Contribution ID: **219**   Type: **presentation**

Improving the use of data quality metadata via a partnership of technologies and resources between the CMS experiment at CERN and industry

Tuesday, 10 July 2018 12:15 (15 minutes)

The CMS experiment dedicates a significant effort to supervise the quality of its data, online and offline. A real-time data quality (DQ) monitoring is in place to spot and diagnose problems as promptly as possible to avoid data loss. The evaluation a posteriori of processed data is designed to categorize the data in term of their usability for physics analysis. These activities produce DQ metadata.

The DQ evaluation relies on a visual inspection of monitoring features. This practice has a high cost in term of human resources and is naturally subject to human arbitration. Potential limitations are linked to the ability to spot a problem within the overwhelming number of quantities to monitor, or to the understanding of detector evolving conditions.

In view of Run III, CMS aims at integrating deep learning technique in the online workflow to promptly recognize and identify anomalies and improve DQ metadata precision.

The CMS experiment engaged in a partnership with IBM with the objective to support, with automatization, the online operations and to generate benchmarking technological results. The research goals, agreed within the CERN Openlab framework, how they matured in a demonstration application and how they are achieved, through a collaborative contribution of technologies and resources, will be presented.

Primary authors: POL, Adrian Alan (Université Paris-Saclay (FR)); JESSOP, Colin (University of Notre Dame (US)); CERMINARA, Gianluca (CERN); VLIMANT, Jean-Roch (California Institute of Technology (US)); PIERINI, Maurizio (CERN); ANDREWS, Michael (Carnegie-Mellon University (US)); DEV, Nabarun (University of Notre Dame (US)); MARINELLI, Nancy (University of Notre Dame (US)); MUDHOLKAR, Tanmay (Carnegie-Mellon University (US)); AZZOLINI, Virginia (Massachusetts Inst. of Technology (US))

Presenter: AZZOLINI, Virginia (Massachusetts Inst. of Technology (US))

Session Classification: T1 - Online computing

Track Classification: Track 1 - Online computing