

Minutes of Kick-Off Meeting of HepData Advisory Board
IPPP, 12.11.2014

- present: Sunje Dallmeier-Tiessen (Inspire, CERN), Bill Murray (ATLAS, Warwick), Jeppe Andersen, Peter Richardson, Graeme Watt, Mike Whalley (all IPPP).
through Vidyó: Kyle Cramner (NYU), Ulrik Egede (LHCB, Imperial), Henning Flaecher (CMS, Bristol), Cesar Sanchez (ALICE, CERN)
apologies: Salvatore Mele (Inspire, CERN), Matthew Wing (smaller experiments, UCL)
- talks can be found on the **Indico page**: <http://conference.ippp.dur.ac.uk/event/425/>
- mixed remarks during the morning talks:
 - HepData Status & Future (Whalley/Watt):
 - downloading the data because of the doi's is becoming obsolete – not valid any more (Cramner)
 - agreed: data reviews as important part of the HepData remit, to contextualise data (all); Inspire also stresses important role of trained particle physicists in HepData (Dallmeier-Tiessen)
 - “data” tab on Inspire because of integration with other sources of data – dominant fraction from HepData so far (Cramner, Dallmeier-Tiessen)
 - better integration with Inspire good for long-term future (Cramner)
 - export formats: suggestion for ROOT output to improve, also integrate with ROOT as allowed input format: allow ROOT (Cramner)
→ plan is for IPPP HepData team to come up with proposal to discuss with Advisory Board
 - Inspire (Dallmeier-Tiessen):
 - references to individual datasets may cause problems with citation counts (Andersen)
 - duplication of effort with Analysis Preservation Framework (APF)? (Andersen) – not quite, HepData, even with Rivet etc. would be one item, the public end, whereas the APF is more internal to the experiments (Dallmeier-Tiessen)
 - HepData Funding Strategy (Krauss):
 - plans for funding application: perfect the web interface, enhance searchability, integrate with Invenio
 - also allow other analysis code, be not exclusive to RIVET (Cramner)
 - transition to Invenio welcomed (Cramner)
 - doi's for data as a contractual question (Egede)
- from experimental talks:
 - generic: for ROOT format, there is no collaboration wide standard → need instructions from HepData how they want to include ROOT → should/could be based on very few ROOT classes → 1D/2D plots (TH1, TH2), TGRAPH etc.
 - ALICE (Sanchez):
 - happy with text format for submission, eager to try out web-interface
 - ATLAS (Murray):

- long-term storage policy: reprocess data with better tools, driven by anticipated development of computer technology
 - errors in HepData data reported with a frequency of 1/month
 - first attempt of upload through web interface apparently successful
- CMS (Flaecher):
 - uploading data part of CMS publication policy – well used by SM measurements, less by searches; responsibility up to now in physics groups
 - plan to include all future data, attempt to also include past datasets
 - first tests of direct uploads under way, quite a load of manual interaction still needed → more input formats would be desirable
 - versioning of inputs to track corrections/additions
- LHCb (Egede):
 - so far cross section data only, fairly time-consuming due to input format and lack of standardisation within LHCb
 - some fear of duplication with CDS/Inspire
 - including figures like Dalitz plots or complicated structures like large dimension correlation matrices very welcome
- Action items for IPPP group:
 - document ROOT → HepData scripts for translating ROOT into HepData format; and turn this into a proposal for accepted ROOT formats
 - add versioning of entries to track corrections/additions
- Integration with Invenio/Inspire
 - added benefit for longevity and automated migration to other formats
 - possible danger: maintain independence
 - decision of where to host the Invenio – some dangers (Richardson, Watt)
 - closer integration with Invenio would be very good, one-stop shop (Egede)
 - important bit is the back-end → searchability, user interface, etc.
- Strategy for renewed funding:
 - main priorities:
 - priority 1: enhance input system (improve input web-interface), add new datatypes
 - priority 2: improve searchability of database
 - both tie in nicely with migration to Invenio/integration with Inspire
 - further items:
 - doi's for additional data
 - forward link to analysis software where applicable
- Next meeting: after STFC results in summer