

Perspectives on Business Intelligence

Synthesis Lectures on Data Management

Editor

M. Tamer Özsu, *University of Waterloo*

Synthesis Lectures on Data Management is edited by Tamer Özsu of the University of Waterloo. The series will publish 50- to 125 page publications on topics pertaining to data management. The scope will largely follow the purview of premier information and computer science conferences, such as ACM SIGMOD, VLDB, ICDE, PODS, ICDT, and ACM KDD. Potential topics include, but not are limited to: query languages, database system architectures, transaction management, data warehousing, XML and databases, data stream systems, wide scale data distribution, multimedia data management, data mining, and related subjects.

Perspectives on Business Intelligence

Raymond T. Ng, Patricia C. Arocena, Denilson Barbosa, Giuseppe Carenini, Luiz Gomes, Jr., Stephan Jou, Rock Anthony Leung, Evangelos Milios, Renée J. Miller, John Mylopoulos, Rachel A. Pottinger, Frank Tompa, and Eric Yu

2013

Semantics Empowered Web 3.0: Managing Enterprise, Social, Sensor, and Cloud-based Data and Services for Advanced Applications

Amit Sheth and Krishnaprasad Thirunarayan

2012

Data Management in the Cloud: Challenges and Opportunities

Divyakant Agrawal, Sudipto Das, and Amr El Abbadi

2012

Query Processing over Uncertain Databases

Lei Chen and Xiang Lian

2012

Foundations of Data Quality Management

Wenfei Fan and Floris Geerts

2012

Incomplete Data and Data Dependencies in Relational Databases
Sergio Greco, Cristian Molinaro, and Francesca Spezzano
2012

Business Processes: A Database Perspective
Daniel Deutch and Tova Milo
2012

Data Protection from Insider Threats
Elisa Bertino
2012

Deep Web Query Interface Understanding and Integration
Eduard C. Dragut, Weiyi Meng, and Clement T. Yu
2012

P2P Techniques for Decentralized Applications
Esther Pacitti, Reza Akbarinia, and Manal El-Dick
2012

Query Answer Authentication
HweeHwa Pang and Kian-Lee Tan
2012

Declarative Networking
Boon Thau Loo and Wenchao Zhou
2012

Full-Text (Substring) Indexes in External Memory
Marina Barsky, Ulrike Stege, and Alex Thomo
2011

Spatial Data Management
Nikos Mamoulis
2011

Database Repairing and Consistent Query Answering
Leopoldo Bertossi
2011

Managing Event Information: Modeling, Retrieval, and Applications
Amarnath Gupta and Ramesh Jain
2011

Fundamentals of Physical Design and Query Compilation
David Toman and Grant Weddell

2011

Methods for Mining and Summarizing Text Conversations
Giuseppe Carenini, Gabriel Murray, and Raymond Ng

2011

Probabilistic Databases

Dan Suciu, Dan Olteanu, Christopher Ré, and Christoph Koch

2011

Peer-to-Peer Data Management

Karl Aberer

2011

Probabilistic Ranking Techniques in Relational Databases

Ihab F. Ilyas and Mohamed A. Soliman

2011

Uncertain Schema Matching

Avigdor Gal

2011

Fundamentals of Object Databases: Object-Oriented and Object-Relational Design

Suzanne W. Dietrich and Susan D. Urban

2010

Advanced Metasearch Engine Technology

Weiyi Meng and Clement T. Yu

2010

Web Page Recommendation Models: Theory and Algorithms

Sule Gündüz-Ögüdücü

2010

Multidimensional Databases and Data Warehousing

Christian S. Jensen, Torben Bach Pedersen, and Christian Thomsen

2010

Database Replication

Bettina Kemme, Ricardo Jimenez-Peris, and Marta Patino-Martinez

2010

Relational and XML Data Exchange

Marcelo Arenas, Pablo Barcelo, Leonid Libkin, and Filip Murlak

2010

User-Centered Data Management

Tiziana Catarci, Alan Dix, Stephen Kimani, and Giuseppe Santucci

2010

Data Stream Management

Lukasz Golab and M. Tamer Özsu

2010

Access Control in Data Management Systems

Elena Ferrari

2010

An Introduction to Duplicate Detection

Felix Naumann and Melanie Herschel

2010

Privacy-Preserving Data Publishing: An Overview

Raymond Chi-Wing Wong and Ada Wai-Chee Fu

2010

Keyword Search in Databases

Jeffrey Xu Yu, Lu Qin, and Lijun Chang

2009

© Springer Nature Switzerland AG 2022
Reprint of original edition © Morgan & Claypool 2013

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means—electronic, mechanical, photocopy, recording, or any other except for brief quotations in printed reviews, without the prior permission of the publisher.

Perspectives on Business Intelligence

Raymond T. Ng, Patricia C. Arocena, Denilson Barbosa, Giuseppe Carenini, Luiz Gomes, Jr., Stephan Jou, Rock Anthony Leung, Evangelos Milios, Renée J. Miller, John Mylopoulos, Rachel A. Pottinger, Frank Tompa, and Eric Yu

ISBN: 978-3-031-00720-0 paperback
ISBN: 978-3-031-01848-0 ebook

DOI 10.1007/978-3-031-01848-0

A Publication in the Springer series
SYNTHESIS LECTURES ON DATA MANAGEMENT

Lecture #32
Series Editor: M. Tamer Özsu, *University of Waterloo*
Series ISSN
Synthesis Lectures on Data Management
Print 2153-5418 Electronic 2153-5426

Perspectives on Business Intelligence

Raymond T. Ng, Patricia C. Arocena, Denilson Barbosa, Giuseppe Carenini, Luiz Gomes, Jr., Stephan Jou, Rock Anthony Leung, Evangelos Milios, Renée J. Miller, John Mylopoulos, Rachel A. Pottinger, Frank Tompa, and Eric Yu

SYNTHESIS LECTURES ON DATA MANAGEMENT #32

ABSTRACT

In the 1980s, traditional Business Intelligence (BI) systems focused on the delivery of reports that describe the state of business activities in the past, such as for questions like “How did our sales perform during the last quarter?” A decade later, there was a shift to more interactive content that presented how the business was performing at the present time, answering questions like “How are we doing right now?” Today the focus of BI users are looking into the future. “Given what I did before and how I am currently doing this quarter, how will I do next quarter?”

Furthermore, fuelled by the demands of Big Data, BI systems are going through a time of incredible change. Predictive analytics, high volume data, unstructured data, social data, mobile, consumable analytics, and data visualization are all examples of demands and capabilities that have become critical within just the past few years, and are growing at an unprecedented pace.

This book introduces research problems and solutions on various aspects central to next-generation BI systems. It begins with a chapter on an industry perspective on how BI has evolved, and discusses how game-changing trends have drastically reshaped the landscape of BI. One of the game changers is the shift toward the consumerization of BI tools. As a result, for BI tools to be successfully used by business users (rather than IT departments), the tools need a business model, rather than a data model. One chapter of the book surveys four different types of business modeling. However, even with the existence of a business model for users to express queries, the data that can meet the needs are still captured within a data model. The next chapter on vivification addresses the problem of closing the gap, which is often significant, between the business and the data models. Moreover, Big Data forces BI systems to integrate and consolidate multiple, and often wildly different, data sources. One chapter gives an overview of several integration architectures for dealing with the challenges that need to be overcome.

While the book so far focuses on the usual structured relational data, the remaining chapters turn to unstructured data, an ever-increasing and important component of Big Data. One chapter on information extraction describes methods for dealing with the extraction of relations from free text and the web. Finally, BI users need tools to visualize and interpret new and complex types of information in a way that is compelling, intuitive, but accurate. The last chapter gives an overview of information visualization for decision support and text.

KEYWORDS

business intelligence, big data, business modeling, vivification, data integration, information extraction, information visualization

Contents

1	Introduction and the Changing Landscape of Business Intelligence	1
	<i>Stephan Jou and Raymond Ng</i>	
1.1	Introduction	1
1.2	The Role of Research and This Book	3
2	BI Game Changers: an Industry Viewpoint	7
	<i>Rock Leung, Chahab Nastar, Frederic Vanborre, Christophe Favart, Gregor Hackenbroich, Philip Taylor, and David Trastour</i>	
2.1	Introduction	7
2.2	Defining Business Intelligence	8
2.3	Early Days of BI	9
2.4	Classic BI	10
2.5	Game-changing Trends	10
2.5.1	Faster Business	11
2.5.2	Bigger Data	13
2.5.3	Better Software	14
2.6	Next-generation BI	16
2.7	Conclusions	17
3	Business Modeling for BI	19
	<i>Eric Yu, Jennifer Horkoff, John Mylopoulos, Gregory Richards, and Daniel Amyot</i>	
3.1	Introduction	19
3.2	Modeling Business Processes	20
3.3	Strategic Business Modeling for Performance Management	22
3.4	Modeling Business Models	24
3.5	Toward Modeling for BI	28
3.5.1	BIM Concepts	28
3.5.2	Reasoning with BIM Models	30
3.6	Conclusions	32

4	Vivification in BI	33
	<i>Patricia C. Arocena, Renée J. Miller, and John Mylopoulos</i>	
4.1	Introduction	33
4.2	A Motivating Example	34
4.3	The Vivification Problem	37
4.3.1	Knowledge Base Vivification	37
4.3.2	Data Exchange	39
4.4	Formal Framework	40
4.5	Current Vivification Strategies	43
4.5.1	Strategies for Dealing with Incompleteness	43
4.5.2	Strategies for Dealing with Uncertainty	44
4.5.3	Summary of Other Relevant Work	45
4.6	Toward Adaptive Vivification Strategies	46
4.6.1	Vivification by Acceptance	46
4.6.2	Vivification by Default	47
4.6.3	Vivification by Resolution	48
4.7	Directions for Future Research	49
4.8	Conclusions	51
5	Information Integration in BI	53
	<i>Rachel A. Pottinger</i>	
5.1	Introduction	53
5.2	Information Integration Goals and Axes	54
5.3	Challenges and Background	56
5.3.1	Schemas and Semantic Heterogeneity	56
5.3.2	Ontologies	57
5.4	Overview of Different Information Integration Architectures	57
5.4.1	Data Integration	57
5.4.2	Data Warehousing	61
5.4.3	Peer Data Management Systems	64
5.5	Information Integration Tools in Industry	65
5.6	Conclusions	66
6	Information Extraction for BI	67
	<i>Denilson Barbosa, Luiz Gomes, Jr., and Frank Wm. Tompa</i>	
6.1	Introduction	67

6.1.1	Levels of Structuredness	68
6.1.2	The Role of IE for BI	69
6.2	IE From Text	71
6.2.1	Patterns in Language	72
6.2.2	Named Entity Recognition	74
6.2.3	Ontology Learning	79
6.2.4	Relation Extraction	80
6.2.5	Factoid Extraction	84
6.3	Data Extraction from the web	85
6.3.1	Wrapper Induction	85
6.3.2	Schema Extraction	86
6.4	BI over Raw Text	88
6.5	Conclusions	89
7	Information Visualization for BI	93
	<i>Giuseppe Carenini and Evangelos Milios</i>	
7.1	Introduction	93
7.2	Information Visualization for Decision Support	93
7.2.1	Information Visualization in the Performance Management Cycle: Information Dashboards	94
7.2.2	Visualization for Preferential Choice	99
7.2.3	Current and Future Trends in Information Visualization for Decision Support	103
7.3	Visualizing Text	105
7.3.1	Text Clouds	105
7.3.2	Topic Models	108
7.3.3	Text Streams	109
7.3.4	Sentiment Analytics	113
7.3.5	Multiview Systems for Document Collections	117
7.4	Conclusions	123
	Bibliography	125
	Authors' Biographies	145