Contribution ID: **25**   Type: **not specified**

Networked data-science for research, academic communities and beyond

Tuesday 18 September 2018 15:30 (30 minutes)

Dr. Andrey Ustyuzhanin, Head of Yandex-CERN Joint Research Projects & Head of the Laboratory of Methods for Big Data Analysis at National Research University Higher School of Economics

Abstract: There is an exceptional way of doing data-driven research employing networked community. The following examples can illustrate the approach: Galaxy Zoo or Tim Gower's blog. However many cases of collaboration with the data-science community within competitions organised on Kaggle or Coda Lab platforms usually get limited by restrictions on those platforms. Common Machine Learning quality metrics do not necessarily correspond to the original research goal. Constraints imposed by the problem statement typically look artificial for ML-community. Preparing a perfect competition takes a considerable amount of efforts. On the contrary research process requires a lot of flexibility and ability to look at the problem from different angles. I'll describe the alternative research collaboration process can bridge the gap between domain-specific research and data science community. Notably, it can involve academic researchers, younger practitioners and all enthusiasts who are willing to contribute. Such research process can be supported by an open computational platform that will be described along with essential examples for the audience of the workshop.

Biography: Dr Andrey Ustyuzhanin - the head of Yandex-CERN joint research projects as well as the head of the Laboratory of Methods for Big Data Analysis at NRU HSE. His team is the member of frontier research international collaborations: LHCb - collaboration at Large Hadron Collider, SHiP (Search for Hidden Particles) - the experiment is designed for the New Physics discovery. His group is unique for both collaborations since the majority of the team members are coming from the Computer and Data Science worlds. The primary priority of his research is the design of new Machine Learning methods and using them to solve tough scientific enigmas thus improving the fundamental understanding of our world. Amongst the project he has been working on are efficiency improvement of online triggers at LHCb, speed up BDT-based online processing formula, the design of custom convolutional neural networks for processing tracks of muon-like particles on smartphone cameras. Development of the algorithm for tracking in scintillators optical fibre detectors and emulsion cloud chambers. Those project aid research at various experiments: LHCb, OPERA, SHiP and CRAY-FIS. Discovering the deeper truth about the Universe by applying data analysis methods is the primary source of inspiration in Andrey's lifelong journey. Andrey is a co-author of the course on the Machine Learning aimed at solving Particle Physics challenges at Coursera and organiser of the annual international summer schools following the similar set of topics. Andrey has graduated from Moscow Institute of Physics and Technology in 2000 and received PhD in 2017 at Institute of System Programming Russian Academy of Sciences.

Presenter: Dr USTYUZHANIN, Andrey

Session Classification: Day 2