

## Guest Editorial: Advances in Databases, Information Systems, and Database Engineered Applications – Recent Trends, and Open Issues

Mirjana Ivanović<sup>1</sup>, Richard Chbeir<sup>2</sup>, Johann Gamper<sup>3</sup>, Sergio Ilarri<sup>4</sup>, Yannis Manolopoulos<sup>5</sup>, and Joe Tekli<sup>6</sup>

<sup>1</sup> University of Novi Sad, Serbia, mira@dmi.uns.ac.rs

<sup>2</sup> University of Pau and the Adour Region, Anglet, France, richard.chbeir@univ-pau.fr

<sup>3</sup> Free University of Bozen-Bolzano, Italy, johann.gamper@unibz.it

<sup>4</sup> University of Zaragoza, Spain, sillari@unizar.es

<sup>5</sup> University of York, Europe Campus, Thessaloniki, Greece, imanolopoulos@city.yorkeuropecampus.eu

<sup>6</sup> Lebanese American University, Lebanon, joe.tekli@lau.edu.lb

Nowadays, modern, technically, and technologically rapidly developing society and the progressive advancement of almost all ICT disciplines bring new intelligent and smart interconnected socio-technological systems, architectures, components and services in everyday life and activities. Two long-lasting events with high reputation: the 28th European Conference on Advances in Databases and Information Systems (ADBIS) and the 28th International Database Engineered Applications Symposium (IDEAS), were held in Bayonne, France, during August 28-31, 2024, attracting high-quality researchers and authors to present their emergent and contemporary scientific results and achievements.

ADBIS conference provides an international platform for the presentation of research on database theory, development of advanced DBMS technologies, solutions for data engineering and analytics, and their advanced applications. The main objective of ADBIS series is to provide a forum for the dissemination of research accomplishments and to promote interaction and collaboration among the database, information systems, data science, artificial intelligence and machine learning communities from European countries and the rest of the world. ADBIS'2024 program included 15 full papers carefully reviewed and selected from 43 submissions. The conference focus was on Algebra, Models and Schemata, Discovery and Data Analysis, Algorithms and Optimization, Access Methods and Query Processing, Advanced Architectures, Machine Learning, Large Language Models.

IDEAS symposium addresses engineering and application aspects of data regardless of the database models. It provides an international forum for discussion of the problems of engineering data driven systems involving not only database technology but also related areas of artificial intelligence, communication, information retrieval, human machine interface, multimedia, NLP, privacy, security and others. The goal of IDEAS is to foster closer interaction among the industrial, research and user communities and provide an opportunity for them to meet, discuss ideas, examine the current ones and develop new solutions and research directions. IDEAS'2024 program included 21 full papers carefully reviewed and selected from 34 submissions. The symposium focused on Theoretical Issues, Classification, Text and Languages, Big Data, Query Processing and Applications, Machine Learning and Rules, Recommendation, and Indexing and Event Detection.

This special section of ComSIS Journal contains revised and extended versions of papers presented at the ADBIS and IDEAS 2024 events. We have selected 10 papers from both events and invite authors to submit extended versions to be considered for publishing in ComSIS Journal. Finally, after rigorous reviewing procedure 4 papers have been accepted for publication.

The first paper entitled “Mitigating Out-of-Vocabulary Challenges in Embedded devices Vulnerability Classification: An Ensemble Embedding Approach with Bidirectional Context Modeling” by Aissa Ben Yahya, Hicham El Akhal and Abdelbaki El Belrhiti El Alaoui, contributes to the area of embedded systems and vulnerable cyberattacks. The effectiveness of commonly used pre-trained word embeddings like Word2Vec, GloVe, and FastText is limited by reliance on training corpora that lack domain specific terminology reducing classification performance. To address this limitation, the authors proposed a novel ensemble embedding technique that combines multiple pre-trained embeddings to improve vulnerability classification in embedded systems. Extensive experiments on benchmark datasets demonstrated significant potential for enhancing cybersecurity in critical infrastructure applications.

The second paper entitled “Towards Enhancing Data Science Agents with Semantics” by Sayed Hoseini, Maximilian Ibbels, Maximilian Knoll1, Christoph Quix, contributes to the area of data lakes. Contemporary approaches in this area still lack essential ML-specific features, limiting their effectiveness for end-to-end automation. However Automated Machine Learning (AutoML) and Large Language Models (LLMs) offer potentially good solutions. Based on that, the paper presents an integration of AutoML frameworks and LLMs within a data lake system. More specifically the authors proposed an innovative vision that integrates multi-agent frameworks for data science with knowledge graphs that capture historical experience from previous ML experiments. Initial version of their data lake system SEDAR for developing comprehensive, context-aware ML agents and their integration is presented.

The third paper entitled “Towards Explainable Sequential Learning” by Giacomo Bergami, Emma Packer, Kirsty Scott and Slivia Del Din, offers a hybridly explainable temporal data processing pipeline, DataFul Explainable MultivariatE coRrelatIonal Temporal Artificial intelligence (EMeriTAtE+DF) which enables human explainable results. Extending the event-based literature to design specification mining algorithms supporting concurrent constituents the authors’ solutions outperform state-of-the-art algorithms for multivariate time series classification. The effectiveness of the proposed methodology premiering the extraction of explainable correlations across multivariate time series dimensions with dataful features are evaluated over four datasets.

The last paper entitled “Siren Federate: Bridging Document, Relational, and Graph Models for Exploratory Graph Analysis” by Georgeta Bordea, Stéphane Campinas, Matteo Catena, and Renaud Delbru, focuses on exploratory graph and graph models. The authors proposed Siren Federate architecture that efficiently supports exploratory graph analysis by bridging document-oriented, relational and graph models. Their contributions include distributed join algorithms, adaptive query planning, query plan folding, semantic caching, and semi-join decomposition for path query. Experimental results show that Siren Federate exhibits low latency and scales well with the amount of data, the number of users, and the number of computing nodes.

We gratefully acknowledge all the hard work and enthusiasm of authors, without whom this special section would not have been possible. In addition, we appreciate for the careful study of the selected papers and the constructive comments made by the reviewers:

Costin Badica, University of Craiova, Romania

Goran Banjac, University of Banja Luka, Bosnia and Hercegovina

Diego Ceccarelli, Bloomberg Company, UK

Zoltan Geler, University of Novi Sad, Serbia

Anastas Misev, Ss. Cyril and Methodius University, North Macedonia

Aleksandra Klasnja Milicevic, University of Novi Sad, Serbia

Vladimir Kurbalija, University of Novi Sad, Serbia

Doina Logofatu, Frankfurt University of Applied Sciences, Germany

Goran Sladic, University of Novi Sad, Serbia

Finally, we hope that readers will enjoy the content of this special section, and that it will inspire them to do high quality research in these modern and attractive areas.

