

Lecture Notes in Business Information Processing

96

Series Editors

Wil van der Aalst

Eindhoven Technical University, The Netherlands

John Mylopoulos

University of Trento, 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

Marie-Aude Aufaure
Esteban Zimányi (Eds.)

Business Intelligence

First European Summer School, eBISS 2011
Paris, France, July 3-8, 2011
Tutorial Lectures

Volume Editors

Marie-Aude Aufaure
Ecole Centrale Paris
MAS Laboratory
Châtenay-Malabry Cedex, France
E-mail: marie-aude.aufaure@ecp.fr

Esteban Zimányi
Universite Libre de Bruxelles
Department of Computer & Decision Engineering (CoDE)
Brussels, Belgium
E-mail: ezimanyi@ulb.ac.be

ISSN 1865-1348
ISBN 978-3-642-27357-5
DOI 10.1007/978-3-642-27358-2
e-ISSN 1865-1356
e-ISBN 978-3-642-27358-2
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2011944231

ACM Computing Classification (1998): J.1, H.2, H.3, D.1

© Springer-Verlag Berlin Heidelberg 2012

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

The use of general descriptive names, registered names, trademarks, 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.

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The First European Business Intelligence Summer School (eBISS 2011) attracted experts from academia and industry as well as PhD students interested in foundational and applicational aspects of business intelligence (BI). This volume contains the lecture notes of the summer school, which took place in Paris, France, in July 2011.

The first chapter reviews consolidated results in data warehouses and describes open research fields. The latter include the need to cope with more complex data, both in structure and semantics, and keeping up with the demands of new application domains such as Web, financial, manufacturing, life sciences, multimedia, and spatiotemporal applications.

The second chapter delves into the issue of data warehouse performance. It reviews three types of data structures, namely, indexes, materialized views, and partitioned tables, which are essential for efficiently answering analytical queries. The chapter also shows how these techniques are applied when executing star queries in three commercial data warehouse systems.

The third chapter shows how popular user-centric techniques, namely, personalization and recommendation, can be applied to OLAP queries. The chapter characterizes the approaches proposed in this respect in terms of formulation effort, prescriptiveness, proactiveness, expressiveness, and in terms of the data leveraged.

The fourth chapter discusses the issue of supporting analytical queries over Web-based textual content. This problem lies at the intersection of the domains of search engines and BI. The chapter shows three recent enabling technologies in this respect, namely, cloud computing, self-supervised keyword generation, and fact extraction. The chapter ends by describing the GoOLAP system, a platform supporting Web-scale business analytics.

The fifth chapter provides an overview of Business Intelligence 2.0, an extension of traditional BI based on the evolution of the Web and emerging technologies such as cloud computing. BI 2.0 promises to enable better decision making by complementing traditional organizational data with information present in the Web, like opinions or information about competitors,

while using collective intelligence, collaborative work through social networks, and supporting BI systems with cloud computing.

The sixth chapter surveys the issue of graph mining over social networks. In this respect, community detection may bring very valuable information about the structure of an existing social network. The chapter defines what is a community in a social network context, and then surveys the most popular techniques to detect such communities.

The seventh chapter presents an overview of the Semantic Web, with a special focus on semantic databases, or triplestores, which are specific databases aimed at integrating, storing, and querying the vast amounts of data generated on the Semantic Web every day. The chapter also provides an overview of BI-related scenarios where semantic technologies and triplestores provide valuable advantage and differentiation.

The eight chapter introduces the service paradigm and analyzes its impacts on BI. Specific techniques to engineering service systems are presented, including cloud computing, service-oriented architectures (SOA), and business process modeling (BPM). The chapter also analyzes whether it is possible to consider BI as a service, and analyzes how to use BI techniques to enhance services.

Finally, the ninth chapter explores collaborative BI, which aims at extending the decision-making process beyond the company boundaries. Due to its inherent distribute nature, collaborative BI requires innovative approaches and architectures. The chapter surveys data warehouse integration as an enabling technique for collaborative BI and outlines a new peer-to-peer framework, called Business Intelligence Network, aiming at sharing business information for the decision-making process.

The lectures of the summer school surveyed established areas in the BI domain. They also also pointed out aspects that are beginning to be explored or are still waiting for a solution. We hope that the school's material will inspire further exciting research in these areas. We are grateful to all the lecturers and their co-authors for their excellent contributions, the participants of the summer school for their enthusiasm, and the external referees for their careful work that helped us to improve the lectures in this volume.

October 2011

Marie-Aude Aufaure
Esteban Zimányi

Organization

The First European Business Intelligence Summer School (eBISS 2011) was organized by the MAS Laboratory of the Ecole Centrale de Paris and the Department of Computer and Decision Engineering (CoDE) of the Université Libre de Bruxelles.

Program Committee

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

Local Organizers

Marie-Aude Aufaure	Ecole Centrale de Paris, France
Nesrine Ben Mustapha	Ecole Centrale de Paris, France
Etienne Cuvelier (Organization Chair)	Ecole Centrale de Paris, France
Micheline Elias	Ecole Centrale de Paris, France
Nicolas Kuchmann-Beauger	Ecole Centrale de Paris, France
Cassio Melo	Ecole Centrale de Paris, France
Nizar Messai	Ecole Centrale de Paris, France
Vincent Mousseau	Ecole Centrale de Paris, France
Rania Soussi	Ecole Centrale de Paris, France
Esteban Zimányi	Université Libre de Bruxelles, Belgium

External Referees

Alfredo Cuzzocrea	ICAR-CNR and University of Calabria, Italy
Cécile Favre	Université Lyon 2, France
Matteo Golfarelli	University of Bologna, Italy
Adriana Marotta	Universidad de la República, Uruguay
Anisoara Nica	SQL Anywhere, Canada
Carlos Ordonez	University of Houston, USA
Raul Ruggia	Universidad de la Republica, Uruguay
Jens Otto Sørensen	Energinet.dk, Denmark
Christian Thomsen	Aalborg University, Denmark
Stijn Vansummeren	Université Libre de Bruxelles, Belgium

Contents

1	Data Warehouses: Next Challenges	1
	<i>Alejandro Vaisman, Esteban Zimányi</i>	
2	Data Warehouse Performance: Selected Techniques and Data Structures	27
	<i>Robert Wrembel</i>	
3	OLAP Query Personalisation and Recommendation: An Introduction	63
	<i>Patrick Marcel</i>	
4	The GoOLAP Fact Retrieval Framework	84
	<i>Alexander Löser, Sebastian Arnold, Tillmann Fiehn</i>	
5	Business Intelligence 2.0: A General Overview	98
	<i>Juan Trujillo, Alejandro Maté</i>	
6	Graph Mining and Communities Detection	117
	<i>Etienne Cuvelier, Marie-Aude Aufaure</i>	
7	Semantic Technologies and Triplestores for Business Intelligence	139
	<i>Marin Dimitrov</i>	
8	Service-Oriented Business Intelligence	156
	<i>Alberto Abelló, Oscar Romero</i>	
9	Collaborative Business Intelligence	186
	<i>Stefano Rizzi</i>	
	Author Index	207