

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

23.10.2024 – 25.10.24

Nizhny Novgorod, HSE University

Bolshaya Pecherskaya street, 25/12, ground floor

To have access to the HSE conference building for the participants a PASSPORT is required.

Links for online participants

Keynotes and Invited Talks Sessions 1, 3, 5, 7, 9, 11, 13, 15 CONNECT Conference ID: 838 1113 4373 Code: 284853	Sessions 2, 4, 6, 8, 10, 12, 14, 16 CONNECT Conference ID: 892 4702 4273 Code: 667679
---	--

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.

WEDNESDAY, OCTOBER 23th			
9:45-10:00	Opening of the Conference		
10:00-11:30	Opening Keynote: Panos Pardalos Introduction to data analytics for networks – a historical perspective and major advances Invited Talk: Marat Gilfanov SRG/eROSITA all-sky survey: from stellar flares and neutrino sources to cosmology Room 125 Chairs: Valery Kalyagin		
11:30-12:00	Coffee Break		
12:00-13:30	<table style="width: 100%; border: none;"> <tr> <td style="width: 50%; border: none;"> SESSION 1 Data Integration Room 125 Chair: Dmitrii Shaposhnikov </td> <td style="width: 50%; border: none;"> SESSION 2 Data analysis in astronomy I Room 126 Chair: Alexey Pozanenko </td> </tr> </table>	SESSION 1 Data Integration Room 125 Chair: Dmitrii Shaposhnikov	SESSION 2 Data analysis in astronomy I Room 126 Chair: Alexey Pozanenko
SESSION 1 Data Integration Room 125 Chair: Dmitrii Shaposhnikov	SESSION 2 Data analysis in astronomy I Room 126 Chair: Alexey Pozanenko		
12:00	<table style="width: 100%; border: none;"> <tr> <td style="width: 50%; border: none;"> Sergey Stupnikov Формальная семантика и верификация программ </td> <td style="width: 50%; border: none;"> Ekaterina Malik, Dana Kovaleva, Oleg Malkov, Pavel Kaygorodov and Bernard Debray </td> </tr> </table>	Sergey Stupnikov Формальная семантика и верификация программ	Ekaterina Malik, Dana Kovaleva, Oleg Malkov, Pavel Kaygorodov and Bernard Debray
Sergey Stupnikov Формальная семантика и верификация программ	Ekaterina Malik, Dana Kovaleva, Oleg Malkov, Pavel Kaygorodov and Bernard Debray		

	материализованной интеграции данных на процедурном диалекте SQL	Developing the optimal cross-matching algorithm for the Gaia and BDB data
12:30	Nikolay Skvortsov Управление качеством данных при решении задач над неоднородными источниками данных	Евгений Щекотихин, Николай Панков, Алексей Позаненко, Сергей Белкин, Павел Минаев and Алина Вольнова Применение нейронных сетей для поиска оптических транзиентов на астрономических изображениях методом вычитания
13:00	Андрей Шепелев and Sergey Stupnikov Сопоставление схем с использованием федеративного обучения на гибридном наборе признаков	Альберт Хабибуллин, Алексей Позаненко and Владимир Лозников Моделирование кривых блеска космических гамма-всплесков
13:30-14:30	Lunch	
14:30-16:10	SESSION 3 Conceptual Modeling and Ontologies Room 125 Chair: Nikolay Skvortsov	SESSION 4 Data analysis in astronomy II Room 126 Chair: Alina Volnova
14:30	Stepan Vinnikov, Anatoly Nardid and Yuriy Garanyuk Metagraph Operations using Bigraph Representation	Matwey Kornilov, Vladimir Korolev, Konstantin Malanchev, Anastasia Lavrukhina, Etienne Russeil, Timofey Semenikhin, Emmanuel Gangler, Emille Ishida, Maria Pruzhinskaya, Alina Volnova and Sreevarsha Sreejith Coniferest: an active anomaly detection framework
15:00	Nikolay Kalinin and Nikolay Skvortsov Подход к анализу предметной области информационной безопасности для построения исследовательских инфраструктур	Timofey Semenikhin, Matwey Kornilov, Maria Pruzhinskaya, Anastasia Lavrukhina, Etienne Russeil, Emmanuel Gangler, Emille Ishida, Vladimir Korolev, Konstantin Malanchev, Alina Volnova and Sreevarsha Sreejith Real-bogus classification for ZTF data releases: two approaches
15:30	[short] Olga Ataeva, Vladimir Serebryakov, Natalia Tuchkova and Ivan Strebkov Онтология и граф знаний математической физики в семантической библиотеке	Art Prosvetov, Sergey Grebenev, Sergey Belkin and Alexey Pozanenko Correlations and Classification in Light Curves of Type Ic Supernovae with GRB Association
15:50	[short] Yury Zagorulko, Galina Zagorulko and Elena Sidorova Approach to Developing a Machine Learning Ontology	
16:10-16:30	Coffee Break	
16:30-18:00	SESSION 5 Machine learning methods Room 125 Chair: Tatiana Levanova	SESSION 6 Machine learning methods and applications II Room 126 Chair: Konstantin Barkalov
16:30	Dmitrii Khliustov and Dmitry Kovalev Aggregation Of Regression Models For Variance Minimization	Irina Dvoretzkaya, Alexey Semenov and Alexander Uvarov Exploring Patterns Of Information Literacy Development In Schools:

		Application Of Multilevel Latent Class Analysis To School Students Survey Data
17:00	Dmitrii Shaposhnikov Minimax Approach for Using the Qualitative Preferences in the Multicriteria Evaluation	Alexander Varnavsky Regression models for the AI gaming chatbot for learning programming based on Wordle-type puzzles
17:30	[short] Олег Валентинович Сенько, Александр Александрович Докукин and Федор Александрович Мельник Использование ансамблей с увеличенной дивергенцией в пространстве прогнозов в рекомендательных системах	[short] Maxim Maron, Arkadiy Maron and Danila Tet'Kov Определения вероятностей реализации рисков по прогнозным и фактическим данным
16:30-18:30	Coordinating Committee Meeting Room 305	
18:30	Welcome reception University Cafeteria at the same building	
THURSDAY, OCTOBER 24th		
9:00-9:45	KEYNOTE: Mikhail Zymbler Time reveals all things: Leveraging behavioral patterns for anomaly detection and event prediction in time series Room 125 Chair: Valeriy Kalyagin	
10:00-11:30	SESSION 7 Image Analysis I Room 125 Chair: Valery Kalyagin	SESSION 8 Data Analysis in Neurophysiology Room 126 Chair: Mikhail Zymbler
10:00	Aleksei Samarin, Alexander Savelev, Aleksei Toropov, Artem Nazarenko, Alexandr Motyko, Elena Mikhailova, Egor Kotenko, Alina Dzestelova and Valentin Malykh Modernized Non-Local Blocks for Infrared Camera Image Segmentation of the Human Eye	Anastasiia Timofeeva, Tatiana V. Avdeenko and Sergei Alkov Robust Partial Correlation between EEG Connectivity and Arithmetic Ability
10:30	Aleksei Samarin, Alina Dzestelova, Egor Kotenko, Valentin Malykh, Elena Mikhailova, Alexandr Motyko, Artem Nazarenko, Alexander Savelev, Aleksei Toropov and Aleksandra Dozortseva Automated Feature Engineering Based Approach for Micrococci Microscopic Image Classification and Taxonomic Characteristics Determination	Almaz Shangareev, Ivan Shanin and Sergey Stupnikov Отслеживание прогресса чтения с использованием глубокого обучения на данных окулографии с высоким уровнем шума
11:00	Art Prosvetov, Alexandr Govorov, Maxim Pupkov, Alexandr Andreev and Vladimir Nazarov Illuminating the Moon: Reconstruction of Lunar Terrain Using Photogrammetry, Neural Radiance Fields, and Gaussian Splatting	Margarita Samburova, Albina Lebedeva, Alexander Naumov, Vyacheslav Razin, Nikolay Gromov, Svetlana Gerasimova, Tatiana Levanova and Lev Smirnov Using two interconnected reservoirs to predict mouse hippocampal local field potentials
11:30-	Coffee Break	

12:00		
12:00-13:40	<p>SESSION 9 Image Analysis II Room 125 Chair: Alexander Ponomarenko</p>	<p>SESSION 10 Data Analysis in Medicine Room 126 Chair: Nikolay Zolotykh</p>
12:00	<p>Aleksei Samarin, Alexander Savelev, Aleksei Toropov, Artem Nazarenko, Alexandr Motyko, Elena Mikhailova, Egor Kotenko, Alina Dzestelova and Valentin Malykh ADSAR: Advanced Dual-Stream Attention and Reweighting for Small Object Detection</p>	<p>Vladislav Kuznetsov, Victor Moskalenko and Nikolay Zolotykh Diagnosis of cardiovascular diseases using recurrent and convolutional neural networks</p>
12:30	<p>Anna Shiyan, Ivan Kozlov, Ildar Baimuratov and Nataly Zhukova Plants and their Diseases Recognition: Multiclass and Multilabel Classification Benchmarks</p>	<p>Alexander Varnavsky Model for assessing the need to involve users of social networks in a healthy lifestyle and giving up bad habits according to the data of a social network</p>
13:00	<p>[short] Павел Архипов, Сергей Филиппских and Максим Цуканов Кластерный анализ данных для оптимизации параметров алгоритмов детектирования объектов мобильными нейросетевыми моделями</p>	<p>[short] Vyacheslav Razin and Alexander Krasnov Deep learning to detect the presence of heart disease on the PTB-XL dataset</p>
13:20	<p>[short] Sergey Stasenko, Andrey Lebedev, Olga Shemagina, Irina Nuidel, Andrey Kovalchuk and Vladimir Yakhno Adaptive Correction of Multi-Cascade Detectors in Biomorphic AI for Pattern Recognition</p>	<p>[short] Maxim Kostyukov, Lev Smirnov and Grigory Osipov Neuromorphic reservoir computing for ECG heartbeats classification</p>
13:40-14:30	Lunch	
14:30	City excursion Walking tour over the city. The duration is 3 hours. Start at 14:30 in the lobby of the HSE building at Bolshaya Pecherskaya 25.	
FRIDAY, OCTOBER 25th		
9:00-11:00	INVITED TALK: Alexander Gorban, Ivan Tyukin Multidimensional navigation in high-dimensional data: The Galileo-Feynman paradigm INVITED TALK: Pavel Velikhov Graphs in Relational Databases: How to beat native graph systems INVITED TALK: Andrey Korotchenko Increasing DBMS performance via system architecture and algorithms Room 125 Chair: Sergey Stupnikov	
11:00-11:30	Coffee Break	
11:30-13:30	<p>SESSION 11 Information extraction from text I: Generative and Transformer-Based Models</p>	<p>SESSION 12 Large Language Models and Applications Room 126</p>

	Room 125 Chair: Boris Dobrov	Chair: Alexander Ponomarenko
11:30	Anna Glazkova and Dmitry Morozov Exploring Fine-tuned Generative Models for Keyphrase Selection: A Case Study for Russian	Dmitry Namiot and Elena Zubareva On Open Datasets for LLM Adversarial Testing
12:00	Alexey Sery, Daria Ilina, Elena Sidorova and Yury Zagorulko Applying Generative Neural Networks to Extract Argument Relations from Scientific Communication Texts	Pujun Xie and Anton Khritankov An LLM Approach to Fixing Common Code Issues in Machine Learning Projects
12:30	Alisher Rogov and Natalia Loukachevitch Explaining Transformer-Based Models: a Comparative Study of flan-T5 and BERT Using Post-Hoc Methods	Vasily Kostyumov, Bulat Nutfullin and Oleg Pilipenko Uncertainty-Aware Evaluation for Vision-Language Models
13:00	Elena Bolshakova and Vladislav Semak An Experimental Study on Cross-domain Transformer-Based Term Recognition for Russian	
13:30 – 14:30	Lunch	
14:30-16:20	SESSION 13 Information extraction from text II: Topic models Room 125 Chair: Natalia Loukachevitch	SESSION 14 Data Extraction and Storage Room 126 Chair: Roman Samarev
14:30	Julian Serdyuk, Konstantin Vorontsov and Murat Apishev Hypergraph topic models of document collections	Alexey Shigarov Regular Table Language for Data Extraction from Tables Presented in Electronic Documents
15:00	Bulat Gizatullin and Olga Nevzorova Comparative analysis of methods for topic modeling of mathematical documents	Evgenii Stepanov and Alexey Mitsyuk bXES: a Binary Format For Storing and Transferring Software Event Logs
15:30	[short] Alexander Sychev Анализ тематической модели для размеченной коллекции текстовых сообщений на основе подхода Word2Vec	[short] Евгений Александров, Игорь Александров, Дмитрий Беяков, Наталья Давыдова, Петр Зрелов, Лидия Калмыкова, Мария Любимова, Татьяна Сапожникова, Татьяна Сыресина and Александр Яковлев Возможности разработанной в ЛИТ информационной системы сопровождения лицензий
15:50 – 16:10	Coffee Break	
16:10-16:00	SESSION 15 Information extraction from text III Room 125 Chair: Natalia Loukachevitch	SESSION 16 Data Analysis in Earth Sciences Room 126 Chair: Nikolay Skvortsov
16:10	Евгений Владимирович Волков and Борис Викторович Добров Автоматическое транскрибирование OOV слов для	Евгений Вязилов and Наталия Пузова Перспективы использования средств искусственного интеллекта в гидрометеорологии

	улучшения распознавания терминов предметной области	
16:40	Artem Prosvetov, Alexey Matveev and Alexandr Andreev Decoding the Past: Building a Comprehensive Glagolitic Dataset for Historical Text Analysis	[short] Nikolai Lavrentiev, Alexey Akhlyostin, Alexey Privezentsev and Alexander Fazliev Data quality assessment in large spectral data collections. States and transitions
17:10	[short] Andrey Lovyagin and Boris Dobrov Verifying Factographic Content in Narrative Texts	[short] Vladimir Budzko and Victor Medennikov Экосистемный Подход к Стратегическому Управлению на Примере Сельского Хозяйства
17:30	[short] Olga Gavenko and Sofia Obersht Development and implementation of software application for comparative analysis of the estimates of the complexity of text data	
17:50	Closing of the Conference Room 125	

Data Integration

Sergey Stupnikov Формальная семантика и верификация программ материализованной интеграции данных на процедурном диалекте SQL

Nikolay Skvortsov Управление качеством данных при решении задач над неоднородными источниками данных

Андрей Шепелев and Sergey Stupnikov Сопоставление схем с использованием федеративного обучения на гибридном наборе признаков

Conceptual Modeling and Ontologies

Stepan Vinnikov, Anatoly Nardid and Yuriy Gapanyuk Metagraph Operations using Bigraph Representation

Nikolay Kalinin and Nikolay Skvortsov Подход к анализу предметной области информационной безопасности для построения исследовательских инфраструктур

[short] Olga Ataeva, Vladimir Serebryakov, Natalia Tuchkova and Ivan Strebkov Онтология и граф знаний математической физики в семантической библиотеке

[short] Yury Zagorulko, Galina Zagorulko and Elena Sidorova Approach to Developing a Machine Learning Ontology

Data analysis in astronomy I

Ekaterina Malik, Dana Kovaleva, Oleg Malkov, Pavel Kaygorodov and Bernard Debray Developing the optimal cross-matching algorithm for the Gaia and BDB data

Евгений Щекотихин, Николай Панков, Алексей Позаненко, Сергей Белкин, Павел Минаев and Алина Вольнова Применение нейронных сетей для поиска оптических транзиентов на астрономических изображениях методом вычитания

Альберт Хабибуллин, Алексей Позаненко and Владимир Лозников Моделирование кривых блеска космических гамма-всплесков

Data analysis in astronomy II

Matwey Kornilov, Vladimir Korolev, Konstantin Malanchev, Anastasia Lavrukhina, Etienne Russeil, Timofey Semenikhin, Emmanuel Gangler, Emille Ishida, Maria Pruzhinskaya, Alina Volnova and Sreevarsha Sreejith
Coniferest: an active anomaly detection framework

Timofey Semenikhin, Matwey Kornilov, Maria Pruzhinskaya, Anastasia Lavrukhina, Etienne Russeil, Emmanuel Gangler, Emille Ishida, Vladimir Korolev, Konstantin Malanchev, Alina Volnova and Sreevarsha Sreejith
Real-bogus classification for ZTF data releases: two approaches

Art Prosvetov, Sergey Grebenev, Sergey Belkin and Alexey Pozanenko Correlations and Classification in Light Curves of Type Ic Supernovae with GRB Association

Machine learning methods

Dmitrii Khliustov and Dmitry Kovalev Aggregation Of Regression Models For Variance Minimization

Dmitrii Shaposhnikov Minimax Approach for Using the Qualitative Preferences in the Multicriteria Evaluation

[short] Олег Валентинович Сенько, Александр Александрович Докукин and Федор Александрович Мельник Использование ансамблей с увеличенной дивергенцией в пространстве прогнозов в рекомендательных системах

Machine Learning and Statistics Applications

Irina Dvoretzskaya, Alexey Semenov and Alexander Uvarov Exploring Patterns Of Information Literacy Development In Schools: Application Of Multilevel Latent Class Analysis To School Students Survey Data

Alexander Varnavsky Regression models for the AI gaming chatbot for learning programming based on Wordle-type puzzles

[short] Maxim Maron, Arkadiy Maron and Danila Tet'Kov Определения вероятностей реализации рисков по прогнозным и фактическим данным

Image Analysis I

Aleksei Samarin, Alexander Savelev, Aleksei Toropov, Artem Nazarenko, Alexandr Motyko, Elena Mikhailova, Egor Kotenko, Alina Dzestelova and Valentin Malykh Modernized Non-Local Blocks for Infrared Camera Image Segmentation of the Human Eye

Aleksei Samarin, Alina Dzestelova, Egor Kotenko, Valentin Malykh, Elena Mikhailova, Alexandr Motyko, Artem Nazarenko, Alexander Savelev, Aleksei Toropov and Aleksandra Dozortseva Automated Feature Engineering Based Approach for Micrococci Microscopic Image Classification and Taxonomic Characteristics Determination

Art Prosvetov, Alexandr Govorov, Maxim Pupkov, Alexandr Andreev and Vladimir Nazarov
Illuminating the Moon: Reconstruction of Lunar Terrain Using Photogrammetry, Neural Radiance Fields, and Gaussian Splatting

Image Analysis II

Aleksei Samarin, Alexander Savelev, Aleksei Toropov, Artem Nazarenko, Alexandr Motyko, Elena Mikhailova, Egor Kotenko, Alina Dzestelova and Valentin Malykh ADSAR: Advanced Dual-Stream Attention and Reweighting for Small Object Detection

Anna Shiyan, Ivan Kozlov, Ildar Baimuratov and Nataly Zhukova Plants and their Diseases Recognition: Multiclass and Multilabel Classification Benchmarks

[short] Павел Архипов, Сергей Филиппских and Максим Цуканов Кластерный анализ данных для оптимизации параметров алгоритмов детектирования объектов мобильными нейросетевыми моделями

[short] Sergey Stasenko, Andrey Lebedev, Olga Shemagina, Irina Nuidel, Andrey Kovalchuk and Vladimir Yakhno Adaptive Correction of Multi-Cascade Detectors in Biomorphic AI for Pattern Recognition

Data Analysis in Neurophysiology

Anastasiia Timofeeva, Tatiana V. Avdeenko and Sergei Alkov Robust Partial Correlation between EEG Connectivity and Arithmetic Ability

Almaz Shangareev, Ivan Shanin and Sergey Stupnikov Отслеживание прогресса чтения с использованием глубокого обучения на данных окулографии с высоким уровнем шума

Margarita Samburova, Albina Lebedeva, Alexander Naumov, Vyacheslav Razin, Nikolay Gromov, Svetlana Gerasimova, Tatiana Levanova and Lev Smirnov Using two interconnected reservoirs to predict mouse hippocampal local field potentials

Data Analysis in Medicine

Vladislav Kuznetsov, Victor Moskalenko and Nikolay Zolotykh Diagnosis of cardiovascular diseases using recurrent and convolutional neural networks

Alexander Varnavsky Model for assessing the need to involve users of social networks in a healthy lifestyle and giving up bad habits according to the data of a social network

[short] Vyacheslav Razin and Alexander Krasnov Deep learning to detect the presence of heart disease on the PTB-XL dataset

[short] Maxim Kostyukov, Lev Smirnov and Grigory Osipov Neuromorphic reservoir computing for ECG heartbeats classification

Information extraction from text I: Generative and Transformer-Based Models

Anna Glazkova and Dmitry Morozov Exploring Fine-tuned Generative Models for Keyphrase Selection: A Case Study for Russian

Alexey Sery, Daria Ilina, Elena Sidorova and Yury Zagorulko Applying Generative Neural Networks to Extract Argument Relations from Scientific Communication Texts

Alisher Rogov and Natalia Loukachevitch Explaining Transformer-Based Models: a Comparative Study of flan-T5 and BERT Using Post-Hoc Methods

Elena Bolshakova and Vladislav Semak An Experimental Study on Cross-domain Transformer-Based Term Recognition for Russian

Information extraction from text II: Topic models

Julian Serdyuk, Konstantin Vorontsov and Murat Apishev
collections

Hypergraph topic models of document

Bulat Gizatullin and Olga Nevzorova Comparative analysis of methods for topic modeling of
mathematical documents

[short] Alexander Sychev Анализ тематической модели для размеченной коллекции текстовых
сообщений на основе подхода Word2Vec

Information extraction from text III

Евгений Владимирович Волков and Борис Викторович Добров Автоматическое
транскрибирование OOV слов для улучшения распознавания терминов предметной области

Artem Prosvetov, Alexey Matveev and Alexandr Andreev Decoding the Past: Building a
Comprehensive Glagolitic Dataset for Historical Text Analysis

[short] Andrey Lovyagin and Boris Dobrov Verifying Factographic Content in Narrative Texts

[short] Olga Gavenko and Sofia Obersht Development and implementation of software application for
comparative analysis of the estimates of the complexity of text data.

Large Language Models and Applications

Dmitry Namiot and Elena Zubareva On Open Datasets for LLM Adversarial Testing

Pujun Xie and Anton Khritankov An LLM Approach to Fixing Common Code Issues in Machine Learning
Projects

Vasily Kostyumov, Bulat Nutfullin and Oleg Pilipenko Uncertainty-Aware Evaluation for Vision-Language
Models

Data Extraction and Storage

Alexey Shigarov Regular Table Language for Data Extraction from Tables Presented in Electronic Documents

Evgenii Stepanov and Alexey Mitsyuk bXES: a Binary Format For Storing and Transferring Software Event
Logs

[short] Евгений Александров, Игорь Александров, Дмитрий Беляков, Наталья Давыдова, Петр Зрелов,
Лидия Калмыкова, Мария Любимова, Татьяна Сапожникова, Татьяна Сыресина and Александр
Яковлев Возможности разработанной в ЛИТ информационной системы сопровождения
лицензий

Data Analysis in Earth Sciences

Евгений Вязилов and Наталия Пузова Перспективы использования средств искусственного интеллекта
в гидрометеорологии

[short] Nikolai Lavrentiev, Alexey Akhlyostin, Alexey Privezentsev and Alexander Fazliev Data quality
assessment in large spectral data collections. States and transitions

[short] Vladimir Budzko and Victor Medennikov Экосистемный Подход к Стратегическому Управлению
на Примере Сельского Хозяйства