21st International Conference on Computing in High Energy and Nuclear Physics **CHEP2015** Okinawa Japan: April 13 - 17, 2015

Contribution ID: **395**   Type: **poster presentation**

Towards Reproducible Experiment Platform

Data analysis in fundamental sciences nowadays is essential process that pushes frontiers of our knowledge and leads to new discoveries. At the same time we can see that complexity of those analysis increases exponentially due to a) enormous volumes of datasets being analyzed, b) variety of techniques and algorithms one have to check inside a single analysis, c) distributed nature of research teams that requires special communication media for knowledge and information exchange between individual researchers. There is a lot of resemblance between techniques and problems arising in the areas of industrial information retrieval and particle physics. To address those problems we propose a Reproducible Experiment Platform (REP) - a software infrastructure to support a collaborative ecosystem for computational science. It is a Python-based solution for research teams that allows running computational experiments on big shared datasets, obtaining reproducible and repeatable results, and consistent comparisons of the obtained results. REP supports many data formats including ROOT, allowing for easy integration with existing HEP software and analyses. We present some key features of REP based on case studies which include trigger optimization and physics analysis studies at the LHCb experiment, as well as an example case of applying the prototype of such a system in Information Retrieval research that led to a performance increase of two orders of magnitude.

Primary authors: BARANOV, Alexander (ITEP Institute for Theoretical and Experimental Physics (RU)); ARTEMOV, Alexey (Yandex School of Data Analysis); ROGOZHNIKOV, Alexey (Yandex School of Data Analysis, Moscow); USTYUZHANIN, Andrey (ITEP Institute for Theoretical and Experimental Physics (RU)); KHAIRULLIN, Egor (Yandex School of Data Analysis, Moscow); LIKHOMANENKO, Tatiana (National Research Centre Kurchatov Institute (RU))

Co-authors: WILLIAMS, J Michael (Massachusetts Inst. of Technology (US)); GLIGOROV, Vladimir (CERN)

Presenter: USTYUZHANIN, Andrey (ITEP Institute for Theoretical and Experimental Physics (RU))

Track Classification: Track6: Facilities, Infrastructure, Network