



Contribution ID: 370

Type: **Poster**

## An Extensible Infrastructure for Querying and Mining Event-level Metadata in ATLAS

*Thursday 24 May 2012 13:30 (4h 45m)*

The ATLAS event-level metadata infrastructure supports applications that range from data quality monitoring, anomaly detection, and fast physics monitoring to event-level selection and navigation to file-resident event data at any processing stage, from raw through analysis object data, in globally distributed analysis. A central component of the infrastructure is a distributed TAG database, which contains event-level metadata records for all ATLAS events, real and simulated.

This resource offers a unique global view of ATLAS data, and provides an opportunity, not only for stream-style mining of event data,

but also for an examination of data across streams, across runs, and across (re)processings.

The TAG database serves as a natural locus for run-level and processing-level integrity checks, for investigations of event duplication and other issues in the trigger and offline systems, for questions about stream overlap, for queries about interesting but out-of-stream events, for statistics, and more. In early ATLAS running, such database queries were largely ad hoc, and were handled manually.

In this paper, we describe an extensible infrastructure for addressing these and other use cases during upload and post-upload processing, and discuss some of the uses to which this infrastructure has been applied.

**Author:** ATLAS, Collaboration (Atlas)

**Co-authors:** Dr MALON, David (Argonne National Laboratory (US)); VINEK, Elisabeth (University of Vienna (AT)); Dr CRANSHAW, Jack (Argonne National Laboratory (US)); ZHANG, Qizhi (Argonne National Laboratory (US))

**Presenter:** Dr CRANSHAW, Jack (Argonne National Laboratory (US))

**Session Classification:** Poster Session

**Track Classification:** Software Engineering, Data Stores and Databases (track 5)