TeRAbytes

Contribution ID: **42**   Type: **Poster**

First LHCb open data release and analysis preservation techniques

LHCb and other experiments at CERN have the commitment to make the collected data available to the public. In December of 2022, LHCb has released the first data set containing 200 Terabytes of data from proton-proton collisions collected in the years 2011 and 2012 (Run1 of the LHC). This data is now available through CERN Open Data Portal. The aim of the open data release is to encourage particle physics research by third parties and aid in building the legacy of detector experiments at CERN. The report on this initial release is provided in the poster. It has become clear that it is important to preserve not only the data or the final results of physics analyses but also the steps taken to obtain these results. These steps include everything from data selection to statistical methods used in the analysis. It is also very important to preserve the computational environment which was used for the analysis, so, at a later date analysis could be rerun from start to finish for cross-checks or when bigger data samples become available. One framework to help with the preservation of the analysis techniques is REANA reproducible analysis platform which will also be briefly introduced in the poster.

Primary author: SARPIS, Mindaugas (Helmholtz-Institut für Strahlen- und Kernphysik)

Presenter: SARPIS, Mindaugas (Helmholtz-Institut für Strahlen- und Kernphysik)

Session Classification: Posters