

Data Management Plan Pierre Auger Observatory

Experiment description:

The Pierre Auger Observatory, located on a high plain near Malargüe in western Argentina, is the world's largest cosmic ray observatory. The objectives of the Observatory are to probe the origin and characteristics of cosmic rays above 10^{17} eV and to study the interactions of them, the most energetic particles observed in nature, with the Earth's atmosphere. The Auger Observatory features an array of 1660 water Cherenkov particle detector stations spread over 3000 km² over-looked by 24 air fluorescence telescopes. In addition, three high elevation fluorescence telescopes overlook a 23.5km², detector infill array. Construction of the full array was begun in 2003 and completed in 2008. The Observatory is currently in data taking mode.

DOE's roles in the experiment:

Fermilab hosted the Auger Project Office until 2014. Fermilab managed DOE construction funds for seven US collaborating institutions. DOE and NSF have shared Auger Observatory operating funds equally.

Partnerships:

An international agreement is entitled: *Agreement for the Organization, Management and Funding of the Pierre Auger Observatory* that involves about 90 institutions from 17 countries.

http://www.auger.org/admin/FinanceBoard/auger_fb_agreement_AUG2008_revised.pdf

In the US DOE and NSF have shared Auger construction and operations costs equally corresponding to about 20% of the costs. Private funds from the US were contributed to the construction.

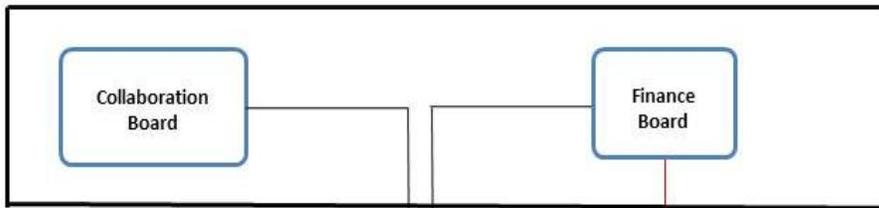
International funding partners may be found at
<http://www.auger.org/admin/Collaborators/agencies.html>

Organization

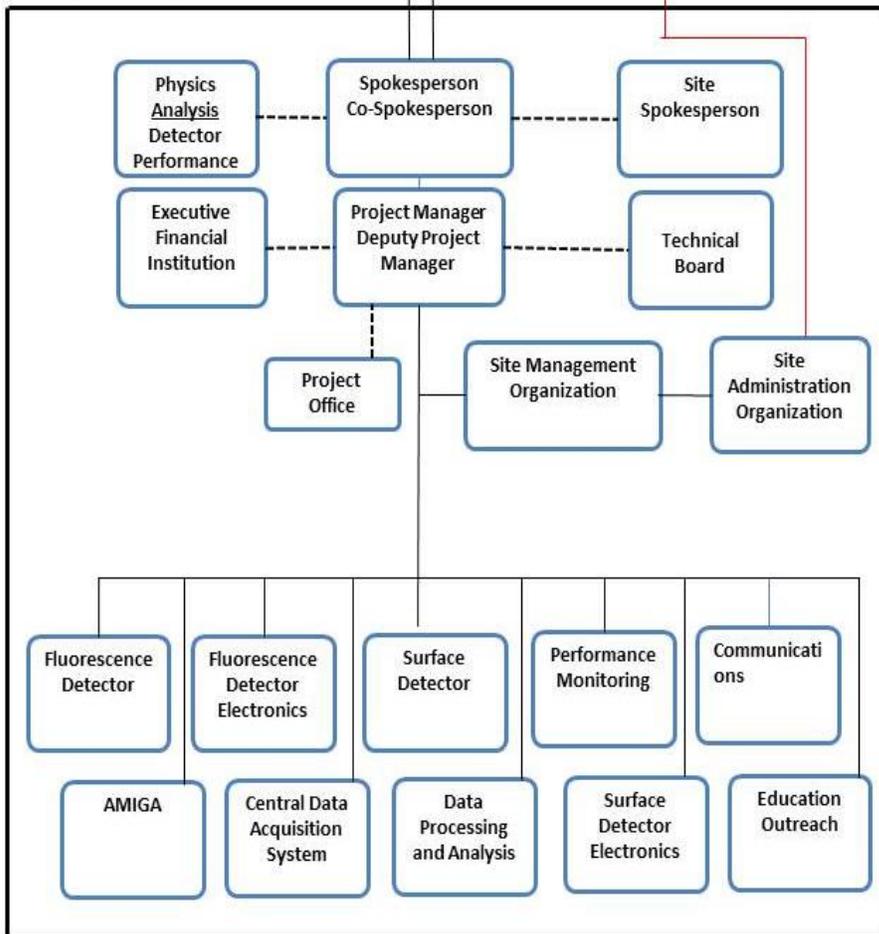
Fermilab was host to the Auger Project Office and overall project management from 1998 until 2014 when the Project Office was transferred to the Karlsruhe Institute of Technology.

Organization – Experiment level

Project Oversight



Project Organization



The organization is detailed in the Project Management Plan referenced above.

Collaboration:

The collaboration consists of 17 countries, 90 institutions, ~ 450 scientists

The description of the collaboration and its governance can be found in the Auger Project Management Plan:

http://www.auger.org/admin/Management/PMP/Project_Management_Plan_Rev_31MAY13_with_track_changes.pdf

Data policy management:

The Auger Collaboration Board is responsible for hearing and discussing proposals related to Collaboration policy. As such, this body will determine the data release and archiving policy. A Data Release Task Force has recently been established which will bear the responsibility of implementing the policy determined by the Collaboration Board, and which will report to the Project Management on progress, problems and manpower requirements.

Data Description & Processing

Description:

We anticipate making three different data streams available. The first includes environmental data recorded by the Auger atmospheric monitoring program. Such data may include temperature, pressure, cloud coverage and aerosol contamination as a function of time. A second data stream will be provided by the surface detector scaler rates, which are sensitive to both atmospheric conditions and extraterrestrial effects including solar activity. These two streams can be made publicly available with a short latency period, given sufficient manpower. The third data stream will comprise a subset of reconstructed data.

The public data will be provided as ASCII files, or in some other common format, with a set of example tools for reading and processing the data. The Data Release Task Force holds the responsibility for determining the best options for data formatting and for providing a portal through which the community can access the public data.

Raw data will not be released, as significant expertise is required to convert this into reconstructed events. Within the collaboration, the raw data and various monitoring streams are merged for processing and production of Advanced Data Summary Trees, which are used for final analysis within the Collaboration.

The data for public release will be made available on dedicated web portal.

Data Products and Releases:

At the present time, 1% of reconstructed events are released to the public with short latency. In the future, 10% of the reconstructed event sample will be made publicly available after a 2 year embargo. This subsample can be used for outreach purposes and as “training data” for the community at large to develop analysis ideas. In case of interesting discoveries in the training sample, the Auger Collaboration will assist outside users in further analysis using the full Auger data sample. This approach is similar to that employed internally by the collaboration, in which a

subset of the data sample is used for exploratory purposes, and in case of an interesting observation, the full sample is subsequently un-blinded.

Plan for Serving Data to the Collaboration and Community:

The infrastructure for serving data to the Collaboration is already mature. Data centers and mirrors are well established, and have been providing event and monitoring data to the Collaboration since the first physics results were published in 2004.

The Data Release Task Force will develop a web portal from which the public data sample and associated tools can be downloaded. Collaboration policy and contacts will also be provided via this portal. The details of the tools provided will be tuned according to community demand and available manpower.

Plan for Archiving Data:

The details of data archiving are currently left open for future discussion, given that funding sources for long-term data storage and tool maintenance have not yet been identified. It is worth noting, however, that the Auger data sample is relatively small compared to a High Energy Physics experiment; the full sample currently consumes 28 Tb and we do not expect to require much more than about 60 Tb by the end of the Observatory lifetime.

Plan for Making Data Used in Publications Available:

Data points from published plots will be made available on the Data Release Task Force web portal, so that community members need not attempt to estimate values by reading plots. Where appropriate, event lists used in published sky maps will also be provided via the portal.

Responsiveness to SC Statement on Digital Data Management

This data management plan fully follows SC Statement on Digital Data Management to the best of our understanding.