

## Preface

The XXII International Conference “Data Analytics and Management in Data Intensive Domains” (DAMDID/RCDL’2020) takes place in 2020 on October 13–16 at the Voronezh State University.

DAMDID is held as a multidisciplinary forum of researchers and practitioners from various domains of science and research, promoting cooperation and exchange of ideas in the area of data analysis and management in domains driven by data-intensive research. Approaches to data analysis and management being developed in specific data-intensive domains (DID) of X-informatics (such as X = astro, bio, chemo, geo, med, neuro, physics, chemistry, material science, etc.), social sciences, as well as in various branches of informatics, industry, new technologies, finance, and business are expected to contribute to the conference content.

The program of DAMDID/RCDL 2020 is oriented towards data science and data-intensive analytics as well as on data management topics. The program of this year included keynotes and invited talks covering database management issues as well as conceptual modeling and data analysis in data intensive domains.

Introduction to GaussDB A, an MPP analytical database by Huawei is provided by Pavel Velikhov (principal software engineer at Huawei) in his invited talk. The research directions on optimization of GaussDB database being conducted in the Moscow Database Lab are described. This research includes applying advanced machine learning methods to generate better query plans, solving the problems of skewed data, and incorporating in-memory database technologies. The invited talk by Alexey Molodchenkov (senior researcher at the Federal research center Computer science and control of RAS) considers an intelligent system for monitoring and controlling factors (health preservation system) that effect the development of diseases. The health preservation system is based on the principle of 4P medicine: prediction, prevention, personalization, and participation. The knowledge base of the system is based on a heterogeneous semantic network, combining the attributes of risk factors into a single network, and including nodes with recommendations corresponding to individually identified combinations of risk factors.

The conference Program Committee reviewed 60 submissions for the conference and 10 submissions for the PhD workshop. For the conference, 29 submissions were accepted as full papers, 14 as short papers, 3 as a short demo, whereas 14 submissions were rejected. For the PhD workshop, 9 papers were accepted and 1 was rejected.

According to the conference and workshops program, 55 oral presentations were structured into 11 sessions. Most of the presentations were dedicated to the results of researches conducted in the research organizations located on the territory of Russia, including Dubna, Ekaterinburg, Innopolis, Kazan, Krasnodar, Moscow, Novosibirsk, Obninsk, Tomsk, Voronezh, Pereslavl, Perm, St. Petersburg, Petrozavodsk, Tver, Tyumen and Yaroslavl. Though the conference featured talks prepared by the foreign researchers from countries such as Armenia, France, Germany, Spain, United Kingdom, USA.

For the sake of better visibility of the conference publications by the international scientific community two post-proceedings volumes are submitted to Springer's Communications in Computer and Information Science (CCIS) and CEUR Workshop Proceedings (<http://ceur-ws.org/>) respectively. Seventeen full papers are included into CCIS volume of the DAMDID/RCDL'2020 conference proceedings. Other 29 full and short papers (demos) accepted for the conference are included into this CEUR volume structured into eight sections: Data Integration, Conceptual Models and Ontologies, Advanced Data Analysis Methods, Digital Platforms and Information Systems, Data Analysis in Medicine, Data Analysis in Astronomy and Spectral Data, Information Extraction from Text, PhD Workshop.

The chairs of Program Committee express their gratitude to the Program Committee members for carrying out the reviewing of the submissions and selection of the papers for presentation, to the authors of the submissions, as well as to the host organizers from Voronezh State University. The Program Committee appreciates the possibility of using the Conference Management Toolkit (CMT) sponsored by Microsoft Research, which provided great support during various phases of the paper submission and reviewing process.

December 2020

Bernhard Thalheim  
Sergey Makhortov  
Alexander Sychev