

## **Data Management Plan for the MINERvA Experiment (FNAL-E938)**

Results from the MINERvA data are posted to the usual archive for high energy physics results (arXiv.org), and are freely available. All dissertations at the are made available online through the Fermilab Publications Office and are publicly available.

MINERvA experiment analyses produces event summary information, summary ntuples, plots/fits and other data that describe the different parts of the final analysis chain. Many of these samples hide the specialized knowledge/characterization of the detectors and instead provide quantities relating to the underlying physics that can be directly used and compared to theory and other experiments. The highest level of these data is often formally published as part of peer review articles or made available as public datasets.

Scientific results data will be cataloged and archived according to the following policies:

- All data presented in graphical formats will be stored in its final published form in digital formats corresponding to the actual format(s) that were used to generate the publication(s) (i.e. if the plot was included in an article as a portal network graphic [PNG] then a PNG copy of that file will be retained. If in addition the graphic was used in a Portal Document Format [PDF] to generate a publication, then a PDF copy will also be retained.)
- The graphical data and their corresponding text representations will be cataloged together in the MINERvA Official plots database together with meta information about the graphic, including a detailed description appropriate for or corresponding to a caption included in a publication.
- All tabular data presented in publications will be stored in a file containing the formatted source of the table (i.e. a latex formatted table object appropriate for direct inclusion in a publication)
- All published data, graphs and tables will be stored by the collaboration on archival medium through the Fermilab mass storage system and will maintain a minimum of two copies in that system. Additional copies may be kept on other storage systems or at other locations.

The intention of this archival policy to ensure that results are preserved in a form that can be used without proprietary software or reliance on MINERvA's intellectual property.

In addition the following items relating to the published data will be cataloged and archived according to the following policies:

- The analysis code, algorithms, scripts and procedures used to generate any published results will be cataloged and indexed in such a way that it is associated with the particular result or publication that it generated.
- These code bases along with their supporting infrastructure will be collected and packaged into a fully redistributable digital form (example: standard Unix tape archive).
- The redistributable form of the code bases will be stored by the collaboration on archival medium through the Fermilab mass storage system and will maintain a

minimum of two copies in that system. Additional copies may be kept on other storage systems or at other locations.

MINERvA uses centrally managed repositories at Fermilab to archive this software. The intention of this archival policy for the code base is to ensure that method by which the published data were derived is fully preserved and that the exact methods and algorithms could be used at a later point in time to verify past, present or future data against the published data.

Data belonging to scientific results for the MINERvA experiment will be retained and supported for the active life of the experimental collaboration. Results data which is the direct output of a published scientific or technical result, will be retained past the dissolution of the MINERvA collaboration at a level corresponding to at least an archival form of the data along with the supporting analysis code base and procedures described previously. The core archival copies of the data will reside at Fermilab in the mass storage systems provided by the lab and will be subject to the published data retention and media migration policies that the lab maintains. Data will be digitally archived using Fermilab facilities following the policies described in [http://computing.fnal.gov/xms/Science\\_&\\_Computing/Policies\\_and\\_Publications](http://computing.fnal.gov/xms/Science_&_Computing/Policies_and_Publications).

Data from scientific results of the MINERvA experiment will be hosted primarily by the Fermilab computing and archive facilities. Replicas of any portions of the scientific results tier (plots, tables, binary data, analysis codes) can be disseminated to collaborating institutions via the standard replication tools and authentication services provided by the Fermilab scientific computing division. This replication can be initiated by members of the MINERvA virtual organization (VO) or by request to the Fermilab computing division staff. In addition, unpublished documents which support the scientific analysis are available for dissemination through the Document Database Service (DocDB) hosted by Fermilab. The DocDB service provides both authenticated and unauthenticated access and permits the MINERvA collaboration to provide both fully public access as well as partially restricted access to these supporting documents.

Dissemination of data from the scientific results, beyond the published plots and data tables (example: individual event kinematics) to non-MINERvA collaboration parties will require approval of the MINERvA collaboration. In addition, due to the nature of the high level analyses that can be performed with these data, and the intellectual property rights of the MINERvA collaboration, these requests may be deferred for a pre-determined, waiting period after the initial publication of MINERvA results. This waiting period is designed to permit the completion of any analysis efforts internal to MINERvA that may be pending or planned for publication to complete (example: a joint fit combining MINERvA data with data published by another experiment). In these cases the data, after approvals, will be disseminated through the standard [authenticated] services.

Requests for MINERvA data from the scientific results that do not fall into the previously described categories will be subject to approval by the MINERvA collaboration and disseminated through the standard authenticated web services provided by Fermilab.

The MINERvA collaboration is currently developing a plan for archiving its complete data set, in a low level form suitable to reproduction of existing analyses or creation of new ones. This data, and access to documentation and code required to use it, will be distributed through the standard unauthenticated web services provided by Fermilab.