

XX International Conference on Data Analytics and Management in Data  
Intensive Domains  
Conference Program

<b>TUESDAY, OCTOBER 9</b>		
Time		
9.00-16.30	Registration	
9.30-10.00	Welcome and Reception	
10.00-10.15	Opening of the Conference	
10.15-11.30	FAIR WORKSHOP SESSION F1 FAIR in the Realm of Open Science (Chair: Peter Wittenburg)	
11.30-12.00	Coffee Break	
12.00-13.30	FAIR WORKSHOP SESSION F2 FAIR Platforms and Interoperability (Chair: Peter Wittenburg)	
13.30-14.30	Lunch	
14.30-16.00	FAIR WORKSHOP SESSION F3 FAIR Implementation Issues and Activities (Chair: Peter Wittenburg)	
16.00-16.30	Coffee Break	
16.30-18.00	FAIR WORKSHOP SESSION F4 Panel: Data Access Challenges for Data Intensive Research in Russia and EOSC (Chair: Sergey Stupnikov)	
<b>WEDNESDAY, OCTOBER 10</b>		
Time		
9.00-16.00	Registration	
9.30-11.00	SESSION C1 Knowledge Representation (Chair: Bernhard Thalheim)	
11.00-11.30	Coffee Break	
11.30-13.00	SESSION C2 Ontologies and Applications (Chair: Alexander Fazliev)	SESSION C3 Advanced Data Analysis Methods (Chair: Dmitry Namiot)
13.00-14.00	Lunch	
14.00-15.30	SESSION C4 Conceptual and Data Models (Chair: Sergey Stupnikov)	SESSION C5 Data Analysis and Applications (Chair: Dmitry Briukhov)
15.30-16.00	Coffee Break	
16.00-17.30	SESSION C6 Data Analysis in Astronomy (Chair: Oleg Malkov)	SESSION C7 Text Search and Processing (Chair: Victor Zakharov)
17.30-19.00	Coordinating Committee Meeting	
19.00-22.00	Conference Dinner	
<b>THURSDAY, OCTOBER 11</b>		

Time		
9.00-14.00	Registration	
9.30-11.00	SESSION C8 Distributed Computing and Applications of Machine Learning (Chair: Nadezhda Kiselyova)	PHD WORKSHOP SESSION W1 (Chair: Roman Samarev) Internet of Things and Cognitive Systems
11.00-11.30	Coffee Break	
11.30-13.30	SESSION C9 Research Data Infrastructures (Chair: Nikolay Skvortsov)	PHD WORKSHOP SESSION W2 (Chair: Roman Samarev) Data Integration and Data Analysis
13.10-14.15	Lunch	
14.15-18.00	Cultural program	
<b>FRIDAY, OCTOBER 12</b>		
Time		
9.00-11.30	Registration	
9.30-11.00	SESSION C11 Information Extraction from Text I (Chair: Boris Dobrov)	
11.00-11.30	Coffee Break	
11.30-13.00	SESSION C12 Information Extraction from Text II (Chair: Natalia Loukachevitch)	SESSION C10 Poster Session (Chair: Mikhail Zymbler)
13.00-13.30	Closing of the Conference	

## FAIR Data and EOSC Workshop

### F1 FAIR in the Realm of Open Science (Chair: Peter Wittenburg)

Michel Schouppe, Jean-Claude Burgelman (European Commission). Relevance of EOSC and FAIR in the Realm of Open Science and Phases of Implementing the EOSC

Simon Hodson (CODATA). FAIR Data, FAIR Services and the FAIR Data Action Plan

### F2 FAIR Platforms and Interoperability (Chair: Peter Wittenburg)

Ari Asmi (Integrated Carbon Observation System ERIC). Building FAIR Environmental Services Platforms in Europe.

Sergey Stupnikov, Leonid Kalinichenko (FRCCSC RAS). FAIR Data Based on Extensible Unifying Data Model Development

Nikolay Skvortsov (FRCCSC RAS). Meaningful Data Interoperability and Reuse among Heterogeneous Scientific Communities

### F3 FAIR Implementation Issues and Activities (Chair: Peter Wittenburg)

Peter Wittenburg (Max Planck Computing and Data Facility). Digital Objects as a Concept to Help Implementing FAIR and EOSC

Erik Schultes (GO FAIR). Ready, Set, GO FAIR: Accelerating Convergence to an Internet of FAIR Data and Services

Damien Lecarpentier (CSC-EUDAT). How FAIR is EOSC? An Overview of FAIR Activities within the Various EOSC funded initiatives

## F4 Panel (Chair: Sergey Stupnikov)

Panel: Data Access Challenges for Data Intensive Research in Russia and EOSC

## PhD Workshop

### W1 Internet of Things and Cognitive Systems (Chair: Roman Samarev)

Mikhail Shpak (Bauman Moscow State Technical University). Stand to Research the Privacy Limitations of the Internet of Things in the DID's Perspective

Evgeniy Latyshev (Lomonosov Moscow State University). Sensor Data Preprocessing, Feature Engineering and Equipment Remaining Lifetime Forecasting for Predictive Maintenance

Dora Mattyasovszky-Philipp (Eotvos Lorand University, Elte, Informatics Faculty). Research on Cognitive Information Systems in Enterprise Environment

### W2 Data Integration and Data Analysis (Chair: Roman Samarev)

Vladimir Sazontev (Lomonosov Moscow State University). Methods for Big Data Integration in Distributed Computation Environments

Lev Bulygin (Lomonosov Moscow State University). Combining Lexical and Semantic Similarity Measures with Machine Learning Approach for Ontology Matching Problem

Kirill Platonov (Vernadsky State Geological Museum). Methods and Technologies for Integration and Processing of Geographically Distributed Quantitative Geological Information

Roman Degtyarev (Lomonosov Moscow State University). Detection of Melanoma with Deep Convolutional Neural Networks

## Conference

Regular talk is 30 min. length, short talk marked by star (\*) is 15 min. length.

### C1 Knowledge Representation (Chair: Bernhard Thalheim)

Nikolai Lavrentiev, Alexey Privezentsev, Alexander Fazliev (Institute of Atmospheric Optics, SB RAS). Systematization of Tabular and Graphical Resources in Quantitative Spectroscopy

Olga Nevzorova, Alexander Kirilovich (Kazan Federal University), Vladimir Nevzorov (A. N. Tupolev Kazan National Research Technical University), Konstantin Nikolaev (Kazan Federal University). The Semantic Context Models of Mathematical Formulas in Scientific Papers

\*Elena Kozerenko, Konstantin Kuznetsov, Nikolay Somin (FRCCSC RAS). LEEN: Integrated Platform for Multilingual Text Knowledge Processing

\*Victor Telnov, Yuri Korovin (National Research Nuclear University MEPhI) Semantic Base of Nuclear Knowledge

### C2 Ontologies and Applications (Chair: Alexander Fazliev)

Manuk Manukyan (Yerevan State University). On an Ontological Modeling Language by a Non-Formal Example

Tatiana Avdeenko, Anastasiia Timofeeva, Ekaterina Makarova (Novosibirsk State Technical University). Integration of Ontological Case-Based Reasoning with Principal Component Analysis: Application to the IT Support Service

\*Victor Zakharov, Yury Nikitin, Alexander Khoroshilov, Alexey Khoroshilov (FRC CSC RAS). Methods for the Automated Creation of Thematic Ontologies for the Tasks of Providing a Semantic Infrastructure of Information Systems

### C3 Advanced Data Analysis Methods (Chair: Dmitry Namiot)

Mikhail Chernoskutov (Krasovskii Institute of Mathematics and Mechanics, Ural Federal University), Dmitry Ustalov (University of Mannheim). Equidistant Nodes Clustering: a Soft Clustering Algorithm Applied for Synset Induction

Archil Maysuradze (Lomonosov Moscow State University), Bharathi Pilar (University College, Mangalore), B. H. Shekar (Mangalore University). A Novel Metric Combination Approach for Verification Problems

Alexei Myshev, Mikhail Turitsyn (OINPE MEPHI). Fractal Methods in Information Technologies of Processing and Analysis of Parametrically Related Data Streams

### C4 Conceptual and Data Models (Chair: Sergey Stupnikov)

Bernhard Thalheim (University of Kiel). Qualitative and Quantitative Models

Valeriy Chernenkiy, Yuriy Gapanyuk, Yuriy Kaganov, Ivan Dunin, Maxim Lyaskovsky, Vadim Larionov (Bauman Moscow State Technical University). Storing Metagraph Model in Relational, Document-Oriented, and Graph Databases

\*Igor Fiodorov (Plekhanov Russian University of Economics), Alexander Sotnikov (Joint Supercomputer Center of the Russian Academy of Sciences). On the Formalization of Semiotic Models

\*Alexander Afanasyev, Nikolay Voit, Sergey Kirillov, Maria Ukhanova, Irina Ionova (Ulyanovsk State Technical University). Processing of Diagrammatic Workflow on the Basis of Finite State Grammars

### C5 Data Analysis and Applications (Chair: Dmitry Briukhov)

Natalya Filatova, Natalya Bodrina, Konstantin Sidorov, Pavel Shemaev (Tver State Technical University). Organization of Information Support for a Bioengineering System of Emotional Response Research

\*Dmitry Namiot (Lomonosov Moscow State University), Oleg Pokusaev (Center of digital high-speed transport systems Russian University of Transport MIIT), Manfred Sneps-Snepe (AbavaNet), Vasily Kupriyanovsky (National Competence Center for Digital Economy). On Analysis of the Use of Railway Stations in Urban Analytics

\*Mikhail Charnine, Irina Galina (FRC CSC RAS), Oleg Zolotarev (Russian New University). Semantic Cyberspace of Scientific Papers Based on WebVR and Neural Network Technologies

\*Aigul Mukhitova (Novosibirsk State University), Oleg Zhizhimov (Institute of Computational Technologies, SB RAS). Application of XML-technologies in the Construction of Input and Editing Data Model in the Adaptive Administrative Graphical Web Interfaces for Heterogeneous Information Systems

### C6 Data Analysis in Astronomy (Chair: Oleg Malkov)

Oleg Malkov, Dmitry Chulkov, Dana Kovaleva, Alexey Sytov, Alexander Tutukov, Lev Yungelson (Institute of Astronomy, RAS), Yikdem Gebrehiwot (Entoto Observatory and Research Center, Mekelle University), Nikolay Skvortsov (FRC CSC RAS), Solomon Belay Tessema (Ethiopian Space Science and Technology Institute, Entoto Observatory and Research Center Astronomy and Astropysics). Insight into Binary Star Formation via Modelling Visual Binaries Datasets

\*Sergei Vereshchagin, Nataliya Chupina, Ekaterina Postnikova (INASAN). What Will Lead the Astrometry Data Accuracy Breakthrough in the Study of Star Clusters?

\*Elena Mazaeva, Alexey Pozanenko, Pavel Minaev (Space Research Institute, RAS). Using Heterogeneous Data of Multi Frequency Astronomy for Testing Physical Models of GRB Emission

\*Alina Volnova, Alexey Pozanenko, Elena Mazaeva (Space Research Institute, RAS). How to Extract Information from What We Cannot Observe?

### C7 Text Search and Processing (Chair: Victor Zakharov)

Alexander Veretennikov (Ural Federal University). Proximity Full-Text Search with a Response Time Guarantee by Means of Additional Indexes with Multi-Component Keys

Matthias Nitsche, Marina Tropmann-Frick (Hamburg University of Applied Sciences). Context and Embeddings in Language Modelling - an Exploration

\*Oleg Zhizhimov (Institute of Computing Technologies of SB RAS). Technology for Extracting Geographical Names from Text Documents Based on the PostgreSQL

\* Feodor Murzin (A.P. Ershov Institute of Informatics Systems SB RAS), Aigerim Yerimbetova (Novosibirsk State University, Institute of Information and Computational Technologies, Kazakhstan), Madina Tussupova, Kirill Aleksandrov (Novosibirsk State University). Development and Analysis of Technologies of Searching Information Relevant to the Search Query Using Linguistic Support

### C8 Distributed Computing and Applications of Machine Learning (Chair: Nadezhda Kiselyova)

Yana Kraeva, Mikhail Zymbler (South Ural State University). An Efficient Subsequence Similarity Search on Modern Intel Many-core Processors for Data Intensive Applications

Oleg Sen'ko (FRCCSC RAS), Nadezhda Kiselyova (IMET RAS), Victor Dudarev (IMET RAS), Alexander Dokukin (FRC CSC RAS), Vladimir Ryazanov (FRC CSC RAS). Various Machine Learning Methods Efficiency Comparison in Application to Inorganic Compounds Design

\*Stepan Evstifeev (Moscow State University), Ivan Shanin (FRCCSC RAS). Music Genre Classification Based on Signal Processing

\*Maxim Stankevich (ISA FRCCSC RAS), Ivan Smirnov (ISA FRCCSC RAS), Nikolay Ignatiev (RUDN University), Oleg Grigoryev (FRC CSC RAS), Natalia Kiselnikova (Psychological Institute of Russian Academy of Education). Analysis of Big Five Personality Traits by Processing of Social Media Users Activity Features

### C9 Research Data Infrastructures (Chair: Nikolay Skvortsov)

Alexander Fazliev (IAO SB RAS), Evgeny Gordov, Igor Okladnikov (IMCES SB RAS), Alexey Privezetsev (IAO SB RAS), Anna Ryazanova (IMCES SB RAS), Alexander Titov (IMCES SB RAS), Andrey Bart (TSU), Alexander Starchenko (TSU). Applied Tasks in Virtual Research Environment based on a web GIS platform 'Climate+'

Johannes Hedtrich, Ekaterina Fabritz, Christian Henning, Bernhard Thalheim (University of Kiel). Digital Playground for Policy Decision Making

\*Vasily Bunakov, Simon Lambert, Brian Matthews (Science and Technology Facilities Council, Harwell, UK). Persistent Identifiers for Facilities Research: Current Practices and Opportunities

\*Victor Dudarev (IMET RAS), Nadezhda Kiselyova (IMET RAS), Igor Temkin (National University of Science and Technology MISIS). Application Programming Interface for Information System on Inorganic Substances Bandgap

\*Evgenii Viazilov, Nikolay Mikhailov, Alexander Gusev (All-Russian Research Institute for Hydrometeorological Information–World Data Centre). National Automated Accounting System for Keeping Record of Observation Units

#### C10 Poster Session (Chair: Mikhail Zymbler)

Anatolii Fedotov (Institute of Computational Technologies SB RAS), Olga Fedotova (State Public Scientific-Technological Library of SB RAS), Madina Sambetbayeva (Novosibirsk State University, Institute of Information and Computational Technologies CS MES RK, Kazakhstan). Multilingual Thesaurus in Information System for Scientific and Educational Activity Support

Yuliya Leonova, Anatolii Fedotov (Institute of computation technologies of SB RAS). Creation of a Prototype of a Scientific Information System for Managing Information Resources

Vasiliy Kireev, Anna Guseva (National Research Nuclear University MEPhI). Application of Markov chains for the Solution of the Predictive Maintenance Tasks in the Fields of IoT and Housing and Communal Services Management

Elena Sokolova, Anna Moskvina, Olga Mitrofanova (St.-Petersburg State University). Keyphrase Extraction from the Russian Corpus on Linguistics by Means of KEA and RAKE Algorithms

Nina Abramova (FGUP "NICI under MFA of Russia"), Valery Abramov (OOO "Soglasie"), Elena Korotaeva (FGUP "NICI under MFA of Russia"), Valentina Gavrilova (VINITI RAS), Elena Korotaeva (FGUP "NICI under MFA of Russia"). Automatic Generation of Emergencies Information Summary

#### C11 Information Extraction from Text I (Chair: Boris Dobrov)

Nicolay Rusnachenko (BMSTU), Natalia Loukachevitch (Lomonosov Moscow State University). Extracting Sentiment Attitudes from Analytical Texts via Piecewise Convolutional Neural Network

Zinaida Apanovich, Vladimir Isachenko (A.P. Ershov Institute of Informatics Systems). Analysis and Visualization Algorithm for Cross-Language Author Names Disambiguation

\*Anna Kriukova (St. Petersburg State University), Olga Mitrofanova (St.-Petersburg State University), Kirill Sukharev (St. Petersburg Electrotechnical University). Measuring Semantic Relatedness of Russian Texts by Means of Explicit Semantic Analysis

\*Vladimir Barakhnin, Olga Kozhemyakina, Elena Rychkova, Ilya Pastushkov, Yulia Borzilova (Institute of Computational Technologies of SB RAS). The Extraction of the Collocations from the Corpus of Texts in the Process of Automated Processing of Russian Poetic Texts

#### C12 Information Extraction from Text II (Chair: Natalia Loukachevitch)

Ivan Mazharov (Lomonosov Moscow State University), Boris Dobrov (Research Computing Center of Lomonosov Moscow State University). Named Entity Recognition for Information Security Domain

Artem Shelmanov, Dmitriy Deviatkin (ISA FRC CSC), Daniil Larionov (People's Friendship University of Russia). Discovering Novel Emergency Events in Text Streams

Sergey Parinov (CEMI RAS). Open Citation Content Data and a Digital Library

## Abbreviations

EOSC	- European Open Science Cloud
FAIR	- Findable, Accessible, Interoperable, Re-Usable
FRCCSC RAS	- Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences
CSC	- IT Center for Science Ltd.
EUDAT	- European Data Infrastructure

INASAN	- Institute of Astronomy, Russian Academy of Sciences
ISA FRC CSC	- Institute for Systems Analysis, Federal Research Center "Computer Science and Control"
CEMI RAS	- Central Economics and Mathematics Institute, Russian Academy of Sciences
IMET RAS	- Baykov Metallurgy and Materials Institute, Russian Academy of Sciences
IAO SB RAS	- Institute of Atmospheric Optics, Siberian Branch of Russian Academy of Sciences
IMCES SB RAS	- Institute of Monitoring of Climatic and Ecological Systems, Siberian Branch of Russian Academy of Sciences
OINPE MEPHI	- Obninsk Institute for Nuclear Power Engineering
TSU	- Tomsk State University