

# Preface

International conference “Data Analytics and Management in Data Intensive Domains” (DAMDID/RCDL’2015) of this year is held on October 13 – 16 in the town of Obninsk, Kaluga region of the Russian Federation. The conference is hosted by the Obninsk education Institute for Nuclear Power Engineering affiliated with the National Research University MEPhI. Obninsk is the first town of science created in USSR in which now many academic and research centers dealing with intensive data analysis in various fields (nuclear physics, modern medicine, oncology, radiology, geophysics, meteorology) are located.

«Data Analytics and Management in Data Intensive Domains» conference (DAMDID) is planned as an interdisciplinary 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 data intensive domains. Approaches to data analysis and management being developed in specific data intensive domains of X-informatics (such as X = astro, bio, chemo, geo, medicine, neuro, physics, 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 the DAMDID/RCDL’2015 conference alongside with traditional data management topics reflects a rapid move into the direction of data science and data intensive analytics. Three conference keynotes form the pivot of the conference program. In the keynote of Peter Wittenburg (Max Planck Data and Compute Center) that opens the conference a survey of the current projects on development of data infrastructures enabling data intensive sciences is given. The second day of the conference is open by the keynote of David Pease (IBM Almaden Research Center). This talk considers objectives and experience of the recently organized IBM Research Lab specifically designed to facilitate complex analytic projects by tackling the challenges of data-intensive scientific discovery. Finally the program of the third day starts with the keynote by Michael Brodie (CSAIL Lab, MIT) in which the author gives analysis and characteristics of the data science as an emerging discipline for data intensive discovery.

Three plenary sessions of the conference can be reckoned as the points of reference of the conference program pivot formed by the keynotes. These are: the invited session on IBM Cognitive Systems with Watson System solutions overview and Watson application examples, particularly in medicine; the panel prepared by the researchers from the eight scientific institutes of the RF devoted to the data access challenges for data intensive research in Russia; and the last session of the conference considering infrastructure solutions intended for support of scientific data and processes.

More than 40 presentations at the scientific sessions at the twelve scientific sessions of the conference cover the problems of data heterogeneity and integration, information extraction from the multistructured data, subject domains modeling (including formation of knowledge bases in medicine), efficiency of computations, semantics of the large textual collections, as well as the specificity of the systems for data analysis (separate session is devoted to the problems of big data analysis in physics), approaches for data intensive problems solving. The majority of these presentations reflect the results of research made in the research institutes, centers and universities located at the different places on the territory of Russia, including: Briansk, Chernogolovka, Dubna, Irkutsk, Jaroslavl, Kazan, Moscow, Nizhny Novgorod, Novosibirsk, Obninsk, Omsk, Pereslavl Zalessky, Saint Petersburg, Tomsk, Chelyabinsk, Vladivostok.

Besides that, the conference includes also several associated events, such as the tutorial on large-scale statistics with MonetDB and R (organized by Hannes Mühleisen (Amsterdam University)); PhD Workshop that includes ten talks related to PhD researches and starts with the keynote by Michael Brodie (CSAIL Lab, MIT) entitled “A 21st Century Applied Computer Science PhD”; open workshop devoted to the social network data analysis.

Special features of the conference DAMDID/RCDL’2015 organization (comparing to previous RCDL conferences) include creation of a new site as well as transfer to the CMT

system use. The chairs of the Program Committee and Organizing Committee of DAMDID/RCDL'2015 express their gratitude to Alexey Vovchenko for the development of the conference site and to Nikolay Skvortsov for the qualified application of the CMT at all stages of the conference preparation.

The chairs of the Organizing Committee and Program Committee of DAMDID/RCDL'2015 express their gratitude to the authors of the submissions as well as to the Russian Foundation for Basic Research and the Department of Nanotechnologies and Information Technologies of the Russian Academy of Sciences for the support of the Conference. The Coordinating committee of the DAMDID/RCDL conferences thanks Director and employees of the Institute for Nuclear Power Engineering of the National Research Nuclear University MEPhI for their hard and responsible work on preparing and carrying out of the Conference as well as the members of the Program Committee for their important work on reviewing and selection of submissions.

**Co-chairs of the Program committee**

Leonid A. Kalinichenko  
(IPI FRC CSC RAS)  
Sergey O. Starkov  
(INPE NRNU MEPhI)

**Co-chairs of the Organizing committee**

Natalia G. Ayrapetova  
(INPE NRNU MEPhI)  
Victor N. Zakharov  
(IPI FRC CSC RAS)