

Dissemination of HEP Research Results

Patricia McBride, Doug Glenzinski
Fermilab
24-June-2011

Outline

- Introduction
- Charge
- Report Summary

Introduction

- **COMPETES Reauthorization Act of 2010 highlights the importance of public access to the results of federally funded research programs**
 - Scholarly publications
 - Digital data
- **DOE Office of Science wants to assess the current policies and practices of their research programs**
 - Charged HEPAP to supply information for OHEP funded research

Introduction

- **HEPAP Sub-committee formed**
Marina Artuso, Andrew Cohen, Lance Dixon,
Doug Glenzinski (chair), Patricia McBride,
Ian Shipsey
 - Worked from Mar-May, met a few times,
surveyed practices, interviewed some of
the stakeholders, drafted the report
 - Preliminary report circulated 20-May
 - Final report issued 08-June

Charge

- Full charge available here
<http://science.energy.gov/hep/hepap/reports/>
 - Summarized on following pages

Charge

- “Please submit to [Brinkman], no later than July 1, 2011, a report describing the current policies and practices for disseminating research results in the fields relevant to the High Energy Physics program.”
- “dissemination” refers to the circulation of research results outside of the originating collaboration

Charge

- “research results” refers to written findings (e.g. publications, conference proceedings or presentations) and digital data
- “practices” refers to accepted practices within a given discipline

Charge

- Things we're asked to consider
 - What are the criteria for dissemination and who makes this determination?
 - How is access provided/controlled?
 - Is access limited?
 - Does access come with additional functionality?
 - What's the official Version of Record and who stewards it?

Charge

- Things we're asked to consider
 - Is peer review a condition of dissemination?
 - What Institution or Body upholds the current policies (e.g. DOE, Host Institution, an established practice)?
 - In addition to dissemination, is long-term stewardship accounted for by the existing policy or practice?

Report Summary

- The dissemination practices of the experimental and theoretical HEP communities are quite different
 - Discussed separately
 - Within each community, a standard set of practices are widely employed

Report Summary

- Both experimentalists and theorists use the arXiv (<http://arXiv.org>) to disseminate their results
 - Open access electronic distribution system of (pre)publication material
 - Complemented by additional functionality provided by SPIRES/INSPIRE
 - Benefits from co-operation of Journals (and vice versa)
 - Together these facilitate and expedite the scientific dialog in HEP

Report Summary

- Both experimentalists and theorists publish in the same journals
 - e.g. Physical Review, Physics Letters, European Journal of HEP, Journal of HEP
 - All require anonymous peer-review by people outside originating collaboration
 - All are subscription based
 - Open access option available for a fee
 - Collaborate with arXiv and INSPIRE

Report Summary

- **Dissemination of Experimental Results**
Written Findings:
 - HEP experiments are mostly international collaborations
 - Dissemination practices reflect long standing practices in HEP
 - Analysis methodology and manuscript both subject to internal peer review
 - Details specified in collaboration By-Laws

Report Summary

- **Dissemination of Experimental Results**
Written Findings:
 - Once approved by collaboration, manuscript posted to the arXiv (and, often, to host institution's pre-print DB), and submitted to a peer-reviewed journal
 - Long term stewardship provided by journal and arXiv
 - Version of Record provided by journals and arXiv

Report Summary

- **Dissemination of Experimental Results**
Written Findings:
 - arXiv, INSPIRE, and the journals all provide additional functionality (e.g. literature, author, title or free text searches, links to references and citations, etc)
 - Occasionally the collaborations will provide more detailed information about the data set or result (e.g. table giving details of events or functional form obtained from a fit to data)

Report Summary

- **Dissemination of Experimental Results**
Digital Data:
 - HEP raw form data sets:
 - Typically quite large 0.1 – 10 PB
 - Require significant additional functionality in order to make use of them
 - Solving the related technical, governance, and support issues for dissemination outside the collaboration requires significant additional personnel and capital resources

Report Summary

- **Dissemination of Experimental Results**
Digital Data:
 - Historically, large scale access to the digital data has not been provided to people outside the originating collaboration
 - Many collaborations provide open access to a sub-set of their data for educational or outreach purposes. Limited additional functionality provided to visualize these data.

Report Summary

- **Dissemination of Experimental Results**
Digital Data:
 - Long term stewardship (>5y after the end of data taking) faces many of the same challenges as dissemination
 - IUPAP / ICFA study group, Data Preservation and Long Term Analysis in High Energy Physics (DPHEP), charged with understanding the challenges and studying proposed solutions
 - Standardized analysis tools designed to facilitate dissemination are in wide use

Report Summary

- **Dissemination of Theoretical Results**
Written Findings:
 - HEP theory paper are typically authored by groups of 6 or fewer people
 - Dissemination practices reflect long standing practices in HEP
 - Decision to disseminate is made by consensus among authors– ultimate responsibility lay with the group leader

Report Summary

- **Dissemination of Theoretical Results**
Written Findings:
 - manuscript posted to the arXiv, and usually to a peer-reviewed journal
 - Version of Record provided by journals and/or arXiv
 - Long term stewardship provided by journal and arXiv

Report Summary

- **Dissemination of Theoretical Results**
Digital Data:
 - Some work generates more than papers
 - Computer source code for Monte Carlo physics simulations (e.g. ALPGEN, PYTHIA)
 - Computer source code for lattice gauge theory calculations (e.g. USQCD, MILC)
 - Computer source code for cross section calculations (e.g. MCFM or MadGraph)
 - Global fits to a large corpus of data (e.g. CTEQ, ZFITTER, CKMFITTER)

Report Summary

- **Dissemination of Theoretical Results**
Digital Data:
 - Usually disseminated in an open access manner via the internet
 - Accompanied by a paper in a peer-reviewed journal describing functionality and/or specific results obtained with the code
 - Version of Record and long term stewardship provided by authors via web pages
 - HepForge offers a common repository used by many

Report Summary

- Dissemination practices reflect long established practices in HEP.
- Final Report delivered to the DOE Office of Science 8-June

- **Some additional comments:**

- The HEP dissemination model (arXiv, INSPIRE, journals) seems to be working well but there are ideas for evolution of tools such as INSPIRE to provide links to additional information (plots, tables, or even datasets).
- Addressing issues associated with dissemination of HEP digital data will require collaboration between the large (and small) collaborations, the host institutions, and the funding agencies.