

XXV International Conference on Data Analytics and Management in Data  
Intensive Domains  
Preliminary Conference Program

24.10.2023 – 27.10.23

Moscow, HSE University

Links for online participants

Sessions CONNECT Conference ID: Code:	Sessions CONNECT Conference ID: Code:
--	--

**Note for chairmen and the authors:** full paper talk is 30 min. (25 min. for the talk, 5 min. for questions and answers); short paper talk is 20 min. (15 min. for the talk, 5 min. for questions and answers). Please follow the timetable strictly.

<b>TUESDAY, OCTOBER 24th</b>	
Time	
9:45-10:00	Opening of the DACOMSIN Workshop
10:00-11:30	<b>SESSION D1</b> <b>AI &amp; ML in Materials Science I</b> Room F201 Chair: Alexandra Khvan
10:00-10:30	Kainan Zhang, Chang Liu, Lixin Tang and Min Xiao Metallographic Structure Recognition with Few Samples Based on Meta-transfer Learning
10:30-11:00	Jingwen Cong, Lixin Tang and Min Xiao A hybrid model based on adaptive active contour and multi-scale attention for pearlite segmentation in metallographic images
11:00-11:30	Hanyu Zhang, Lixin Tang and Yong Shuai GCr15 Bearing Steel Metallographic Image Fine Segmentation Model Based on Transformer Unet Network and CV model
11:30-12:00	<b>Coffee Break</b> <b>Room G109</b>
12:00-13:30	<b>SESSION D2</b> <b>AI &amp; ML in Materials Science II</b> Room F201 Chair: Nadezhda Kiselyova
12:00-12:30	Shicheng Xie, Lixin Tang and Yong Shuai A migration learning based multi-headed attentional convolutional neural network applied to metallographic image denoising

12:30-13:00	Nadezhda Kiselyova, Victor Dudarev, Oleg Senko, Alexander Dokukin, Andrey Stolyarenko and Yuliana Kuznetsova Double halides physical properties prediction by means of machine learning methods	
13:00-13:30	Alexander Glushko, Alexey Neznanov, Olga Maksimenkova, Sergei Kuznetsov and Galina Kuzmicheva Interactive Research Toolbox for Chemical Compounds Analysis Based on Well-interpretable ML Methods	
13:30-14:30	<b>Lunch</b>	
14:30-16:00	<b>SESSION D3 Materials Informatics &amp; Databases</b> Room F201 Chair: Victor Dudarev	
14:30-15:00	Sergey Dyachkov, Adilbek Erkimbaev, Sergey Grigoryev, Pavel Korotaev, Andrey Kosinov, Pavel Levashov, Maxim Maltsev, Dmitry Minakov, Igor Morozov, Mikhail Paramonov, Aleksey Yanilkin and Vladimir Zitserman Database for properties of nuclear reactor materials based on the ontology and NoSQL data format	
15:00-15:30	Victor Dudarev, Nadezhda Kiselyova and Alfred Ludwig Flexible Materials Properties Management System as a Basis for Data-centric Systems in Inorganic Materials Science	
15:30-16:00	Alexander Shevchenko Structure building units of homoleptic coordination compounds	
16:00-16:15	<b>Closing of the DACOMSIN Workshop</b>	
16:30-18:00	<b>Welcome Reception</b>	
<b>WEDNESDAY, OCTOBER 25th</b>		
9:00-9:15	Opening of the Conference	
9:15-11:00	Alexander Petrenko <b>Memorial speech on Sergey Dmitrievich Kuznetsov</b>  Invited Talk: Pavel Velikhov <b>Современная оптимизация запросов в аналитических системах</b>  Invited Talk: Andrey Borodin <b>Scientific and technical dualism: collaborative work of PostgreSQL community and researchers in the field of data management</b>  Room R401 Chair: Sergey Kuznetsov	
11:00-11:30	<b>Coffee Break</b>	
11:30-13.30	SESSION 1 Image Analysis Room R401 Chair: TBA	SESSION 2 Data analysis in medicine and cognitive science Room R207 Chair: TBA

13.30-14.30	<b>Lunch</b>	
14.30-16:00	SESSION 3 Machine learning methods I Room R401 Chair: Dmitry Namiot	SESSION 4 Machine learning applications I Room R207 Chair: TBA
16:00-16.30	<b>Coffee Break</b>	
16.30-18.30	SESSION 5 Machine learning methods II Room R401 Chair: Dmitry Ignatov	SESSION 6 Machine learning applications II Room R207 Chair: TBA
19.00-22:00	Conference Dinner	
<b>THURSDAY, OCTOBER 26th</b>		
9:30-11:00	KEYNOTE: Bernhard Thalheim <b>Kinds of models in science, engineering, and daily life including generic, specific, special</b>  INVITED TALK: Ildar Baimuratov <b>Open Research Knowledge Graph: a semantic approach to scientific communication</b>  Room R204 Chair: Sergey Stupnikov	
11:00-11:30	<b>Coffee Break</b>	
11:30-13:30	SESSION 7 Conceptual models, Ontologies, Semantic Web Room R204 Chair: Sergey Stupnikov	SESSION 8 Data analysis in astronomy I Room R307 Chair: Alexey Pozanenko
13:30-14:30	<b>Lunch</b>	
14:30-16:10	SESSION 9 Research Infrastructures and Models in Earth Sciences Room R204 Chair: Viktor Zakharov	SESSION 10 Data analysis in astronomy II Room R307 Chair: Oleg Malkov
16:30-18:00	Coordinating Committee Meeting	
<b>FRIDAY, OCTOBER 27th</b>		
9:30-11:00	KEYNOTE: Mikhail Zymbler <b>Discord discovery in time series, or Can we detect all anomalies of an anomalously long time series in an anomalously short time?</b>  KEYNOTE: Vladislav A. Blatov <b>Topological Methods and Tools for the Analysis of Big Crystallographic Data</b>  Room R401 Chair: Nadezhda Kiselyova	

11:00-11:30	<b>Coffee Break</b>	
11:30-13:00	<p style="text-align: center;">SESSION 11 Information extraction from text I Room R401 Chair: Boris Dobrov</p>	<p style="text-align: center;">SESSION 12 Social Networks Analysis Room R407 Chair: Nikolay Skvortsov</p>
13:00-14:00	<b>Lunch</b>	
14:00-15:20	<p style="text-align: center;">SESSION 13 Information extraction from text II Room R401 Chair: Natalia Loukachevitch</p>	<p style="text-align: center;">SESSION 14 Advanced algorithms Room R407 Chair: Mikhail Zymbler</p>
15:20-15:50	<b>Coffee Break</b>	
15:50-17:10	<p style="text-align: center;">SESSION 15 Information extraction from text III Room R401 Chair: TBA</p>	
17:10	<b>Closing of the Conference</b>	

### Image Analysis

- Aleksei Samarin, Artyom Nazarenko, Alina Dzestelova, Alexandr Motyko, Alexander Savelev, Aleksei Toropov, Elena Mikhailova and Valentin Malykh UniFi: Universal Filter Model for Image Enhancement
- Elnur Abbasov and Maxim Bakaev Does UI Labeling Data Quality Matter for Predicting Website Aesthetics
- Aleksei Samarin, Aleksei Toropov, Alina Dzestelova, Alexander Savelev, Alexandr Motyko, Valentin Malykh, Elena Mikhailova, Artem Nazarenko, Pavel Dmitriev and Anastasia Golovatiuk Human eye iris and pupil segmentation using infrared camera snapshots
- [short] Pavel Arkhipov, Sergey Philippskih and Maxim Tsukanov A method for creating realistic synthetic examples using a generative deep learning model for classifying anomalies in panoramas

### Advanced algorithms

- Majid Sohrabi, Amir M Fathollahi-Fard and Vasili A Gromov Genetic Engineering Algorithm (GEA): An Efficient Metaheuristic Algorithm for Solving Combinatorial Optimization Problems
- Andrew Soroka, German Mikhelson, Alex Meshcheryakov and Sergey Gerasimov Smart Routes: a system for developing and comparing algorithms for solving vehicle routing problems with realistic constraints
- [short] Ivan Deykin, Vladimir Syuzev and Elena Smirnova Spectral Theory for Multidimensional Digital Data Matrices' Processing

### Machine learning methods I

- Igor Buyanov, Vasiliy Vasiliy Yadrntsev and Ilya Sochenkov Using Autoencoders to improve Nearest Neighbor Search on large Datasets
- Maxim Bakaev Distributional Ground Truth: a Non-Redundant Data Quality Control Method
- Zaur Shibzukhov Some robust variants of the principal components analysis

## Machine learning methods II

- Oleg Pilipenko, Bulat Nutfullin and Vasily Kostyumov TrojanInterpret: A detecting backdoors method in DNN based on neural network interpretation methods
- Temirlan Bidzhiev and Dmitry Namiot Attacks on machine learning models based on the PyTorch framework
- Mariia Zueva and Sergei Kuznetsov Индексы интересности как инструмент отбора формальных понятий для построения нейронной сети на основе решётки формальных понятий
- Xenia Naidenova, Vladimir Parkhomenko, Tatiyana Martirova and Alexander Schukin Plausible Reasoning in an Algorithm of Good Classification Test Generating

## Machine learning applications I

- Anton Dubovskoi, Maksim Sokolov and Ildar Baimuratov Informed Object Detection for Computer Games
- Art Prosvetov and Vladislav Balaev Degradation detection for steam turbines
- Vyacheslav Manevich Corrected Triple Correction Method, CNN and transfer learning for prediction the realized volatility of Bitcoin and E-mini S&P500

## Machine learning applications II

- Erchimen Gavriliev and Tatiana Avdeenko Towards an Approach to Formulating Personal Development Plan for Developers based on Competency Framework and Data Mining
- Artyom Abakumov and Sergey Ereemeev Segmentation of graphical user interface elements based on topological decomposition for GUI testing tasks
- [short] Антон Черненко Использование искусственного интеллекта в патентных исследованиях

## Conceptual models, Ontologies, Semantic Web

- Stepan Vinnikov, Anatoly Nardid and Yuriy Gapanyuk Analysis of the Metagraph Data Model in Terms of Category Theory
- Olga Ataeva, Vladimir Serebriakov and Natalia Tuchkova Построение персонального графа знаний в цифровой семантической библиотеке
- [short] Manuk Manukyan Conceptual Data Model : Concept, Formal Bases, and Implementation Issues
- [short] Victor Telnov and Konstantin Odintsov Experience in Integrating Domain-Specific Knowledge Bases Founded on Semantic Web Standards
- [short] Irina Sobolevskaya, Nikolay Kalenov, Alexander Sotnikov and Svetlana Vlasova Ontology of the Universal Subspace of the Common Digital Space of Scientific Knowledge

## Research infrastructures

- Alexander Fazliev and Nikolay Lavrentiev Quality control of scientific plot collections. The problem and methods of its solution
- [short] Nikolay Kalinin and Nikolay Skvortsov Модель инфраструктуры исследовательских данных для исследований в области компьютерных наук

## Models in Earth Sciences

- Evgenii Viazilov and Denis Melnikov About using models in digital twins of earth sciences

- [short] Владимир Будзко and Виктор Мединников Математическое Моделирование Оценки Влияния Глобальных Изменений в Мире на Цифровой След Российских Аграрных Университетов

#### Data analysis in astronomy I

- Georgiy Mozgunov, Alexei Pozanenko, Pavel Minaev, Ivan Chelovekov and Sergei Grebenev Classification of long gamma-ray transients from INTEGRAL data using machine learning approach
- Aleksandra Avdeeva, Dana Kovaleva and Oleg Malkov Assessing the Reliability of Gaia DR3 Effective Temperatures
- Nicolai Pankov, Artem Prokhorenko, Eugene Schekotihin, Alexei Pozanenko, Pavel Minaev, Sergei Belkin and Alina Volnova AWARE: Alert Watcher and Astronomical Rapid Exploration

#### Data analysis in astronomy II

- Eugene Schekotihin, Nicolai Pankov, Sergey Belkin, Alexei Pozanenko, Pavel Minaev, Alina Volnova and Artem Prokhorenko Применение нейронных сетей для поиска оптических транзиентов на астрономических изображениях методом вычитания
- Volnova Alina, Patrick D. Aleo, Anastasia Lavrukhina, Etienne Russeil, Timofey Semenikhin, Emmanuel Gangler, Emille E. O. Ishida, Matwey V. Kornilov, Vladimir Korolev, Konstantin Malanchev, Maria V. Pruzhinskaya and Sreevarsha Sreejith Exploring the Universe with SNAD: Anomaly Detection in Astronomy.
- [short] Timofey Semenikhin, Matwey Kornilov and Maria Pruzhinskaya Neural network architecture for artifacts detection in ZTF survey
- [short] Anastasia Lavrukhina and Konstantin Malanchev Tree-based machine-learning classifier for stellar flares in The Zwicky Transient Facility survey

#### Information extraction from text I

- Muhammad Shahid Iqbal Malik. Analysis of Cross-lingual and Multi-lingual Text Classification Approaches using Transfer Learning: for high and low-resource languages
- Alisher Rogov and Natalia Loukachevitch Evaluating the Performance of Interpretability Methods in Text Categorization Task
- Anna Glazkova and Dmitry Morozov Cross-Domain Robustness of Transformer-based Keyphrase Generation

#### Information extraction from text II

- Ivan Pimenov and Natalia Salomatina Evaluating the Influence of Argumentation Markers on the Identification of Reasoning Models
- Rodion Sulzhenko and Boris Dobrov Approximation of the meaning for thematic subject headings by simple interpretable representations
- [short] Alexander Sychev Diagnostics of the Topical Model for a Collection of Text Messages Based on Latent Dirichlet Allocation

#### Information extraction from text III

- Vladimir Bochkarev, Stanislav Khristoforov, Anna Shevlyakova and Valery Solovyev Diachronic Analysis of a Word Concreteness Rating: Impact of Semantic Change
- Valentin Malykh, Alexander Kukushkin, Maria Tikhonova and Tatiana Shavrina MOROCCO: Model Resource Comparison Framework
- [short] Konstantin Nikolaev Web-based visualization of semantic annotation of mathematical PDF documents

## Data analysis in medicine and cognitive science

- Dmitry Lyutkin, Dmitry Ignatov, Andrey Soloviev, Dmitry Zhukov and Denis Pozdnyakov  
Transformer-based classification of user queries for medical consultancy with respect to expert specialisation
- Stefan Nikolic, Dmitriy Ignatov and Peter Fedorov Acne severity grading with deep learning
- Konstantin Sorokin, Alexander Levin, Andrey Zaitsev, German Magai, Maxim Beketov and Vladimir Sotskov Global cognitive graph properties dynamics of hippocampal formation
- Almaz Shangareev Reading Progress Tracking: A Novel Autoencoder Model Approach

## Social Networks Analysis

- Samyak Jain, Sarthak Johari and Radhakrishnan Delhibabu Analyzing Cryptocurrency trends using Tweet Sentiment Data and User Meta-Data
- [short] Evelina Bronnikova de Menezes Identifying Life Satisfaction of Social Network Users
- [short] Nagim Davletshin and Konstantin Nikolaev Information System for Predicting Personal Success Based on Open Data from Social Networks
- [short] Kishankumar Bhimani and Khushbu Saradva Decoding the Workplace & EOR: An Employee Survey Analysis by Data Science Techniques and Visualization