

Data Preservation Challenges For High Energy and Astro Physics...

Mike Hildreth, University of Notre Dame
Elisabeth Long, University of Chicago
Rob Roser, Ruth Pordes, Fermilab

Data and Software

- ✱ 100s of PB of data and 10s million lines of code:
 - ✱ Irreplaceable resource
 - ✱ Re-use necessary when find new signals and to compare/validate new results.
- ✱ DPHEP Working Group
 - ✱ Convened by International Committee on Future Accelerators (ICFA)
 - ✱ ~ 100 members from different HEP Experiments and Facilities
 - ✱ Two Reports:
 - ✱ DPHEP-2009-00, <http://arxiv.org/pdf/0912.0255>
 - ✱ DPHEP-2012-01, May 2012, [arXiv:1205.4667v1](http://arxiv.org/abs/1205.4667v1)
 - ✱ Conclusions:
 - ✱ “an urgent and vigorous action is needed to ensure data preservation in HEP”
 - ✱ “A clear and internationally coherent policy should be defined and implemented
- ✱ Several projects in the USA and Europe initiated to do

Goals of this BOF

- ✱ To tell you what we are thinking about
- ✱ To talk with you and learn from your experience
- ✱ To get some ideas of common needs and challenges

- ✱ Introducing efforts for the Fermilab Tevatron Experiments (CDF, D0)
- ✱ Giving experiences from preservation/archiving activities for Sloan Digital Sky Survey (SDSS)
- ✱ Describing goals of Data and Software Preservation for Open Science (DASPOS) project focusing initially on Large Hadron Collider ATLAS and CMS experiments
- ✱ Discussion

Summary

- ✿ **Data Preservation and Access and Software Preservation and Re-Use will be major issues**
 - ✿ Merely preserving the data for re-use within the experiments will be a major challenge
 - ✿ Preserving validated software systems will need ongoing focused activities
 - ✿ No technical infrastructure in place to handle public release, access to data

Additional Slide: Q's..

What other data preservation efforts do people have going?

Are there any levels of commonality that you think are worth shooting for?

Do you have some urls that give your plans?

What technologies could industrial partners bring to the table?

How do you see information can be useful after the initial research by people who did not collect the data originally?