Preface

This volume contains the refereed proceedings of the 7th International Conference on Analysis of Images, Social Networks, and Texts (AIST 2018)[1]. The previous conferences during 2012–2017 attracted a significant number of data scientists – students, researchers, academics, and engineers working on interdisciplinary data analysis of images, texts, and social networks.

The broad scope of AIST made it an event where researchers from different domains, such as image and text processing, exploiting various data analysis techniques, can meet and exchange ideas. We strongly believe that this may lead to cross-fertilisation of ideas between researchers relying on modern data analysis machinery.

Therefore, AIST brought together all kinds of applications of data mining and machine learning techniques. The conference allowed specialists from different fields to meet each other, present their work, and discuss both theoretical and practical aspects of their data analysis problems. Another important aim of the conference was to stimulate scientists and people from industry to benefit from the knowledge exchange and identify possible grounds for fruitful collaboration.

The conference was held during July 5–7, 2018. The conference was organized in Moscow, the capital of Russia, on the campus of Moscow Polytechnic University (MPU)[2].

This year, the key topics of AIST were grouped into six tracks:

1. **General topics of data analysis** chaired by Sergei Kuznetsov (Higher School of Economics, Russia) and Amedeo Napoli (LORIA, France);
2. **Natural Language Processing** chaired by Natalia Loukachevitch (Lomonosov Moscow State University, Russia) and Goran Glavaš (University of Mannheim, Mannheim, Germany);
3. **Social Network Analysis** chaired by Vladimir Batagelj (University of Ljubljana, Slovenia) and Olessia Koltsova (Higher School of Economics, Russia);
4. **Analysis of images and video** chaired by Marcello Pelillo (University of Venice, Italy) and Andrey Savchenko (Higher School of Economics, Russia);
5. **Optimization problems on graphs and network structures** chaired by Panos Pardalos (University of Florida, USA) and Michael Khachay (IMM UB RAS and Ural Federal University, Russia);
6. **Analysis of dynamic behavior through event data** chaired by Wil van der Aalst (RWTH Aachen University, Germany) and Irina Lomazova (Higher School of Economics, Russia).

The Programme Committee and the reviewers of the conference included 154 well-known experts in data mining and machine learning, natural language pro-

[1] http://aistconf.org/
[2] http://mospolytech.ru/?eng

CEUR-WS.org/Vol-2268/preface.pdf

cessing, image processing, social network analysis, and related areas from leading institutions of 32 countries including Argentina, Australia, Austria, Bangladesh, Canada, China, Croatia, Czech Republic, Denmark, Egypt, Finland, France, Germany, Greece, India, Italy, Kazakhstan, Lithuania, Norway, Qatar, Romania, Russia, Slovenia, Spain, Switzerland, Vietnam, Taiwan, Turkey, Ukraine, The Netherlands, UK, and USA. This year we received 107 submissions mostly from Russia but also from Austria, Bangladesh, Finland, Germany, , Norway, Switzerland, Taiwan, Turkey, Ukraine, USA, and Vietnam.

Out of 107 submissions, only 29 papers were accepted as regular oral papers. Thus, the acceptance rate of this volume was around 36% (not taking into account 26 automatically rejected papers). In order to encourage young practitioners and researchers, we included 28 papers to the supplementary proceedings after their poster presentation at the conference. Each submission was reviewed by at least three reviewers, experts in their fields, in order to supply detailed and helpful comments.

The conference featured several invited talks and an industry session dedicated to current trends and challenges.

The topical opening talk was presented by Rostislav Yavorskiy on "Visualization of Data Science Community in Russia."

The invited talks were:

– Anne Rozinat (Fluxicon), "Process Mining: Discovering Process Maps from Data";

– Goran Glavaš (University of Mannheim), "Cars, Drivers, Vehicles, and Wheels: Specializing Distributional Word Vectors for Lexico-Semantic Relations";

– Gleb Gusev (Yandex and Moscow Institute of Physics and Technology), "Machine Learning Research in Yandex and CatBoost";

– Ekaterina Chernyak of HSE and Sberbank presented a tutorial on "Supervised Methods in Natural Language Processing."

The business speakers also covered a wide variety of topics. We list those invited talks below:

– Iosif Itkin (Exactpro Systems), "The World's Most Valuable Commodity";

– Maksim Balashevich (Santiment), "Crypto assets valuation";

– Ilya Utekhin (the European University at St. Petersburg), "On understanding users for human-centered design";

– Denis Stepanov (JetBrains), "Workshop: Data visualization with datalore.plot in research, education and just for fun";

– Eugeny Malyutin (OK.ru), "Trend detection at OK";

– Oleg Klepikov (Center for Applied Neuroeconomics), "Psychometric behavorial classification of banking clients – a case study";

– Dmitriy Skougarevskiy (the European University at St. Petersburg), "Semantic representation of accident reports to understand the dark figure of crime."

Under a special agreement between the responsible program co-chairs and the proceedings chair, due to its interdisciplinary nature, one paperwas set to a special section on Innovative Systems (outside of the main tracks).

We would like to thank the authors for submitting their papers and the members of the Programme Committee for their efforts in providing exhaustive reviews.

According to the Program Chairs taking into account the obtained reviews and presentation quality, the best paper awards were granted to the following papers:

Track 1. **General topics of data analysis**: "Lookup Lateration: Mapping of Received Signal Strength to Position for Geo-Localization in Outdoor Urban Areas" by Andrey Shestakov, Attila Kertesz-Farkas, Dmitri Shmelkin and Danila Doroshin;

Track 2. **Natural language processing:** "Learning representations for soft skill matching" by Luiza Sayfullina, Eric Malmi and Juho Kannala;

Track 3. **Social network analysis**: "Organizational networks revisited: Predictors of headquarters-subsidiary relationship perception" by Antonina Milekhina, Valentina Kuskova and Elena Artyukhova;

Track 4. **Analysis of Images and Video**: "Application of Fully Convolutional Neural Networks to Mapping Industrial Oil Palm Plantations" by Artem Baklanov, Mikhail Khachay and Maxim Pasynkov;

Track 5. **Optimization problems on graphsand network structures**: "Efficient PTAS for the Euclidean CVRP with Time Windows" by Michael Khachay and Yuri Ogorodnikov;

Track 6. **Analysis of dynamic behavior through event data**: "Neural Approach to the Discovery Problem in Process Mining" by Timofey Shunin, Natalia Zubkova and Sergey Shershakov.

We would also like to express our special gratitude to all the invited speakers and industry representatives.

We deeply thank all the partners and sponsors. Our golden sponsors is Exactpro.

Exactpro, a fully owned subsidiary of the London Stock Exchange Group, specializes in quality assurance for exchanges, investment banks, brokers, and other financial sector organizations. Our special thanks go to Springer for their help, starting from the first conference call to the final version of the proceedings. JetBrains covered important conference expenses while Springer sponsored the best paper awards. Last but not least, we are grateful to the dean of Information Technologies Faculty of MPU, Andrey Phillipovich, and all the organizers, especially to Andrey Novikov and Marina Danshina, and the volunteers, whose endless energy saved us at the most critical stages of the conference preparation.

Here, we would like to mention the Russian word "aist" is more than just a simple abbreviation (in Cyrillic) – it means a "stork." Since it is a wonderful free bird, a symbol of happiness and peace, this stork gave us the inspiration to organize the AIST conference. So we believe that this young and rapidly growing conference will likewise be bringing inspiration to data scientists around the world!

October 2018

Wil van der Aalst
Vladimir Batagelj
Goran Glavaš
Dmitry Ignatov
Michael Khachay
Olessia Koltsova
Sergei Kuznetsov
Irina Lomazova
Natalia Loukachevitch
Amedeo Napoli
Alexander Panchenko
Panos Pardalos
Marcello Pelillo
Andrey Savchenko

Organisation

Program Committee Chairs

Wil van der Aalst	RWTH Aachen University, Germany
Vladimir Batagelj	University of Ljubljana, Ljubljana, Slovenia
Goran Glavaš	University of Mannheim, Mannheim, Germany
Michael Khachay	Krasovskii Institute of Mathematics and Mechanics of Russian Academy of Sciences, Russia & Ural Federal University, Yekaterinburg, Russia
Sergei Kuznetsov	National Research University Higher School of Economics, Moscow, Russia
Olessia Koltsova	National Research University Higher School of Economics, St. Petersburg, Russia
Amedeo Napoli	LORIA CNRS, University of Lorraine, and INRIA, Nancy, France
Irina Lomazova	National Research University Higher School of Economics, Moscow, Russia
Natalia Loukachevitch	Computing Centre of Lomonosov Moscow State University, Russia
Alexander Panchenko	University of Hamburg, Germany & Université catholique de Louvain, Belgium
Panos Pardalos	University of Florida, US
Marcello Pelillo	University of Venice, Italy
Andrey Savchenko	National Research University Higher School of Economics, Nizhny Novgorod, Russia

Proceedings Chair

Dmitry I. Ignatov	National Research University Higher School of Economics, Russia

Business Day Chair

Rostislav Yavorskiy	National Research University Higher School of Economics, Moscow

Steering Committee

Dmitry I. Ignatov	National Research University Higher School of Economics, Russia
Michael Khachay	Krasovskii Institute of Mathematics and Mechanics of Russian Academy of Sciences, Russia & Ural Federal University, Yekaterinburg, Russia
Alexander Panchenko	University of Hamburg, Germany & Université catholique de Louvain, Belgium
Rostislav Yavorskiy	National Research University Higher School of Economics, Russia

Program Committee

Ekaterina Arkhangelskaya	St. Petersburg Department of Steklov Institute of Mathematics of Russian Academy of Sciences, Russia
Aleksey Artamonov	Neuromation, US
Xiang Bai	Huazhong University of Science and Technology, China
Jaume Baixeries	Universitat Politècnica de Catalunya, Spain
Amir Bakarov	National Research University Higher School of Economics, Russia
Artem Baklanov	International Institute for Applied Systems Analysis, Austria
Lamberto Ballan	University of Padova, Italy
Vladimir Batagelj	Institute of Mathematics, Physics and Mechanics, Slovenia
Timo Baumann	University of Hamburg, Germany
Malay Bhattacharyya	Indian Statistical Institute, Kolkata
Chris Biemann	University of Hamburg, Germany
Battista Biggio	University of Cagliari, Italy
Elena Bolshakova	Lomonosov Moscow State University, Russia
Samuel Bulò	Fondazione Bruno Kessler, Italy
Andrea Burattin	Technical University of Denmark, Denmark
Evgeny Burnaev	Skolkovo Institute of Science and Technology, Russia
Aleksey Buzmakov	National Research University Higher School of Economics, Perm, Russia
Ignacio Cassol	Universidad Austral, Argentina
Artem Chernodub	Institute of Mathematical Machines and Systems Problems of National Academy of Sciences of Ukraine, Ukraine
Mikhail Chernoskutov	Krasovskii Institute of Mathematics and Mechanics of the Ural Branch of the Russian Academy of Sciences and Ural Federal University, Russia

Bonaventura Coppola	University of Trento, Italy
Hernani Costa	University of Malaga, Spain
Massimiliano de Leoni	Eindhoven University of Technology, the Netherlands
Boris Dobrov	Recearch Computing Center of Lomonosov Moscow State University, Russia
Sofia Dokuka	National Research University Higher School of Economics, Russia
Dirk Fahland	Eindhoven University of Technology, the Netherlands
Stefano Faralli	University of Mannheim, Germany
Victor Fedoseev	Samara National Research University, Russia
Elena Filatova	City University of New York, US
Olga Gerasimova	National Research University Higher School of Economics, Russia
Edward Kh. Gimadi	Sobolev Institute of Mathematics of the Siberian Branch of the Russian Academy of Sciences, Russia
Goran Glavaš	University of Mannheim, Germany
Ivan Gostev	National Research University Higher School of Economics, Russia
Natalia Grabar	STL CNRS Université Lille 3, France
Artem Grachev	Samsung R&D, Russia
Dmitry Granovsky	Yandex, Russia
Anna Gromova	Exactpro, Russia
Marianne Huchard	LIRMM, Université Montpellier 2 et CNRS, France
Dmitry Ignatov	National Research University Higher School of Economics, Russia
Dmitry Ilvovsky	National Research University Higher School of Economics, Russia
Vladimir Ivanov	Innopolis University, Russia
Pei Jun	Hefei University of Technology, China
Alex Jung	Aalto University, Finland
Anna Kalenkova	National Research University Higher School of Economics, Russia
Ilia Karpov	National Research University Higher School of Economics, Russia
Nikolay Karpov	National Research University Higher School of Economics, Nizhniy Novgorod, Russia
Egor Kashkin	Vinogradov Russian Language Institute of Russian Academy of Sciences, Russia
Alexander Kelmanov	Sobolev Institute of Mathematics of Siberian Branch of Russian Academy of Sciences, Russia
Attila Kertesz-Farkas	National Research University Higher School of Economics, Russia

Michael Khachay	Krasovskii Institute of Mathematics and Mechanics, Ural Branch of Russian Academy of Sciences and Ural Federal University, Russia
Sergey Khamidullin	Sobolev Institute of Mathematics, Siberian Branch of Russian Academy of Sciences, Russia
Oleg Khamisov	Melentiev Energy Systems Institute, Siberian Branch of the Russian Academy of Sciences, Russia
Vladimir Khandeev	Sobolev Institute of Mathematics, Siberian Branch of Russian Academy of Sciences, Russia
Alexander Kharlamov	National Research University Higher School of Economics, Russia
Javad Khodadoust	National Research University Higher School of Economics, Russia
Konstantin Kobylkin	Krasovskii Institute of Mathematics and Mechanics, Ural Branch of Russian Academy of Sciences, Russia
Yury Kochetov	Sobolev Institute of Mathematics, Siberian Branch of Russian Academy of Sciences, Russia
Sergei Koltcov	National Research University Higher School of Economics, Russia
Olessia Koltsova	National Research University Higher School of Economics, Russia
Jan Konecny	Palacky University, Czech Republic
Andrey Kopylov	Tula State University, Russia
Mikhail Korobov	ScrapingHub Inc., Ireland
Evgeny Kotelnikov	Vyatka State University, Russia
Ilias Kotsireas	Wilfrid Laurier University and Maplesoft, Canada
Fedor Krasnov	Gazprom Neft, Russia
Olga Krasotkina	Markov Processes International, US
Ekaterina Krekhovets	National Research University Higher School of Economics, Russia
Tomas Krilavicius	Vytautas Magnus University, Lithuania
Rauf Kurbanov	Neuromation, US
Andrey Kutuzov	University of Oslo, Norway
Andrey Kuznetsov	Samara National Research University, Russia
Sergei Kuznetsov	National Research University Higher School of Economics, Russia
Florence Le Ber	ICube Laboratory and Ecole Nationale du Génie de l'Eau et de l'Environnement de Strasbourg, France
Alexander Lepskiy	National Research University Higher School of Economics, Russia
Benjamin Lind	Anglo-American School of St. Petersburg, Russia
Ma Lizhuang	Shanghai Jiao Tong University, China
Nikola Ljubešić	Jožef Stefan Institute, Slovenia and University of Zagreb, Croatia

Irina Lomazova	National Research University Higher School of Economics, Russia
Natalia Loukashevich	Lomonosov Moscow State University, Russia
Ilya Makarov	National Research University Higher School of Economics, Russia
Alexey Malafeev	National Research University Higher School of Economics, Russia
Yury Malkov	Institute of Applied Physics of the Russian Academy of Sciences, Russia
Eric Malmi	Google, Switzerland
Luis Marujo	Snap Research and Carnegie Mellon University, US
Eyasu Zemene Mequanint	Ca'Focasri University of Venice, Italy
Nizar Messai	LI - Université François Rabelais, France
Tristan Miller	Technische Universität Darmstadt, Germany
Olga Mitrofanova	St. Petersburg State University, Russia
Evgeny Myasnikov	Samara National Research University, Russia
Federico Nanni	University of Mannheim, Germany
Amedeo Napoli	Université de Lorraine, Nancy, France
Kirill Nikolaev	National Research University Higher School of Economics, Russia
Damien Nouvel	Institut National des Langues et Civilisations Orientales, France
Evgeniy M. Ozhegov	Higher School of Econommics
Alina Ozhegova	National Research University Higher School of Economics
Anna Panasenko	Novosibirsk State University, Russia
Alexander Panchenko	University of Hamburg, Germany
Panos Pardalos	University of Florida, US
Marcello Pelillo	University of Venice, Italy
Georgios Petasis	National Centre for Scientific Research "Demokritos"
Stefan Pickl	Universität der Bundeswehr München
Lidia Pivovarova	University of Helsinki, Finland
Vladimir Pleshko	RCO LLC, Russia
Alexander Porshnev	National Research University Higher School of Economics, Nizhniy Novgorod, Russia
Surya Prasath	Cincinnati Children's Hospital Medical Center, US
Andrea Prati	University of Parma, Italy
Artem Pyatkin	Novosibirsk State University and Sobolev Institute of Mathematics, Siberian Branch of Russian Academy of Sciences, Russia
Evgeniy Riabenko	Facebook, UK
Martin Riedl	University of Stuttgart, Germany
Alexey Romanov	University of Massachusetts Lowell, US
Yuliya Rubtsova	Ershov Institute of Informatics Systems, Siberian Branch of Russian Academy of Sciences, Russia

Alexey Ruchay	Chelyabinsk State University, Russia
Eugen Ruppert	University of Hamburg, Germany
Christian Sacarea	Babes-Bolyai University, Romania
Mohammed Abdel-Mgeed M. Salem	German University in Cairo, Egypt
Andrey Savchenko	National Research University Higher School of Economics, Nizhniy Novgorod, Russia
Friedhelm Schwenker	Ulm University, Germany
Oleg Seredin	Tula State University, Russia
Andrey Shcherbakov	University of Melbourne, Australia
Henry Soldano	Université Paris-Nord, France
Alexey Sorokin	Lomonosov Moscow State University, Russia
Tobias Staron	University of Hamburg, Germany
Dmitry Stepanov	Ailamazyan Program Systems Institute of Russian Academy of Sciences, Russia
Vadim Strijov	Dorodnicyn Computing Centre of Russian Academy of Sciences, Russia
Diana Sungatullina	Skolkovo Institute of Science and Technology, Russia
Yonatan Tariku	University of Central Florida, US
Irina Temnikova	Qatar Computing Research Institute, Qatar
Rocco Tripodi	University of Venice, Italy
Diana Troanca	Babes-Bolyai University, Romania
Christos Tryfonopoulos	University of Peloponnese, Greece
Elena Tutubalina	Kazan Federal University, Russia
Dmitry Ustalov	University of Mannheim, Germany
Wil van der Aalst	RWTH Aachen University, Germnay
Sebastiano Vascon	University Ca' Foscari of Venice, Italy
Natalia Vassilieva	Hewlett Packard Labs, Russia
Dmitry Vetrov	Lomonosov Moscow State University and National Research University Higher School of Economics, Russia
Ekaterina Vylomova	University of Melbourne, Australia
Gui-Song Xia	Wuhan University, China
Roman Yangarber	University of Helsinki, Finland
Marcos Zampieri	University of Wolverhampton, UK
Alexey Zobnin	National Research University Higher School of Economics, Russia
Nikolai Zolotykh	University of Nizhniy Novgorod, Russia

Additional Reviewers

Mikhail Batsyn
Marc Chaumont
Ivan Grechikhin
Arsenis Tsokas
Alexander Rassadin
Yuri Rykov
François Suro

Organising Committee

Rostislav Yavorskiy, Conference Chair	National Research University Higher School of Economics, Russia
Andrey Novikov, Head of Organization	National Research University Higher School of Economics, Russia
Marina Danshina, Venue Organization & Management	Moscow Polytechnic University, Russia
Anna Ukhanaeva, Information Partners and Communications	National Research University Higher School of Economics, Russia
Anna Kalenkova, Visa Support and International Communications	National Research University Higher School of Economics, Russia
Alexander Gnevshev, Venue Organization & Management	Moscow Polytechnic University, Russia

Volunteers

Daniil Bannyh	Moscow Polytechnic University, Russia
Ksenia Belkova	Moscow Polytechnic University, Russia
Tatiana Mishina	Moscow Polytechnic University, Russia
Zahar Kuhtenkov	Moscow Polytechnic University, Russia
Maxim Pasynkov	Krasovsky Institute of Mathematics and Mechanics of RAS, Russia, Yekaterinburg
Ivan Poylov	Moscow Polytechnic University, Russia

Sponsors

Golden sponsor:	Exactpro	
Bronze sponsors:	JetBrains	Springer

Table of Contents

Unsupervised Co-Authorship Based Algorithm for Clustering of R&D Trends at Science and Technology Centers in Oil and Gas Industry 1
Fedor Krasnov and Mars Khasanov

Method for Determining Ridge Lines of a Latent Fingerprint 13
Kirill Milovidov, Vladimir Gudkov and Daria Lepikhova

The Framework for Hypothesis Verification and Analysis of Natural Language Processing Results for the Russian Language 25
Ekaterina Politsyna, Sergey Politsyn and Alexander Porechny

Russian Person Names Recognition Using the Hybrid Approach 34
Anna Glazkova

A Hybrid Chatbot System Combining Question Answering and Knowledge-Base Approaches 42
Yuriy Gapanyuk, Sergey Chernobrovkin, Aleksey Leontiev, Igor Latkin, Marina Belyanova and Oleg Morozenkov

Automated Analysis of Micro-contexts of Word for Construction of Its Lexicographic Description 54
Nadezhda Lagutina, Yuliya Tsofina, Ilya Paramonov, Ksenia Lagutina and Natalia Kasatkina

Data Mining on the Use of Railway Stations 66
Dmitry Namiot, Oleg Pokusaev and Vasily Kupriyanovsky

Aspect-Based Sentiment Analysis of Russian Hotel Reviews 75
Valery Rybakov and Alexey Malafeev

Automatic Morphemic Analysis of Russian Words 85
Lyudmila Maltina and Alexey Malafeev

Improving the Efficiency of Entropy Coding Method in Video Encoder H.265/HEVC 95
Andrey Aleksandrovich Tropchenko, Tien Ban Doan and Van Truong Nguyen

Text Augmentation for Neural Networks 104
Anna Mosolova, Vadim Fomin and Ivan Bondarenko

Characterizing Text Complexity with Core Vocabulary Distributional Patterns: Corpus-based Approach 110
Marina Solnyshkina, Vladimir Ivanov and Valery Solovyev

Comparative Analysis of Classification Methods for Human Identification by gait 118
Lubov Shiripova and Evgeny Myasnikov

Voronoi-based Path Planning based on Visibility and Kill/death Ratio Tactical Component 129
Ilya Makarov, Pavel Polyakov and Roman Karpichev

Security System Event Log Analysis 141
Dmitry Chalyy, Nikolai Ovchenkov, Ekaterina Lazareva and Rafael Yaikov

Evaluation of Nonlinear Dimensionality Reduction Techniques for Classification of Hyperspectral Images 147
Evgeny Myasnikov

Fourty Years of Network Science: Analysis of Journal Contribution to the Field 155
Marina Kalinina, Vladimir Kuznetsov and Valentina Kuskova

Image and Video Tag Aggregation 161
Olga Kanishcheva and Natalia Sharonova

Text Scene Detection with Transfer Learning in Price Detection Task 173
Vladimir Fomenko, Dmitry Botov and Julius Klenin

Verification of the Formal Approach to Data Fusion 179
Ildar Baimuratov, Stefan Morozov and Nataly Zhukova

Sentiment Analysis of Short Texts from Social Networks Using Sentiment Dictionaries and Blending of Machine Learning Algorithms . . . 190
Anastasia Novikova and Sergey Stupnikov

Probability Analysis of the Vocabulary Size Dynamics Using Google Books Ngram Corpus 202
Yulia Maslennikova, Anastasia Pekina and Vladimir Bochkarev

Neural Networks with Attention for Word Sense Induction 208
Oleg Struyanskiy and Nikolay Arefyev

Combined Machine-Learning Approach to PoS-Tagging of Middle English and Old Norse Texts 214
Raoul Karimov, Andrei Akinin and Dmytro Yakymets

Text Normalization and Spelling Correction In Kazakh Language 221
Gaukhar Slamova and Meruyert Mukhanova

Static Gestures Classification Using Convolutional Neural Networks on the Example of the Russian Sign Language 229
Oleg Potkin and Andrey Philippovich

Evaluation of Methods for Cell Nuclear Structure Analysis from Microscopy Data 235
Alexandr A. Kalinin, Brian D. Athey and Ivo D. Dinov

Reconstructing an Artificial Society on the Basis of Big Open Data 241
Aleksandra L. Mashkova