Introduction to the special issue of Business Intelligence and the Web

Over the last decade we have been witnessing an increasing use of Business Intelligence (BI) solutions, which allow business people to query, understand, and analyze their business data in order to make better decisions. Traditionally, BI applications allowed business people to acquire useful knowledge from the data of their organization by means of a variety of technologies, such as data warehousing, data mining, business performance management, OLAP, periodical business reports, and the like. Yet, in the very recent years, a new trend emerged: BI applications no longer limit their analysis to the data inside one company. Increasingly, they also source their data from the outside, i.e., from the Web, and complement company-internal data with value-adding information from the Web (e.g., retail prices of products sold by competitors), in order to provide richer insights into the dynamics of today's business. In parallel to the move of data from the Web into BI applications, we are now assisting to the move of BI applications from company-internal information systems to the Web: BI as a service (e.g., hosted BI platforms for small- and medium-size companies) is the target of huge investments and the focus of large research efforts by industry. The idea is that of outsourcing the processing and analysis of large bodies of data and consuming BI from the cloud.

The International Workshop on Business intelligencE and the WEB (BEWEB), held in conjunction with the 13th International Conference on Extending Database Technology (EDBT 2010), intended to be a forum for exchanging ideas on (i) how to leverage the huge amount of data that is available on the Web in BI applications, (ii) how to apply Web engineering methods and techniques to the design of BI applications, and (iii) how to use BI knowledge in the design of Web applications.

Authors of the best papers were invited to extend their papers and resubmit them for this special issue. These extended papers were strongly reviewed by paying special attention on the new material and quality of the research. In the following, we summarize these selected papers.

The first paper, "Building Data Warehouses with Semantic Web Data", by Victoria Nebot and Rafael Berlanga, presents a semi-automatic approach for identifying and extracting multidimensional elements of a data warehouse by analyzing semantic data. Authors use well-known standards, ontology inferences, as well as a case study on the biomedical domain for evaluation purposes, thus showing the scalability and applicability of their approach.

The second paper, "A Platform for Situational Awareness in Operational BI", by Malu Castellanos, Chetan Gupta, Song Wang, Umeshwar Dayal and Miguel Durazo, relies on the fact that companies must extract more value when dealing with unstructured data. In this paper, authors argue that it is highly important to correlate streaming unstructured Web data with internal document data in near real time in order to be aware of external events that can affect business operations of companies. To this aim, authors describe SIE-OBI, a platform being developed at HP Labs, and illustrate its use via a case study.

The third paper, entitled "A Personalization Process for Spatial Data Warehouses Development", by Octavio Glorio, Jose-Norberto Mazón, Irene Garrigós, and Juan Trujillo, argues that spatial needs of each decision maker could change over time or depending on the context, thus requiring the multidimensional schema of a data warehouse to be continuously updated with changes related to spatiality. This may hamper decision making process, since analyst may get frustrated when spatial features are required. To solve this problem, authors present a Web-based approach for considering spatiality as a personalization feature within a formal design process. In this way, each decision maker will be able to access its own personalized multidimensional schema with its required spatial structures and instances, suitable to be properly analyzed at a glance.

The fourth paper is "A Practical Application of our MDD approach for Modeling Secure XML Data Warehouses", by Belén Vela, Carlos Blanco, Eduardo Fernández-Medina, and Esperanza Marcos. In this paper, the authors claim that the extensive use of XML in the implementation of data warehouses is motivated by the fact that data warehouses integrate heterogeneous sources which are not only limited to their internal business data but also include data from the Web. Since business information that data warehouses manage is crucial and highly sensitive, then data must be carefully protected, and security becomes a key issue in the design of data warehouses, regardless of the implementation technology used. Therefore, in this paper authors propose a methodological approach for the model driven development of secure XML data warehouses.

We would like to thank all the authors who revised and extended their papers for this special issue and the reviewers for their hard work in reviewing these extended papers and providing critical and useful comments which helped authors in improving their papers. Absolutely, all of them have contributed in having this special issue of a high quality. We hope the readers will enjoy reading this issue and find the content beneficial to their research.

Finally, we would like to recall that BEWEB 2010 and this special issue were dedicated to the memory of Dr. Juan Manuel Pérez who sadly passed away on December 18, 2009. Juan Manuel made contributions in the area of information retrieval, multidimensional databases, and web-based technologies. He not only was a valuable colleague but also a charming friend.

Jose-Norberto Mazón obtained his Ph.D. in Computer Science from the University of Alicante (Spain). He currently enjoys a research grant from the Spanish Ministry of Education and Science. He has published several papers about data warehouses in national and international workshops and conferences, (such as DAWAK, ER, DOLAP, BNCOD, JISBD and so on) and in several journals such as Decision Support Systems (DSS) or Data and Knowledge Engineering (DKE). He is co-organizer of the international workshops WeRE and BEWEB. His research interests are: business intelligence, design of data warehouses, multidimensional databases, and model driven development. Email: jnmazon@dlsi.ua.es

Irene Garrigós is an assistant professor and post-doc researcher at the University of Alicante, (Spain), from which she holds a Ph.D. and a master's degree in Computer Science. She has published several papers in national and international workshops and conferences, such as ICWE, ER, WISE, APWEB, JISBD and so on. Dr. Garrigós has served as a Program Committee member of several workshops and conferences such as ER, WISM, MDA, FPUML, UWA and has served as assistant referee in several international conferences such as WWW and ICWE. She has done research stays in Belgium (Vrije Universiteit Brussel) and the Netherlands (Technische Universiteit Eindhoven). Irene is co-organizer of the international workshops WeRE and BEWEB. Her research interests are: Web engineering, personalization, model driven development, Web and business intelligence, adaptive systems. Email: igarrigos@dlsi.ua.es

Florian Daniel is a postdoctoral researcher at the Department of Information Engineering and Computer Science of the University of Trento, Italy. He has a Ph.D. in Information Technology from Politecnico di Milano, Italy. His main research interests are mash-ups and web/services engineering, and compliance, quality, and privacy in business intelligence applications. He is co-author of the book "Engineering Web Applications" (Springer, 2009) and has published more than 50 scientific papers in international conferences and journals. Florian is co-organizer of the international workshops ComposableWeb and BEWEB and involved in the organization of international conferences like BPM, ICSOC, and ICWE. Email: daniel@disi.unitn.it

Juan Trujillo is an associated professor at the Computer Science School at the University of Alicante, Spain. Trujillo received a Ph.D. in Computer Science from the University of Alicante (Spain) in 2001. His research interests include database modeling, data warehouses, conceptual design of data warehouses, multidimensional databases, data warehouse security and quality, mining data warehouses, OLAP, as well as object-oriented analysis and design with UML. He has published many papers in high quality international conferences such as ER, UML, ADBIS, CAiSE, WAIM or DAWAK. He has also published papers in highly cited international journals such as IEEE Computer, Decision Support Systems (DSS), Data and Knowledge Engineering (DKE) or Information Systems (IS). Dr. Trujillo has served as a Program Committee member of several workshops and conferences such as ER, DOLAP, DAWAK, DSS, JISBD and SCI and has also spent some time as a reviewer of several journals such as JDM, KAIS, ISOFT and JODS. He has been Program Chair of DOLAP' 05 and BP-UML' 05, and Program Co-chair of DAWAK' 05, DAWAK' 06, BP-UML' 06, FP-UML' 07, FP-UML' 08, and FP-UML' 09. Email: jtrujillo@dlsi.ua.es

Jose-Norberto Mazón
Lucentia, DLSI, University of Alicante, Spain
Corresponding author.
E-mail address: jnmazon@dlsi.ua.es

Irene Garrigós
Lucentia, DLSI, University of Alicante, Spain
E-mail address: igarrigos@dlsi.ua.es

Florian Daniel
dbTrento, DISI, University of Trento, Italy
E-mail address: daniel@disi.unitn.it

Juan Trujillo
Lucentia, DLSI, University of Alicante, Spain
E-mail address: jtrujillo@dlsi.ua.es