

## Data intensive Science in High Energy Physics

The new challenging era of scientific data management in the coming decade of “Big Data” requires a new paradigm named data intensive science to deal with petabyte data scale in many modern social, economic and scientific areas. Experimental High Energy Physics, in particular, demands two main issues: (1) the giant complexes for distributed computing and corresponding grid-cloud internet services as the [Worldwide LHC Computing Grid](#) (WLCG); (2) Machine Learning (ML) approaches to search for unevident regularities in data for getting the most probable forecast of phenomena under study. The first issue is considered on examples of the grid/cloud systems developed at the Laboratory of Information Technologies (JINR, Dubna) and focused on improving their development efficiency. For this purpose a simulation program is proposed which is designed to use work quality indicators of some real system to design and predict its evolution. The second issue is also considered on examples of such ML applications in the JINR practice as artificial neural networks, two stage clustering and some others.