



Contribution ID: 366

Type: **Poster**

## File and Dataset Metadata Collection and Use in Atlas

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

The ATLAS Metadata Interface (“AMI”) was designed as a generic cataloguing system, and as such it has found many uses in the experiment including software release management, tracking of reconstructed event sizes and control of dataset nomenclature. In this paper we will discuss the primary use of AMI which is to provide a catalogue of datasets (file collections) which is searchable using physics criteria.

The AMI dataset catalogues are filled from several sources:

- The Tier 0 database for raw data and first pass reconstruction.
- The Production System database for Monte Carlo and reprocessed data.
- The Distributed Data Management system.
- Direct input from the physicist community.

We will summarize the information taken from each source, and discuss the different mechanisms used to obtain it.

By correlating information from different sources we can derive aggregate information which is important for physics analysis; for example the total number of events contained in dataset, and possible reasons for missing events such as a lost file.

Finally we will describe some specialized interfaces which were developed for the Data Preparation and reprocessing coordinators. These interfaces manipulate information from both the dataset domain held in AMI, and the run-indexed information held in the ATLAS COMA application (Conditions and Configuration Metadata).

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

**Co-authors:** GALLAS, Elizabeth (University of Oxford (GB)); LAMBERT, Fabian (Universite Joseph Fourier (FR)); FULACHIER, Jerome (Universite Joseph Fourier (FR)); Dr ALBRAND, Solveig (Universite Joseph Fourier (FR))

**Presenter:** GALLAS, Elizabeth (University of Oxford (GB))

**Session Classification:** Poster Session

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