Discovering Analytical Concepts from User Profiles”

We would like to thank the attendees of the summer school for their active participation, as well as the speakers and their co-authors for the high quality of their contribution in a constantly evolving and highly competitive domain. Finally, the lectures in this volume benefited greatly from the comments of the external reviewers.

March 2016

Esteban Zimányi  
Alberto Abelló

# Organization

The 5th European Business Intelligence Summer School (eBISS 2015) was organized by the Department of Computer and Decision Engineering (CoDE) of the Université Libre de Bruxelles, Belgium, and the Database Technologies and Information Management Group of the Universitat Politècnica de Catalunya.

## Program Committee

|                     |  |
|---------------------|--|
| Alberto Abelló      | Universitat Politècnica de Catalunya, BarcelonaTech, Spain |
| Marie-Aude Aufaure  | Ecole Centrale de Paris, France                            |
| Ralf-Detlef Kutsche | Technische Universität Berlin, Germany                     |
| Patrick Marcel      | Université François Rabelais de Tours, France              |
| Esteban Zimányi     | Université Libre de Bruxelles, Belgique                    |

## External Reviewers

|                       |   |
|-----------------------|---|
| Ahmed Ahmedov         | Technische Universität Dresden, Germany     |
| Waqas Ahmed           | Université Libre de Bruxelles, Belgium      |
| Gastón Bakkalian      | Poznan University of Technology, Poland     |
| Besim Bilalli         | Universitat Politècnica de Catalunya, Spain |
| Dilshod Ibragimov     | Université Libre de Bruxelles, Belgium      |
| Mohammed Idris        | Université Libre de Bruxelles, Belgium      |
| Faisal Orakzai        | Université Libre de Bruxelles, Belgium      |
| Mahmoud Sakr          | Université Libre de Bruxelles, Belgium      |
| Muhammad Aamir Saleem | Aalborg University, Denmark                 |
| Vasileios Theodorou   | Universitat Politècnica de Catalunya        |
| Martin Ugarte         | Université Libre de Bruxelles, Belgium      |

# Contents

|  |     |
|--|-----|
| Schema Evolution for Databases and Data Warehouses . . . . .                               | 1   |
| <i>Petros Manousis, Panos Vassiliadis, Apostolos Zarras,<br/>and George Papastefanatos</i> |     |
| Publishing OLAP Cubes on the Semantic Web . . . . .  | 32  |
| <i>Alejandro Vaisman</i>   |     |
| Design Issues in Social Business Intelligence Projects . . . . .                           | 62  |
| <i>Matteo Golfarelli</i>   |     |
| Context-Aware Business Intelligence . . . . .  | 87  |
| <i>Rafael Berlanga and Victoria Nebot</i>  |     |
| Key Performance Indicators in Data Warehouses . . . . .                                    | 111 |
| <i>Manfred A. Jeusfeld and Samsethy Thoun</i>  |     |
| <b>Author Index</b> . . . . .  | 131 |