ICFA Panel Discussion

HEP Data Ownership and Usage

Important note: All my remarks are my <u>private opinion</u>

No serious discussion has yet taken place within the ATLAS Collaboration

Summary: I am not *a priori* against sharing data with the full community

However there are good reasons that this should only be done after a certain time

I am also rather sceptical that complete outsiders could in practice fully understand the data

Incentive and motivation for the detector construction

People as individuals and as teams in the collaborations make extreme efforts during the long phases of conception, prototyping, design and construction of the detectors

Creative ideas and technologies are developed to the limits in order to arrive at an optimum instrument, within cost constraints, for the physics

By large people are _driven by their motivation for the physics_ to come, they want to maximize their chance to take part in the front-line physics

Without this prospect of getting the physics first I am convinced that a much less broad effort would go into detector developments, and the field would lose some of the physics potential

(Detectors conceived by instrumentation experts only would not be optimal for the physics …)

Therefore a significant 'first-use time' for the data should be granted to the collaborations

How long?
- Equivalent to the construction time? (order 6-8 years)
- Typical phase of data sets at colliders before upgrades? (hadron collider order 5 years)

Would it be possible for an outsider to use the data?

I am rather sceptical that outsiders could in practice fully understand the data

     - All the data selection chains (trigger and event filter) at a hadron collider are of crucial importance, and need to be understood for any physics analysis

     (In general one cannot 'observe' passively an unbiased sample)

     - Physics analyses imply often very specific and topics-dedicated large Monte Carlo simulation and calibration efforts

     - Software and analysis tools improve over the years based on the collective experience in the collaboration

All this knowledge and the software tools must of course be kept available and documented within the collaborations

However I think it would be a *huge further effort* to maintain this at a level such that outsiders can use it independently

My experience is that a lot of the knowledge, checks, critical evaluation of the analyses, come from the interactions with the community having built and operated the detector

Some suggestions

A much more _realistic way to share our data_ would be in my mind to make sure that people will have also in the future a chance to work with (or join) the collaborations

- guest scientist in one of the collaborating institutes for a given topic
- the new publication channel of 'Scientific Notes' could be used to document individual ideas in addition to the final publication with the full collaboration

Nevertheless data could be on 'free access' after several years, but then publications based on them should be clearly marked as such, and journals should take a strong commitment for proper refereeing

It would have to be understood that the responsibility of these publications is fully with the 'outside user'

Finally, I am all for having the collaborations making available _simplified event samples for educational purposes!_