Summary of Neutrino Cross-Section Data Usage and Archival Mini-Workshop

https://indico.fnal.gov/event/45059/

NF06 Conveners: Baha Balantekin, Jason Newby, Jonathan Asaadi, Kendall Mahn

Workshop Summary:

- **From the workshop indico:** "The purpose of this workshop is to discuss how neutrino cross section measurement data is used (data usage and data releases) and how it can be reanalyzed in the future (data archival). We will also present ideas for the subsequent mini workshops for feedback."
  - 43 participants total (34 on day 1, 23 on day 2)
- **Presentations from people who have done**
  - Archaeology analyses: Callum Wilkinson (LBNL)
  - Use historic x-sections in generators: Julia Tena Vedal (Liverpool)
  - Current efforts to archive MINERvA: Rob Fine (Rochester)
- **Notes from the meeting can be found here**

Some takeaways from the workshop: Data Archeology

- **When data is released, correlation matrices are essential!**
  - Need to be included either in papers or central repositories
- **Efficiency corrections and limitations of the simulation should be acknowledged in the published work**
  - Independent MC package/tune/versions should be pushed through analyses
- **When unfolding model comparisons should be carried out in both reco and true spaces**
  - Some usefulness in showing data which hasn't been unfolded

There is a lot of value in the ability to do data archeology, but it requires incentives to both the analyzers and the experiments to perform additional work and to provide continuity in conversation (who do you talk to when the analyzers have "moved on")

Ideas put forward on Data Archaeology

- **A neutrino scattering center (similar to what comes from nuclear physics) could be established to provide both "incentives" (e.g. funding and jobs) to provide do the work for longevity of measurements and to provide "resources" (e.g. computing and disk space) to make data useful long into the future**
  - There needs to be DOE support to make something like this happen
  - A lot of work has been done on the LHC which can provide a useful model for standards and expectations
- **Having access to data which is "uncontaminated" (e.g. no generator models or hard to track down corrections applied) is helpful for long term use**
  - Lots of practical difficulties in publishing and using this data, but should be looked into
- **Community standards and expectations need also to be in place to incentivize the analyzers (and their advisors) to do the hard work of making measurements have long-term usefulness**
  - Might need more than what is currently in our "Data Management" Plans