Message from the Chairs

We are delighted to present to you, on behalf of the entire conference organizing committee and the workshop organizers, the proceedings of the Workshops of the EDBT/ICDT 2016 Joint Conference, held on March 15, 2016, in Bordeaux, France.

The International Conference on Extending Database Technology (EDBT) and the International Conference on Database Theory (ICDT) are two prestigious forums for the exchange of the latest research results in data management and the theoretical foundations of database systems. While having the same overarching goal of presenting cutting-edge results, ideas, techniques, and theoretical advances in databases, the workshops of the EDBT/ICDT joint conference are separately tasked by focusing on emerging topics that complement the areas covered by the main technical program.

This year, the workshops cover the following topics:

- **Multi-engine and workflow management for big data analytics**. Big data analytics has become a necessity in the majority of industries, enabling engineers, domain experts and scientists alike to tap the potential of vast amounts of data that are critical for business and science. The success and effectiveness of such analysis depend on numerous challenges related to the data itself, the nature of the analytics tasks, as well as the computing environment over which the analysis is performed. These issues have given rise to many diverse programming models, execution engines and data stores to enable large-scale data management. To harvest the benefits of this plethora of data and compute engines as well as programming models, libraries and tools available, we need coordinated, adaptive and integrative efforts. These efforts include the definition of versatile programming models, engine performance modeling and monitoring, extended planning and optimization algorithms, deployment/execution on multiple engines, as well as workflow management and visualization techniques, for complex analytics queries over large, heterogeneous, irregular or unstructured data over diverse compute environments. (Workshop name: *MEDAL: 1st International Workshop on Multi-Engine Data AnaLytics*.)
- **Interplay between big data processing and stream processing solutions**. Big data processing has gained much attention in the recent years raising the need for stream processing solutions. At the same time, there has been substantial research and development of methods and technologies that have high potential for further progress in stream processing. This workshop discusses the latest developments in this area and examines how they relate to each other. The workshop is especially interested in the following topics: exploring how those new developments can be exploited and combined; what the useful development approaches for implementing and adapting complex and efficient processing pipelines on stream processing infrastructures are; how stream processing approaches and algorithms can be applied in specific domains; how evidences gained from different types of data streams (e.g., Twitter stream vs. financial data streams) can be combined for improved and consistent analysis results; how the processing tasks can be seamlessly offloaded to co-processors, including FPGAs, GPUs, etc., and which algorithms qualify for this. (Workshop name: *Big Data Processing – Reloaded*.)
- **Management and query answering for graph data and linked web data**. The growing scale and importance of graph data in several database application areas has recently driven much research efforts towards the development of data models and technologies for graph-data management. Life science databases, social networks, Semantic Web data, bibliographical networks, knowledge bases and ontologies, are prominent examples of application domains exhibiting data that is natural to represent in graph-based form. Datasets in these domains are often characterized by heterogeneity, complexity and largeness of contents that make the querying experience a really challenging task. Some of the issues that will be addressed are the following: the efficient and effective support of graph queries in different application domains; data management approaches related to the Linked Data, how these may evolve to address the big data issues, and the relationships of these approaches to other Semantic Web technologies. (Workshop names: *GraphQ: 5th international workshop on Querying Graph Structured Data*, and *LWDM: 6th International Workshop on Linked Web Data Management*.)

- **Management and mining for energy and smart city data**. In today's urbanizing world, cities have not only physical infrastructures, such as road networks, utilities, or buildings, but also comprise a knowledge infrastructure ranging from low-level sensor networks to public databases and social media streams. The data emerging from all those sources is a very precious resource to make cities more intelligent, innovative and integrated beyond the boundaries of isolated applications, thus enabling us to turn the city data into insightful information. This will require advancements in a broad range of topics rooted in different scientific fields, in order to enable novel research to mine important patterns from city data, and to apply them in various emerging application areas, such as smart mobility/transportation, smart tourism or smart participation. The special topic of energy data management will be addressed in its own right, aiming at sparking the discussion within the academic database community and bridging the gap between practitioners coming from the energy domain and database experts. (Workshop names: *DAMASCA: DAta Mining And Smart Cities Applications*, and *EnDM: 5th workshop on Energy Data Management*.)
- **Preservation of big data**. Recently, there has been a vast and rapidly increasing quantity of scientific, corporate, government and crowd-sourced knowledge bases published and curated independently on the Data Web in various data formats. Traditional closed-world settings impose that data management and evolution are performed within a well-defined controlled environment where change operations and dependencies on data can be easily monitored and handled. On the other hand, web and digital preservation techniques assume that preservation subjects, such as web pages, are plain digital assets that are collected (usually via a crawling mechanism) and individually archived for future reference. In contrast with these two approaches, the Data Web poses new requirements for revisiting and adjusting traditional closed-world data management techniques such as temporal management and change detection, data archiving and preservation, data ingestion, integration and enrichment, data provenance and quality, data visualization and exploratory analysis to the characteristics of multi-curated knowledge bases. (Workshop name: *DIACHRON: 2nd International Workshop on Preservation of Evolving Big Data*.)
- **Privacy and anonymity in a digital world**. Organizations collect vast amounts of information on individuals, and at the same time they have access to ever-increasing levels of computational power. Although this conjunction of information and power provides great benefits to society, it also threatens individual privacy. As a result legislators for many countries try to regulate the use and the disclosure of confidential information. While privacy is a topic discussed everywhere, data anonymity recently established itself as an emerging area of computer science. Its goal is to produce useful computational solutions for releasing data, while providing scientific guarantees that the identities and other sensitive information of the individuals who are the subjects of the data are protected. This workshop brings together researchers and practitioners from computer science and other fields that are interacting with computer science in the privacy area, such as statistics, healthcare informatics, and law, in order to discuss and present current research challenges and advances in data privacy and anonymity research. (Workshop name: *PAIS: 9th International Workshop on Privacy and Anonymity in the Information Society*.)

We thank the workshop chairs for their effort in organizing these workshops, and for putting together an exciting program. We also thank the authors for continuing to submit their high-quality work to the EDBT/ICDT workshops, and the PC members and external reviewers for their contributions that made these workshops possible. Finally, we would like to acknowledge the help of the EDBT Proceedings Chair, Dr. Kostas Stefanidis, who edited the proceedings of the workshops, and the conference organizers and volunteers for the realization of this event.

Sincerely,

Themis Palpanas, *Workshops Chair*
Evaggelia Pitoura and Wim Martens, *EDBT and ICDT Program Chairs*
Sofian Maabout, *General Chair*