Data Preservation in High Energy Physics: a collaborative perspective

Cristinel Diaconu, Ulrich Schwickerath
on behalf of the DPHEP collaboration

Outline

- Apropos data preservation
  - Why preserve data
  - DPHEP collaboration history
  - Principles, data kinds and **risks**, preservation levels, status
  - Relationship to Open data/ open science
- Highlights
  - from our last workshop in autumn 2024
- Lessons learned and conclusions

Quick intro about myself

- Until 2002, particle physicist at LEP, DELPHI experiment
- Working in CERN IT since 2005
- IT coordinator for scientific data preservation
- Member of the DPHEP collaboration
- About **24 years of history in DP**

http://dphep.org

Data preservation: why?

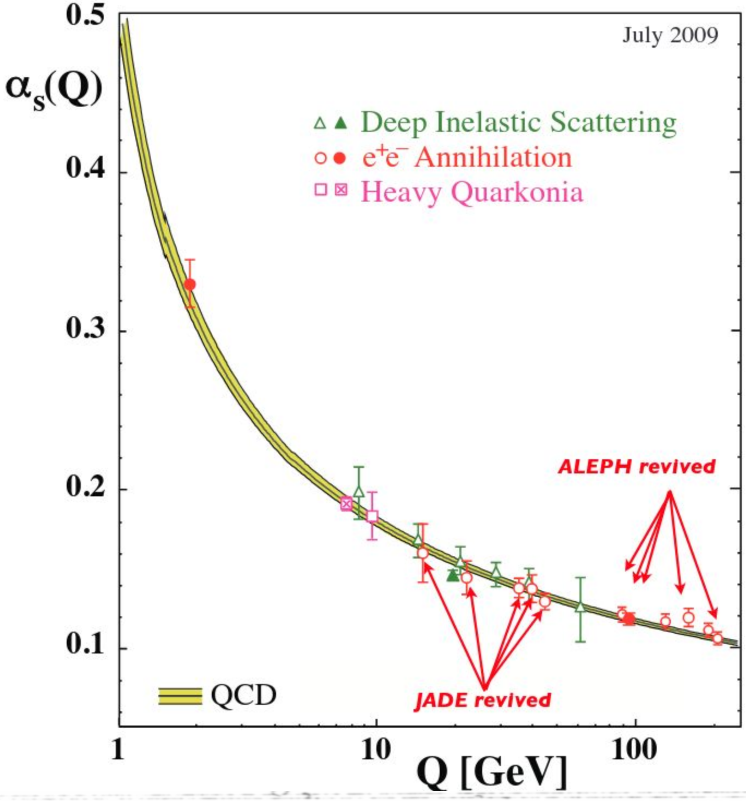
Example: the JADE experiment

- Experiment at Petra @ DESY
- **1978-1986**

Took about **8 years** to recover the data and the full software stack, **1995-2003**

Full story from Siggi Bethke (Talk at KEK)

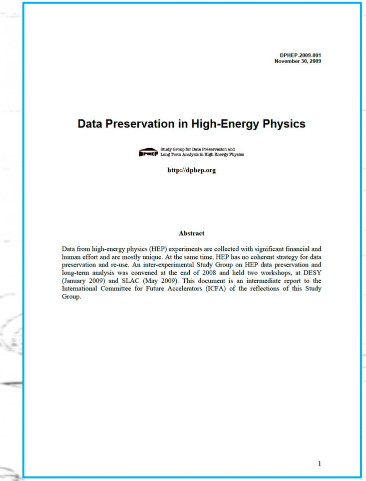
Proof of the energy dependence of the strong coupling constant



The DPHEP Collaboration

2009
Letter of intent

arXiv:0912.0255



DPHEP-2009-001
November 30, 2009

Data Preservation in High-Energy Physics

DPHEP Study Group for Data Preservation and Long Term Analysis in High Energy Physics

http://dphep.org

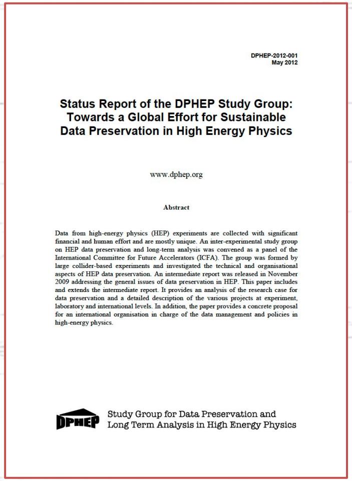
Abstract

Data from high-energy physics (HEP) experiments are collected with significant financial and human effort and are mostly unique. At the same time, HEP has no coherent strategy for data preservation and re-use. An inter-experimental Study Group on HEP data preservation and long-term analysis was convened at the end of 2008 and held two workshops, at DESY (January 2009) and SLAC (May 2009). This document is an intermediate report to the International Committee for Future Accelerators (ICFA) of the reflections of this Study Group.

1

2012
Blueprint

arXiv:1205.4667



DPHEP-2012-001
May 2012

Status Report of the DPHEP Study Group: Towards a Global Effort for Sustainable Data Preservation in High Energy Physics

www.dphep.org

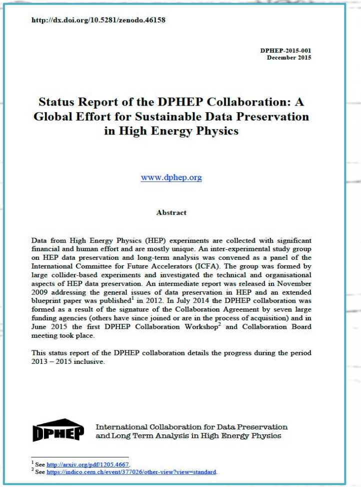
Abstract

Data from high-energy physics (HEP) experiments are collected with significant financial and human effort and are mostly unique. An inter-experimental study group on HEP data preservation and long-term analysis was convened as a panel of the International Committee for Future Accelerators (ICFA). The group was formed by large collider-based experiments and investigated the technical and organisational aspects of HEP data preservation. An intermediate report was released in November 2009 addressing the general issues of data preservation in HEP. This paper includes and extends the intermediate report. It provides an analysis of the research case for data preservation and a detailed description of the various projects at experiment, laboratory and international levels. In addition, the paper provides a concrete proposal for an international organisation in charge of the data management and policies in high-energy physics.

DPHEP Study Group for Data Preservation and Long Term Analysis in High Energy Physics

2015
Collaboration MoU

arXiv: 1512.02019



http://dx.doi.org/10.5281/zenodo.46158

DPHEP-2015-001
December 2015

Status Report of the DPHEP Collaboration: A Global Effort for Sustainable Data Preservation in High Energy Physics

www.dphep.org

Abstract

Data from High Energy Physics (HEP) experiments are collected with significant financial and human effort and are mostly unique. An inter-experimental study group on HEP data preservation and long-term analysis was convened as a panel of the International Committee for Future Accelerators (ICFA). The group was formed by large collider-based experiments and investigated the technical and organisational aspects of HEP data preservation. An intermediate report was released in November 2009 addressing the general issues of data preservation in HEP and an extended blueprint paper was published[1] in 2012. In July 2014 the DPHEP collaboration was formed as a result of the signature of the Collaboration Agreement by seven large funding agencies (others have since joined or are in the process of acquisition) and in June 2015 the first DPHEP Collaboration Workshop[2] and Collaboration Board meeting took place.

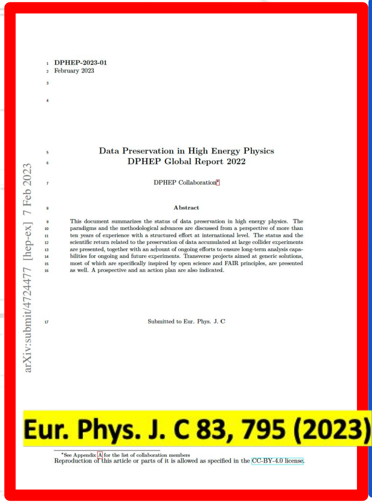
This status report of the DPHEP collaboration details the progress during the period 2013 – 2015 inclusive.

DPHEP International Collaboration for Data Preservation and Long Term Analysis in High Energy Physics

[1] See http://arxiv.org/pdf/1205.4667.
[2] See https://indico.cern.ch/event/377026/other-view?view=standard.

2023
Decade report

arXiv: 2302.03583



arXiv:submit/4724477 [hep-ex] 7 Feb 2023

DPHEP-2023-01
February 2023

Data Preservation in High Energy Physics
DPHEP Global Report 2022

DPHEP Collaboration*

Abstract

This document summarizes the status of data preservation in high energy physics. The paradigms and the methodological advances are discussed from a perspective of more than ten years of experience with a structured effort at international level. The status and the scientific return related to the preservation of data accumulated at large collider experiments are presented, together with an account of ongoing efforts to ensure long-term analysis capabilities for ongoing and future experiments. Transverse projects aimed at generic solutions, most of which are specifically inspired by open science and FAIR principles, are presented as well. A prospective and an action plan are also indicated.

Submitted to Eur. Phys. J. C

Eur. Phys. J. C 83, 795 (2023)

*See Appendix A for the list of collaboration members
Reproduction of this article or parts of it is allowed as specified in the CC-BY-4.0 license.

2025
EPPS 2026 submission

2503.23619 [hep-ex]

arXiv:2503.23619v1 [hep-ex] 30 Mar 2025

DPHEP-2025-01

Data Preservation in High Energy Physics

DPHEP Collaboration

Alexandre Arbey[11], Jamie Boyd[5], Daniel Britzger[19], Concetta Cartaro[24], Gang Chen[14], Gabor David[25], Dmitri Denisov[4], Cristinel Diaconu[6], Dirk Duellmann[5], Marcus Ebert[7], Eckhard Elsen[9], Jacopo Fanini[5], Dillon S. Fitzgerald[27], Benjamin Fuks[17], Gerardo Ganis[5], Achim Geiser[9], Takanori Hara[15], Lukas Heinrich[26], Michael D. Hildreth[28], Julie M. Hogan[3], Henry Klest[1], Sabine Kraml[18], Eric Lançon[4], Clemens Lange[21], Kati Lassila-Perini[10], Sergey Levonian[9], Dietrich Liko[2], Chiara Mariotti[13], Zach Marshall[16], Thomas McCauley[28], François Le Diberder[12], Jean-Yves Le Meur[5], Gerald Myatt[8], Maxim Potekhin[4], Michael Roney[7], Pablo Saiz[5], Heidi Schellman[22], Jose Benito Gonzalez Lopez[5], Matthias Schröder[5], Ulrich Schwickerath[5], Tim Smith[5], David South[9], Giordon Stark[23], Tibor Šimko[5], Jan Timmermans[20], Andrii Verbytskyi[19], Arne Wiebalck[5], and Zhiqing Zhang[12]

[1]Argonne National Laboratory, USA
[2]Austrian Academy of Sciences, Austria
[3]Bethel University, USA
[4]Brookhaven National Laboratory, USA
[5]CERN, Geneva, Switzerland
[6]CPPM, Aix Marseille Université, CNRS/IN2P3, France
[7]Department of Physics and Astronomy, Univ. of Victoria, Canada
[8]Department of Physics, University of Oxford, United Kingdom
[9]Deutsches Elektronen Synchorotron, DESY, Hamburg, Germany
[10]Helsinki Institute of Physics, Finland
[11]IP2I, Université Claude Bernard Lyon 1, CNRS/IN2P3, France
[12]IJCLab, Paris-Saclay University, CNRS/IN2P3, France
[13]INFN Torino, Italy
[14]Institute of High Energy Physics, China
[15]KEK IPNS / SOKENDAI, Japan
[16]Lawrence Berkeley National Laboratory, USA
[17]LPTHE, CNRS, Sorbonne Université, France
[18]LPSC, Université Grenoble-Alpes, CNRS/IN2P3, France
[19]Max Planck Institute for Physics, Münich, Germany
[20]Nikhef, Amsterdam, Netherland
[21]Paul Scherrer Institute, Switzerland
[22]Physics Department, Oregon State University, USA
[23]Santa Cruz Inst. for Particle Physics, Univ. of California, USA
[24]SLAC National Accelerator Laboratory, California, USA
[25]Stony Brook University, USA
[26]Technical University Munich, Germany
[27]University of Michigan, USA
[28]University of Notre Dame, Indiana, USA

April 1, 2025

Abstract

Data preservation significantly increases the scientific output of high-energy physics experiments during and after data acquisition. For new and ongoing experiments, the careful consideration of long-term data preservation in the experimental design contributes to improving computational efficiency and strengthening the scientific activity in HEP through Open Science methodologies. This contribution is based on 15 years of experience of the DPHEP collaboration in the field of data preservation and focuses on aspects relevant for the strategic programming of particle physics in Europe: the preparation of future programs using data sets preserved from previous similar experiments (e.g. HERA for EIC), and the use of LHC data long after the end of the data taking. The lessons learned from past collider experiments and recent developments open the way to a number of recommendations for the full exploitation of the investments made in large HEP experiments.[1]

1 Introduction and context

The need for data preservation demands long-term collaboration to support the continuing analysis program of complex data produced by large experiments. This was highlighted at the conclusion of operation of several large collider experiments at the beginning of this century, in the years before the LHC began. A study group was convened and produced in 2009 a report [1] that proposed a coordinated initiative in HEP, further detailed in a 2012 blueprint [2]. The document outlined initial data preservation projects and solidified the collaborative framework.

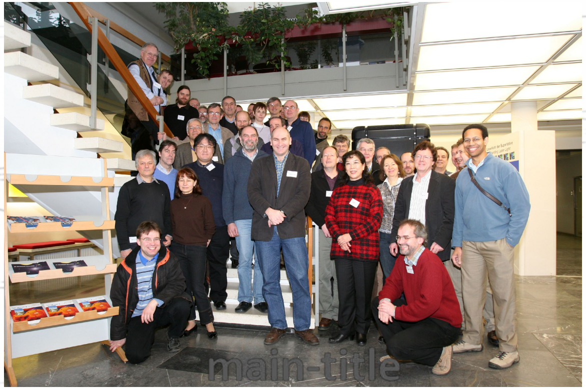
[1]This document has been submitted to the European Strategy for Particle Physics 2026.

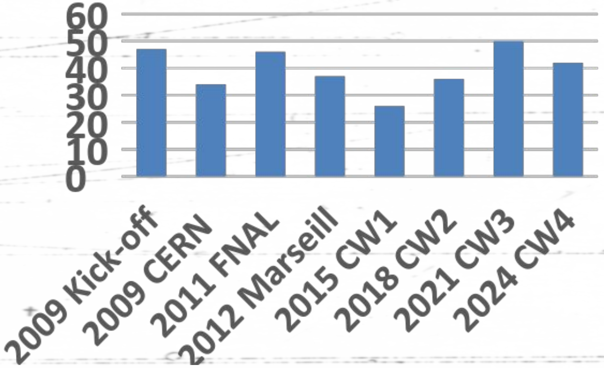
ICFA Panel 2024: DPHEP joined Data Lifecycle Panel

Workshops and activities

- DPHEP reports to the “International Committee for Future Accelerators” (ICFA)
- Regular workshops:
  - Every 2-3 years
  - 4th DPHEP workshop October 2024
  - Upcoming: 5th DPHEP workshop March 2026
- LEP and CERNLIB TF meetings

https://indico.cern.ch/category/4458/





What is data ?	What is preservation ?
NOT only “files”	NOT a freezer, herbarium, museum, album, cellar …
Every digitally encoded information that was created as a result of planning, running and exploiting an experiment	The **process** of transforming a “high intensity/rapidly changing” computing system into a low intensity/slowly evolving computing system, while conserving the capability to extract new science from the data
	Requires a clear plan and long term organisation ● Within each collaboration ● and at international level (DPHEP)

Areas of data DP

Bit - Preservation

Maintain and keep accessible:

- Raw data
- Reconstructed data
- Simulated data
- Meta data
- databases, conditions data

Software preservation

Includes:

- Analysis framework
- Reconstruction software
- Simulation software
- Visualisation

Note: Keep the sources (alive). Binaries may stop working

Documentation

- Published results
- Notes and internal documentation
- Technical documentation
- Manuals
- Web pages
- ….

Analysis preservation

- What has been done ?
- How has it been done ?
- Which data sets were used ?
- Analysis software sources

Ability to reproduce published results long term

Areas of DP: Risks

Bit - Preservation

- physical media failure (e.g. tape damage)
- system errors
- human errors

Mitigation: dual-tape, external copies

Software preservation

- Ever changing IT infrastructure
- Architecture and technology changes
- Storage systems access changes
- Security enhancements, e.g. deprecation of ciphers
- Compiler and computing language changes
- **Deprecation of required external dependencies**

Documentation

- Quality of internal documentation
  - May not be intended for general public
  - Use of jargon
- Author agreements, copyright
- Paper only copies with bad quality
- Document format long term support

Mitigation: QC from the start of the experiment, digitization

Analysis preservation

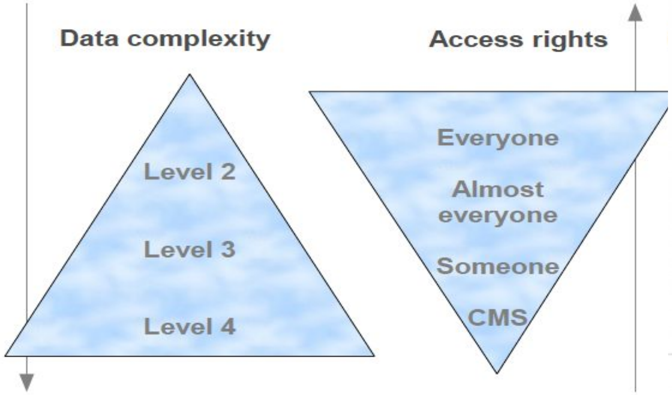
- Badly written code
- Code not being shared and lost
- Code not maintained long term
- Proper documentation of the code and the analysis details

Mitigation: Guidelines for analyses, enforce these if possible

Guidance into data complexity

The “DPHEP Preservation Levels”

Preservation Model		Use Case	
1	Provide additional documentation	Publication related info search	**Documentation**
2	Preserve the data in a simplified format	Outreach, simple analyses	**Outreach, reanalysis**
3	Preserve the analysis level software and data format, frameworks	Full scientific analysis, based on the existing reconstruction	**Technical Preservation Projects**
4	Preserve the reconstruction and simulation software as well as the basic level data	Retain the full potential of the experimental data	

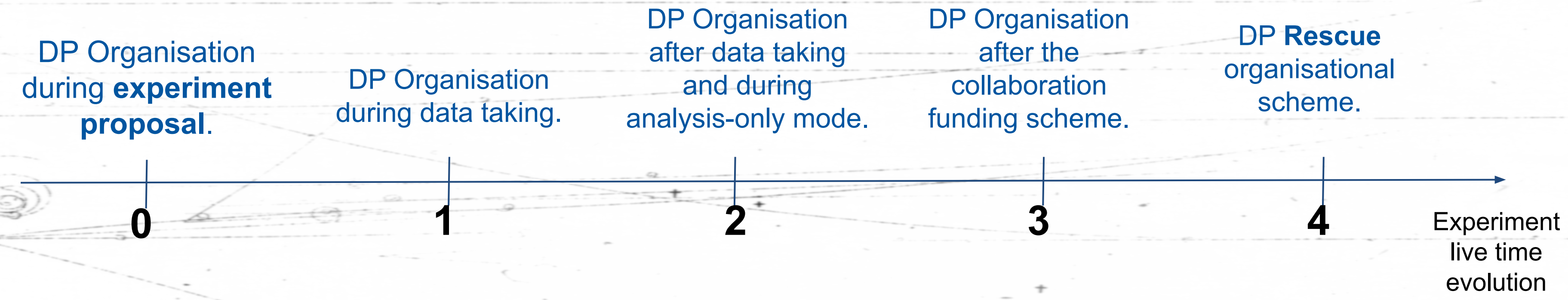


A matter of collaboration as well (I)

The supervision and knowledge transfer/**capture** is essential at long term

Need to clarify the status and the rules

Various stages of organisation (CL) can be defined:



0 — DP Organisation during **experiment proposal**.

1 — DP Organisation during data taking.

2 — DP Organisation after data taking and during analysis-only mode.

3 — DP Organisation after the collaboration funding scheme.

4 — DP **Rescue** organisational scheme.

Experiment live time evolution

A matter of collaboration as well (II)

Level 4: This organisation scheme is to be activated when:

The host laboratory stops support and announce no long-term commitment.

The official collaboration/ data stewardship is stopped with no further plans (no step 3 is clearly defined).

Taking no action == decommissioning (deleting) the data.

“Securely” storing/freezing the files and the latest version of the software is certainly not a substitute for a preservation project.

Laboratory/ Collider	Experiment	Period	Preservation Level	Data Volume	Present status	CL
DESY/PETRA	JADE	1979–1986	4	1 TB	Analysis running on preserved data; migrated from DESY to MPP	4
CERN/LEP	ALEPH, DELPHI, OPAL L3	1989-2000	4 2	0.5 PB	ADO: Analysis running on preserved data	4
DESY/HERA	H1 ZEUS	1992 – 2007	4 3/4	0.5 PB 0.2 PB	Analysis running on preserved data	3
SLAC/PEP II	BABAR	1999–2008	4	2 PB	Analysis running on preserved data; migrated from home lab to different centers	4
KEK/KEKB	Belle I	1999-2010	4	4 PB	Analysis running on preserved data; Compatible with Belle II computing	2
FNAL/TeVatron	DØ CDF	1983–2011	4 4	8.5 PB 9 PB	Archived on tapes	4
BNL/RHIC	PHENIX	2000–2016	3	25 PB	Analysis running on preserved data	3
FNAL/v-beam	Minerva	2010–2019	3	10 TB	Analysis running	2
IHEP/BEPCII	BESIII	2009–2030	4	6 PB	Collecting and analyzing data	1
CERN/LHC	ALICE, ATLAS, CMS, LHCb	2010-2041	4	O(1EB-10EB)	Collecting and analyzing data	1

Cost and Benefit

C1	Host laboratories allocate **person power** and **computing resources** - specifically to DP. (in % to the construction/operation costs)
C2	Collaborating laboratories participate in the effort: **replicate or take over** data and computing systems and provide technical assistance.
C3	Researchers and engineers participate outside their main research area.
C4	Innovative computing projects, including pluri-disciplinary open science initiatives, may offer attractive opportunities for data preservation and are therefore an indirect source of support.
C5	The proximity of a **follow-up experiment** clearly helps in structuring and supporting a data preservation project.

B1	**New publications** – counting here those executed with a strong involvement of the dedicated DP systems.
B2	**Derivative work**: Publications made by other groups/people using the new publications produced at B1.
B3	**Preserving** the **scientific expertise** and the leadership in the field of the experiment, possibly **boosting the transition** to a new experiment
B4	Technology expertise in robust data preservation. Improved ability to **plan for new experiments** and preserve their scientific potential at long term.

Figure of Merit:

FoM = B1/C1

Conclusions after 10 years: the scientific output

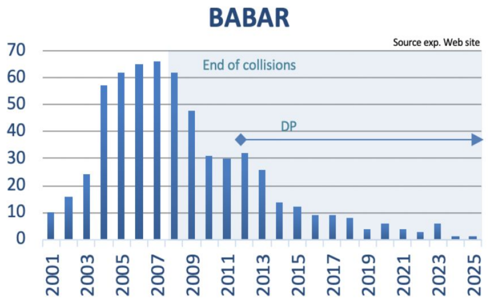
DP is a **cost-effective way of doing fundamental research** by exploiting unique data sets in the light of the increasing theoretical understanding.

DP leads to

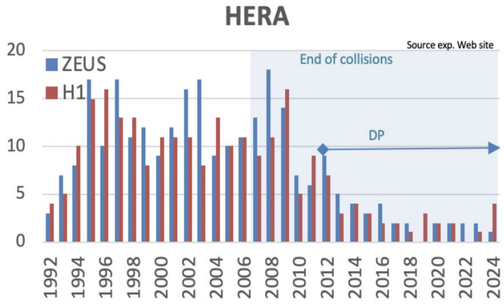
- a **significant increase in the scientific output** (>10% typically)
- for a minimal investment overhead (0.1%).
- …. As predicted in 2013

Lesson: When collisions are stopped, **~20% of the publications are still to come**, and half of them are unknown/unplanned!

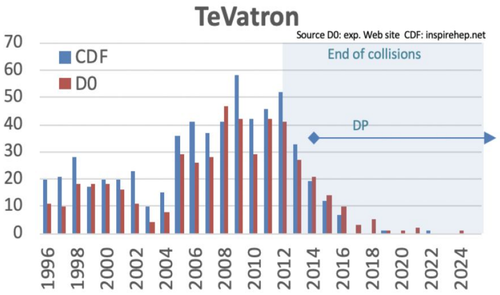
LHC will have 3 decades ahead after the end of collisions !


BABAR

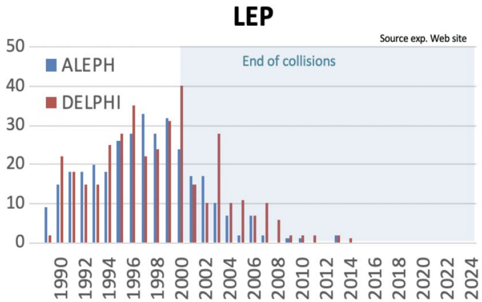
Source exp. Web site


HERA

Source exp. Web site


TeVatron

Source D0: exp. Web site CDF: inspirehep.net


LEP

Source exp. Web site

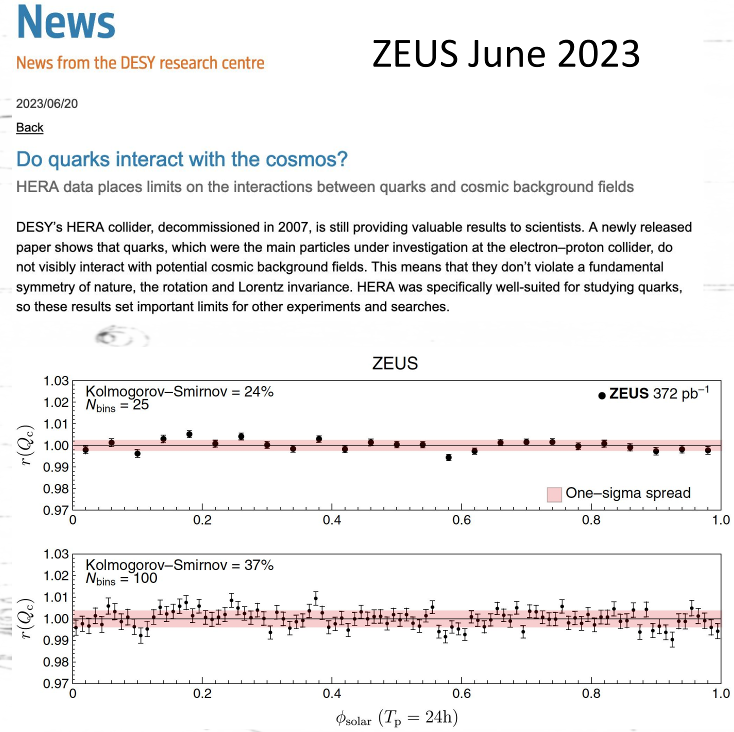
Figure Updated 15/03/2025

	Data taking stopped	Publications before 2012	Publications 2012-2022	Scientific return increase %
Babar	2008	471	154	33%
H1+ZEUS	2007	436	62	14%

… and ongoing

BaBar April 2023
The 600th paper

ZEUS June 2023



News

News from the DESY research centre

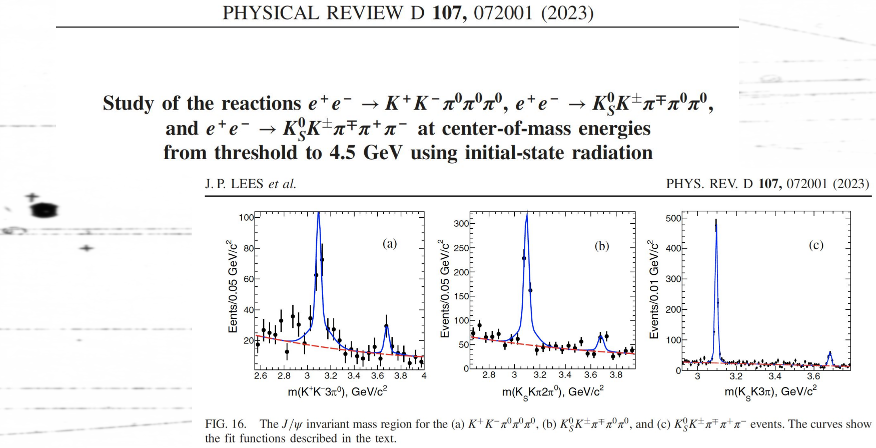
2023/06/20

Back

Do quarks interact with the cosmos?

HERA data places limits on the interactions between quarks and cosmic background fields

DESY's HERA collider, decommissioned in 2007, is still providing valuable results to scientists. A newly released paper shows that quarks, which were the main particles under investigation at the electron–proton collider, do not visibly interact with potential cosmic background fields. This means that they don't violate a fundamental symmetry of nature, the rotation and Lorentz invariance. HERA was specifically well-suited for studying quarks, so these results set important limits for other experiments and searches.



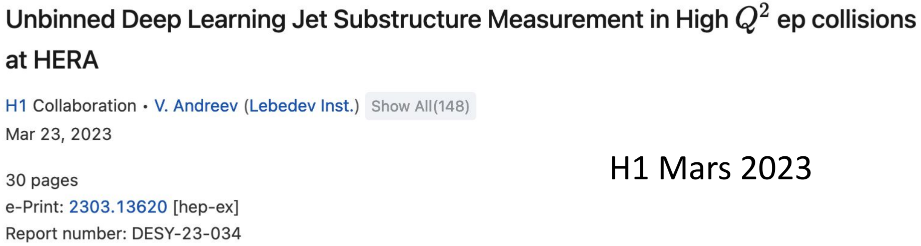
PHYSICAL REVIEW D **107**, 072001 (2023)

Study of the reactions $e^+e^- \to K^+K^-\pi^0\pi^0\pi^0$, $e^+e^- \to K^0_SK^\pm\pi^\mp\pi^0\pi^0$, and $e^+e^- \to K^0_SK^\pm\pi^\mp\pi^+\pi^-$ at center-of-mass energies from threshold to 4.5 GeV using initial-state radiation

J. P. LEES *et al.* PHYS. REV. D **107**, 072001 (2023)

(a)
(b)
(c)
Eents/0.05 GeV/c^2
Events/0.05 GeV/c^2
Events/0.01 GeV/c^2
m(K$^+$K$^-$3π^0), GeV/c^2
m(K$_S$Kπ2π^0), GeV/c^2
m(K$_S$K3π), GeV/c^2

FIG. 16. The J/ψ invariant mass region for the (a) $K^+K^-\pi^0\pi^0\pi^0$, (b) $K^0_SK^\pm\pi^\mp\pi^0\pi^0$, and (c) $K^0_SK^\pm\pi^\mp\pi^+\pi^-$ events. The curves show the fit functions described in the text.



Unbinned Deep Learning Jet Substructure Measurement in High Q^2 ep collisions at HERA

H1 Collaboration • V. Andreev (Lebedev Inst.) Show All(148)

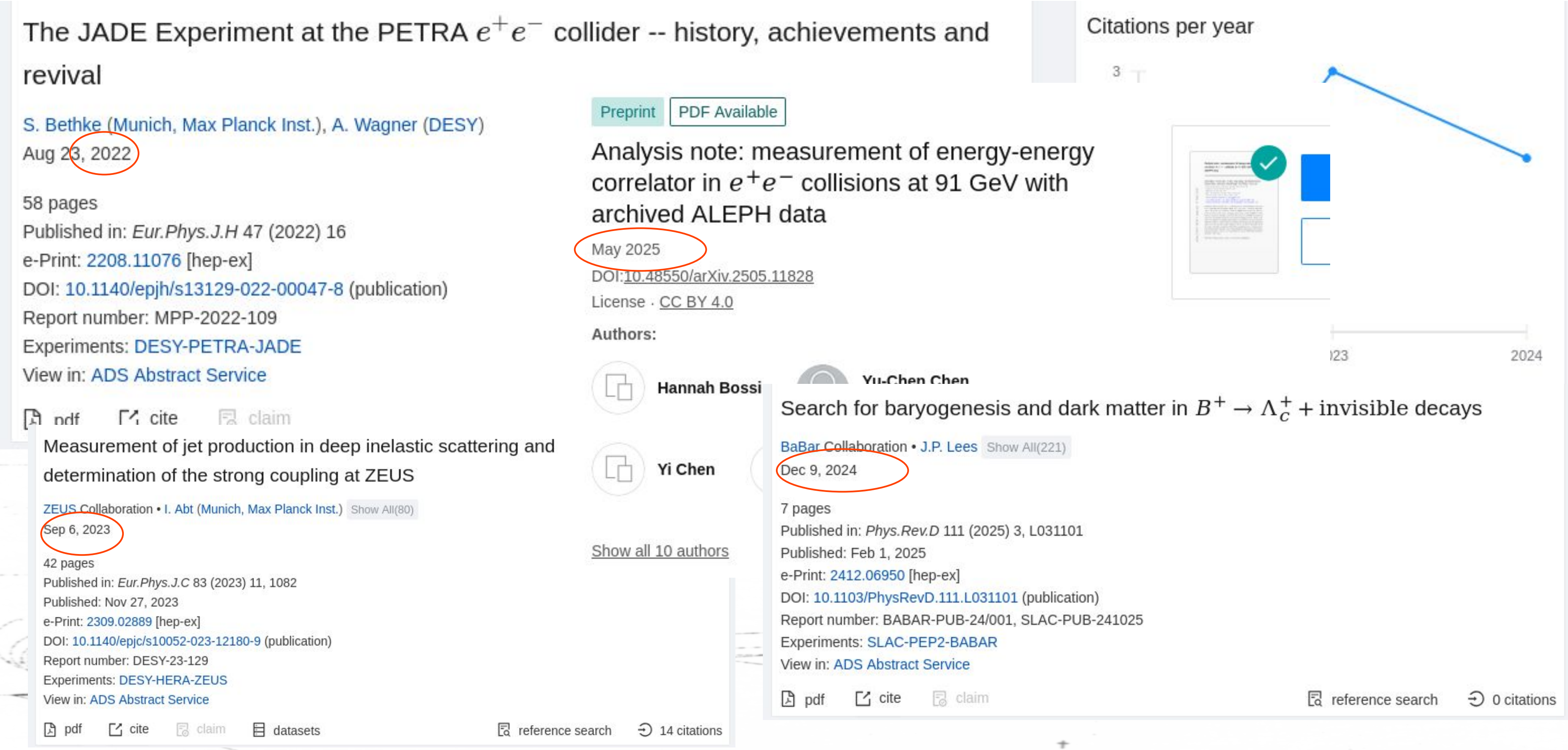
Mar 23, 2023

30 pages

e-Print: 2303.13620 [hep-ex]

Report number: DESY-23-034

H1 Mars 2023



The JADE Experiment at the PETRA e^+e^- collider -- history, achievements and revival

S. Bethke (Munich, Max Planck Inst.), A. Wagner (DESY)

Aug 23, 2022

58 pages

Published in: *Eur.Phys.J.H* 47 (2022) 16

e-Print: 2208.11076 [hep-ex]

DOI: 10.1140/epjh/s13129-022-00047-8 (publication)

Report number: MPP-2022-109

Experiments: DESY-PETRA-JADE

View in: ADS Abstract Service

Analysis note: measurement of energy-energy correlator in e^+e^- collisions at 91 GeV with archived ALEPH data

May 2025

DOI:10.48550/arXiv.2505.11828

License · CC BY 4.0

Authors:

Hannah Bossi

Yu-Chen Chen

Yi Chen

Show all 10 authors

Measurement of jet production in deep inelastic scattering and determination of the strong coupling at ZEUS

ZEUS Collaboration • I. Abt (Munich, Max Planck Inst.) Show All(80)

Sep 6, 2023

42 pages

Published in: *Eur.Phys.J.C* 83 (2023) 11, 1082

Published: Nov 27, 2023

e-Print: 2309.02889 [hep-ex]

DOI: 10.1140/epjc/s10052-023-12180-9 (publication)

Report number: DESY-23-129

Experiments: DESY-HERA-ZEUS

View in: ADS Abstract Service

Search for baryogenesis and dark matter in $B^+ \to \Lambda_c^+ +$ invisible decays

BaBar Collaboration • J.P. Lees Show All(221)

Dec 9, 2024

7 pages

Published in: *Phys.Rev.D* 111 (2025) 3, L031101

Published: Feb 1, 2025

e-Print: 2412.06950 [hep-ex]

DOI: 10.1103/PhysRevD.111.L031101 (publication)

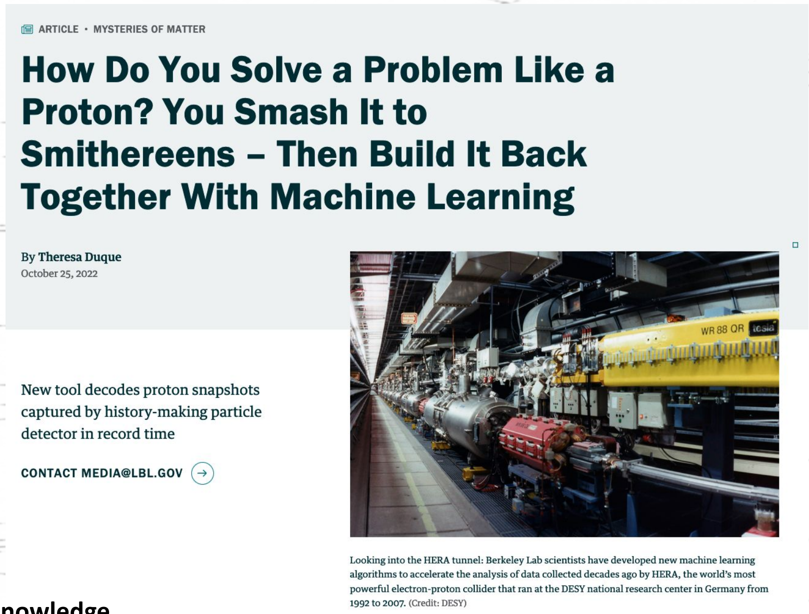
Report number: BABAR-PUB-24/001, SLAC-PUB-241025

Experiments: SLAC-PEP2-BABAR

View in: ADS Abstract Service

Boosting the future experiments

- **HERA EIC**
  - “Scientists today have a **renewed interest in HERA’s particle experiments**, as they hope to use the data – and more precise computer simulations informed by tools like OmniFold – to aid in the analysis of results from future electron-proton experiments, such as at the Department of Energy’s next-generation **Electron-Ion Collider (EIC)**. “
- **Possibly**
  - **LHC FCChh**
  - **LEP FCCee**



ARTICLE • MYSTERIES OF MATTER

How Do You Solve a Problem Like a Proton? You Smash It to Smithereens – Then Build It Back Together With Machine Learning

By Theresa Duque
October 25, 2022

New tool decodes proton snapshots captured by history-making particle detector in record time

CONTACT MEDIA@LBL.GOV

Looking into the HERA tunnel: Berkeley Lab scientists have developed new machine learning algorithms to accelerate the analysis of data collected decades ago by HERA, the world’s most powerful electron-proton collider that ran at the DESY national research center in Germany from 1992 to 2007. (Credit: DESY)

Preserved data can be used to transfer knowledge, training/teaching, outreach or boosting new research programs

Relationship with Open Data

- Two different beasts, albeit related
- Not the same but complementary to each other
  - Open Data can be serve as a long term data preservation strategy
  - Can be made as easy as agreeing on a data access strategy, and then dropping access restrictions
- Open Data requires a certain level of data preservation
  - Can be restricted to a subset or an extract (ntuple) of the preserved data
- Open Data solves the long term data access problem

4th DPHEP workshop Oct 2024: Agenda

Wednesday Oct 2nd

- 14:25 **ALEPH**; Jacopo Fanini
- 14:45 **CERNLIB;** Andrii Verbytskyi, Ulrich Schwickerath
- 15:00 **DELPHI** ; Dietrich Liko, Dr Ulrich Schwickerath
- 15:15 **OPAL** ; Matthias Schroeder
- 15:30 **DELPHI and OPAL event displays;** M.Schroeder

- 16:00 **ZEUS**; Achim Geiser
- 16:20 **H1** ; Speaker: Henry Klest
- 16:40 **JADE** Andrii Verbytskyi /Richard Hildebrandt
- 17:00 **PHENIX** ; Maxim Potekhin
- 17:20 **BaBar** ; Marcus Ebert

Thursday Oct 3rd

- 09:00 **KEK / Belle I & II** ; Takanori Hara
- 09:20 **BESIII** Gang Chen
- 09:40 **CERN Open Data portal** Pablo Saiz
- 10:00 **REANA** Marco Donadoni (CERN)
- 10:20 **CERN Analysis Preservation porta** P. Fokianos
- 11:00 **CERN Open Data: Policy/implementation**; J. Boyd
- 11:20 **ALICE** ; David Dobrigkeit Chinellato
- **11:40 ATLAS**; Zach Marshall
- 12:00 **LHCb**; Dillon Fitzgerald
- 14:00 **Preserving ANTARES legacy data** ; Jutta Schnabel **14:20 PUNCH4NFDI** ; Achim Geiser
- 14: **40 CMS** ; Julie Hogan
- 15:00 **ICFA Data Lifecycle Panel** ; Kati Lassila-Perini
- 15:25 **DPHEP Collaboration**

4th DPHEP workshop

Key points

Significant progress has been made since the last workshop, both on data preservation and opening of data.

While CMS has pioneered the publication of **open data**, the other LHC experiments are rapidly catching up now: ATLAS, LHCb, ALICE

Open data policies are increasingly applied also **beyond LHC**; useful synergy and data opening calendar

LEP Data is (mostly) **alive and active** ! DELPHI data has been recently released as open data

Observations

In many contributions continued funding of DP was mentioned as an issue. An example is the BaBar experiment, whose software is running on outdated hardware but there is no funding to replace them.

Transfer across generations is visible

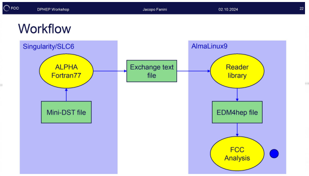
- HERA, RHIC -> EIC
- LEP -> FCC

Bottom up approach starting from the people involved in the practical work

Packed agenda, with input from HEP and also astromomy

Highlights:LEP data is back !

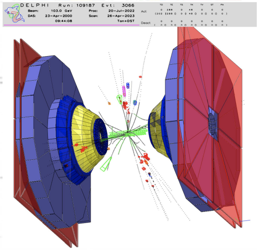
- **ALEPH:**
  - Full stack on containers
  - EDM4HEP migration on



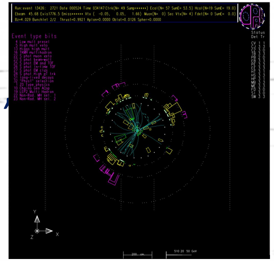
Workflow

Singularity/SLC6: Mini-DST file → ALPHA Fortran77 → Exchange text file → AlmaLinux9: Reader library → EDM4hep file → FCC Analysis

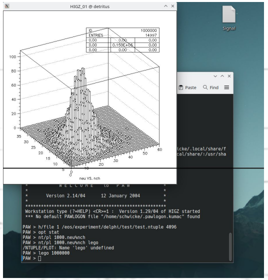
- **DELPHI:**
  - Full stack available, based on community CERNLIB
  - Available on Opendata.org



- **OPAL:**
  - Full stack resurrected, based on community CERNLIB
  - Plan to open the data as DELPHI did, work on this ongoing



(Community) CERNLIB



```
*          W E L C O M E   to   P A W                *
*                                                    *
*       Version 2.14/04      12 January 2004         *
*                                                    *
******************************************************
Workstation type (?=HELP) <CR>=1 :  Version 1.29/04 of HIGZ started
*** No default PAWLOGON file "/home/schwicke/.pawlogon.kumac" found

PAW > h/file 1 /eos/experiment/delphi/test/test.ntuple 4096
PAW > opt stat
PAW > nt/pl 1000.neu%nch
PAW > nt/pl 1000.neu%nch lego
/NTUPLE/PLOT: Name 'lego' undefined
PAW > lego 1000000
PAW >
```

- 64bit support
- Works on decent OS versions
- Binaries soon available from /cvmfs/dphep.cern.ch
- Support for xrootd

Collaborative effort:

- Frequent LEP TF meetings
- CERNLIB and LEP collaboration
- Synergies e.g. for event displays of DELPHI and OPAL
- Interest in DELPHI about EDM4hep

Best practices: Learn from the past

- ICFA panel on best practices being collected
  - Coming soon! Stay tuned!
- Learn from past experiments
  - Remember Murphies law

Keep it simple	
Rely on open and free software	Remove license cost and avoid 3rd party closed source software
Reduce external dependencies as much as possible	Removed site specific dependencies as much as possible
Preserve documentation and knowledge	Ensure documentation quality
Use Automation	Continuous integration
	Automated testing
Plan for and enforce analysis preservation	Used to be a difficult task, and still is !

Conclusions

Significant/measurable impact of dedicated DP projects @expts./labs	– Production of high quality and unique scientific results at very low (non-zero) cost • **10%** output for less than **1%** investment: ✓ • **Long term organisation proves to be productive** – Signs of re-vigorating collaborations in the context of new projects, e.g. HERA-EIC; LEP-FCCee – Case for longer term preservation: data sets parking • CDF, D0, Babar, LEP, JADE : carefully follow the usability in time
There is full coherence (but not total overlap) between **DP** and **Open Data/Science**	– LHC experiments consider both, looking forward to 2041+20/30 years – Lesson: When collisions are stopped, 20% of the publications are still to come, and half of them are unknown/unplanned!
The (DP)HEP future is also considered	FCC, EIC : transfer of knowledge in DP from LHC/oldies
And more is possible on education, training, outreach … via open data	