Editorial: Big Data Exploration, Visualization and Analytics *

Nikos Bikakis[1] George Papastefanatos[2]
Olga Papaemmanouil[3]

[1]University of Ioannina, Greece
[2]ATHENA Research Center, Greece
[3]Brandeis University, USA

The Big Data era has realized the availability of a great amount and variety of datasets for analysis by non-corporate data analysts, such as research scientists, data journalists, policy makers, SMEs and individuals. They are characterized by high volumes which make traditional database and system infrastructure incapable of efficiently storing and processing them; they are accessible in very different, usually raw formats (e.g., plain text, json, rdf); they are generated or modified in high rates, and they exhibit different levels of data quality and schema representations. The level of difficulty in transforming a data-curious user into someone who can access and analyze that data is even more burdensome now for a great number of users with little or no support and expertise on the data processing part. The goal of visual data exploration and analysis is to facilitate information perception and manipulation, knowledge extraction and inference by non- expert users. The visualization techniques, used in a variety of modern systems, provide users with intuitive means to interactively explore the content of the data, identify interesting patterns, infer correlations and causalities, and supports sense-making activities that are not always possible with traditional data traditional data analysis techniques.

Several challenges arise in the field of information visualization and data management, due to the new Big Data characteristics. First, the modern exploration and visualization systems should offer scalable data management techniques in order to efficiently handle billion objects datasets, limiting the system response in a few milliseconds. Nowadays systems must, also, address the challenge of on-the-fly scalable visualizations over

* Special Issue on "*Big Data Exploration, Visualization and Analytics*", Big Data Research, Elsevier, 2019

large and dynamic sets of volatile raw data, offering efficient interactive exploration techniques, as well as mechanisms for information abstraction, sampling and summarization for addressing problems related to visual information overplotting. Further, they must encourage user comprehension offering customization capabilities to different user-defined exploration scenarios and preferences according to the analysis needs. Overall, the challenge is to enable users to gain value and insights out of the data as rapidly as possible, minimizing the role of IT-experts in the loop.

This special issue aimed to publish work on multidisciplinary research areas spanning from Data Management and Mining to Information Visualization and Human-Computer Interaction. In addition to the normal submissions, this special issue considered to invite some of the best papers from the 1st International Workshop on Big Data Visual Exploration and Analytics[1], held in conjunction with the 21th Intl. Conference on Extending Database Technology & 21th Intl. Conference on Database Theory (EDBT/ICDT 2018). After a rigorous peer review process we accepted four papers out of nine.

The paper "*HiePaCo: Scalable Hierarchical Exploration in Abstract Parallel Coordinates Under Budget Constraints*" by Gaëlle Richer, Joris Sansen, Frédéric Lalanne, David Auber, and Romain Bourqui, proposes an aggregation-based multiscale visual representation, in order to visualize and interact with large multidimensional datasets using parallel coordinates. Further, a client-server architecture is adopted, allowing the system to scale to billion-items datasets using a cluster of 15 computers.

In the paper "*Joint Contour Net Analysis for Feature Detection in Lattice Quantum Chromodynamics Data*", Dean P. Thomas, Rita Borgo, Robert S. Laramee and Simon J. Hands, study the benefits of using multivariate topological visual representations to analyze Lattice Quantum Chromodynamics data. Particularly, the study uses the Joint Contour Net representation, that had been proposed to simultaneously quantize the variation of multiple variables. The paper demonstrates that multivariate topological representations can assist scientists to locate and track important data characteristics in a temporal setting.

In "*Visual Exploration of Geolocated Time Series with Hybrid Indexing*", Georgios Chatzigeorgakidis, Kostas Patroumpas, Dimitrios Skoutas, Spiros Athanasiou, and Spiros Skiadopoulos, introduce spatial-time series indexes which are used to offer scalable multilevel map-based visualizations of geolocated time series. The scalability of the proposed methods is demonstrated using two real-world and one synthetic dataset.

Finally, the paper "*Interactive Visual Analytics for Sensemaking with Big Text*" by Michelle Dowling, Nathan Wycoff, Brian Mayer, John E. Wenskovitch, Scotland Leman, Leanna House, Nicholas F. Polys, Chris North,

[1] https://bigvis.imsi.athenarc.gr/bigvis2018

and Peter Hauck, presents a human-in-the-loop computational model that enables sensemaking tasks over textual data. The model consists of two subprocess: (a) the foraging loop, which is implemented as an interactive ranking of results, integrated with topic modeling; and (b) the synthesis loop, which offers an interactive spatial projection. A system prototype is implemented offering sensemaking tasks over 30K news articles.

In closing, we would like to thank the authors for their high-quality contributions to this special issue and all the reviewers for their generous help and valuable suggestions. We also like to thank the Editors-in-Chief of the Journal of Big Data Research, Themis Palpanas and Zhaohui Wu, for giving us the opportunity to publish this special issue.

Guest Editors

Nikos Bikakis
University of Ioannina, Greece

George Papastefanatos
ATHENA Research Center, Greece

Olga Papaemmanouil
Brandeis University, USA