Preface

The rapid growth of organizational and business process managed via information systems made available a big variety of data, which as a consequence created a high demand for making available data analysis techniques more effective and valuable. The third edition of the International Symposium on Data-driven Process Discovery and Analysis (SIMPDA 2013) was conceived to offer a forum where researchers from different communities and the industry can share their insights in this hot new field. The symposium features a number of advanced keynotes illustrating new approaches, presentations on recent research, a competitive PhD seminar, and selected research and industrial demonstrations. The goal is fostering exchanges among academic researchers, industry, and a wider audience interested on process discovery and analysis. The event is organized jointly by the IFIP WG 2.6 and W.G 2.12. In 2015 the symposium was inserted among the VLDB 2013 workshops, in the charming setting of Riva del Garda at the northwestern corner of Lake Garda, at the southern edge of the Italian Alps, near the Dolomites.

Submissions cover theoretical issues related to process representation, discovery, and analysis or provide practical and operational experiences in process discovery and analysis. To improve the quality of the contributions the symposium fostered the discussion during the presentation, letting authors the opportunity to improve their work extending the presented results. For this reason, authors of accepted papers and keynote speakers were invited to submit extended articles to this post-symposium volume of LNBIP. There were 18 submissions and 6 papers were accepted for publication.

In the first paper "The Effect of Noise on Mined Declarative Constraints," Claudio Di Ciccio et al. present an experimental analysis on the robustness to noise of the different types of Declare constraints, establishing a constraint hierarchy based on formal relationships between the different types.

In the second paper by Gregor Grambow et al. "Towards Collecting Sustainability Data in Supply Chains with Flexible Data Collection Processes," a dedicated information system for supporting data collection processes is developed. In combination with potentially long-running processes, these issues result in great process variability.

The third paper by Helen Balinsky et al., "Handling Environment for Publicly Posted Composite Documents," focuses on Publicly Posted Composite Documents (PPCD) – a multipart document format retaining parts in their original formatting for usability, user convenience, and information availability, while providing document security and built-in access control for inter- and intra-organizational document workflows distributed over low-security channels. In particular, this paper illustrates a methodology to use OLE Automation, Add-Ons, and System Call Interception technique to preserve the overall document composition, while the handling of individual parts is delegated to external tools, native for individual document formats.

The fourth paper by Roberto Espinosa et al., "Enabling Non-expert Users to Apply Data Mining for Bridging the Big Data Divide," offers to nonexpert miners a tool that just by uploading their data sets, returns them the more accurate mining pattern without dealing with algorithms or settings, thanks to the use of a data mining algorithm recommender. The authors also incorporate a previous task to help nonexpert users to specify data mining requirements and a later task in which users are guided in interpreting data mining results. Furthermore, they experimentally test the feasibility of our approach, in particular, the method to build recommenders in an educational context, where instructors of e-learning courses are nonexpert data miners who need to discover how their courses are used in order to make informed decisions to improve them.

The fifth paper by Antonia Azzini et al., "Combining Semantic Lifting and Ad Hoc Contextual Analysis in a Data Loss Scenario," investigates the role of semantic lifting in discovering behavioral patterns of the executed process. Conclusions spotlight the impact that semantics lifting has on support and confidence of the inferred probabilities of observing these behavioral patterns.

The sixth paper by Wil M.P. van der Aalst et al., "Comparative Process Mining in Education: An Approach Based on Process Cubes," starts by observing that existing process mining techniques focus on the analysis of a single process rather than the comparison of different processes. This paper proposes comparative process mining using process cubes. An event has attributes referring to the dimensions of the process cube. Through slicing, dicing, rolling-up, and drilling-down we can view event data from different angles and produce process mining results that can be compared.

November 2014   Paolo Ceravolo
                                                                          Rafael Accorsi
                                                                  Philippe Cudre-Mauroux

Springer