

The PHENIX DAP update

- Website
- Open Data & Zenodo
- Keywords and Conferences
- HEPData/GitHub

Maxim Potekhin
(BNL, NPPS)
10/29/2020

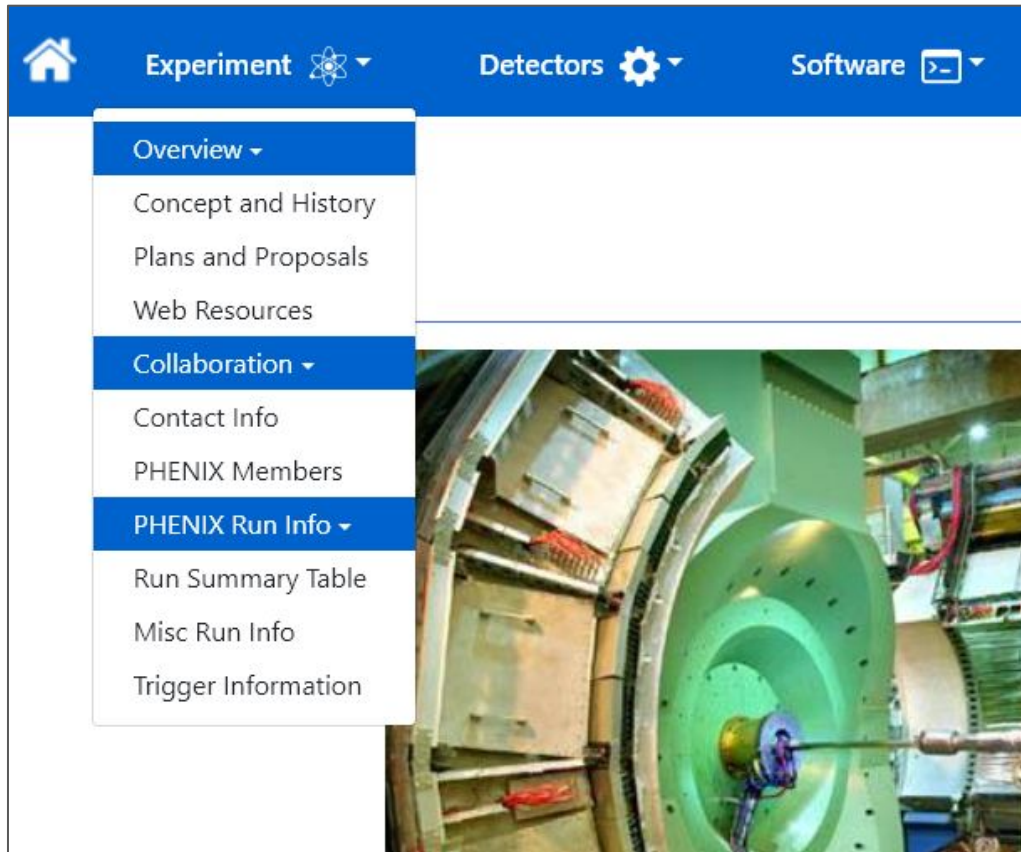
Status of the website

- Development instance: <https://phenixcollaboration.github.io/web/>
- Production instance updated every ~10 days: <https://www.phenix.bnl.gov/>
- Site survey results:
 - Redundant content e.g. overlapping sets of links on multiple pages
 - Pages with minimal content that can be refactored
 - Menu readability not optimal, needs to be improved
 - HEPData policy needs another update






Website updates

- Menus updated for better readability, added category sections (bars)
- Major cleanup, consolidation of content
 - Removed the “documents” and “publications” page from “Resources” as redundant, ensured all links are placed on relevant pages elsewhere
 - Removed a couple of “overview” pages and consolidated content elsewhere
- Updated the HEPData policy page
- Created a separate page for archival of the long-term PHENIX planning documents e.g. decadal plans and BUPs which are very informative

The updated “experiment” menu - more content



The decadal/BUF page

[Experiment](#)  [Results](#)  [Detectors](#)  [Software](#)  [Analysis](#) 

Plans and Proposals







Decadal Plans

- DOI [10.5281/zenodo.3840266](https://doi.org/10.5281/zenodo.3840266) The PHENIX Experiment at RHIC - Decadal Plan 2004–2013 (W.Zajc et al)
- DOI [10.5281/zenodo.3842204](https://doi.org/10.5281/zenodo.3842204) The PHENIX Experiment at RHIC - Decadal Plan 2011–2020 (B.Jacak et al)

Beam Use Proposals

- DOI [10.5281/zenodo.3886886](https://doi.org/10.5281/zenodo.3886886) PHENIX Beam Use Proposal Run-15 and Run-16 (K. Dehmelt)
- DOI [10.5281/zenodo.3887304](https://doi.org/10.5281/zenodo.3887304) PHENIX Beam Use Proposal Run-14 and Run-15 (H. van Hecke)
- DOI [10.5281/zenodo.4122366](https://doi.org/10.5281/zenodo.4122366) PHENIX Beam Use Proposal Run-13 and Run-14 (B. Jacak)
- DOI [10.5281/zenodo.4134966](https://doi.org/10.5281/zenodo.4134966) PHENIX Beam Use Proposal Run-12 and Run-13 (B. Jacak)
- DOI [10.5281/zenodo.4135020](https://doi.org/10.5281/zenodo.4135020) PHENIX Beam Use Proposal Run-11 and Run-12 (B. Jacak)
- DOI [10.5281/zenodo.4135043](https://doi.org/10.5281/zenodo.4135043) PHENIX Beam Use Proposal Run-10 and Run-11 (B. Jacak)
- DOI [10.5281/zenodo.4148873](https://doi.org/10.5281/zenodo.4148873) PHENIX Beam Use Proposal for Run-9 and Beyond (B. Jacak)
- DOI [10.5281/zenodo.4148892](https://doi.org/10.5281/zenodo.4148892) PHENIX Beam Use Proposal Update for Run-7 and Beyond (B. Jacak)

The “Web Resources” and papers.html







[Experiment](#)  [Detectors](#)  [Software](#)  [Analysis](#)  [Results](#)  [Resources](#) 

Web Resources

- [PHENIX on Inspire](#)
- [PHENIX on HEPData](#)
- [PHENIX Community on Zenodo](#)
- [PHENIX Collaboration on GitHub](#)

Legacy Links

The legacy website of the PHENIX Collaboration contained the “papers” page which served both as a catalog of its publications and a collection of links to the publication-related data tables. Please see [this page](#) for details. Please notify the [Data and Analysis Preservation team](#) should you encounter the 404 page such as [this one](#).

[Experiment](#)  [Detectors](#)  [Software](#)  [Analysis](#)  [Results](#)  [Resources](#) 

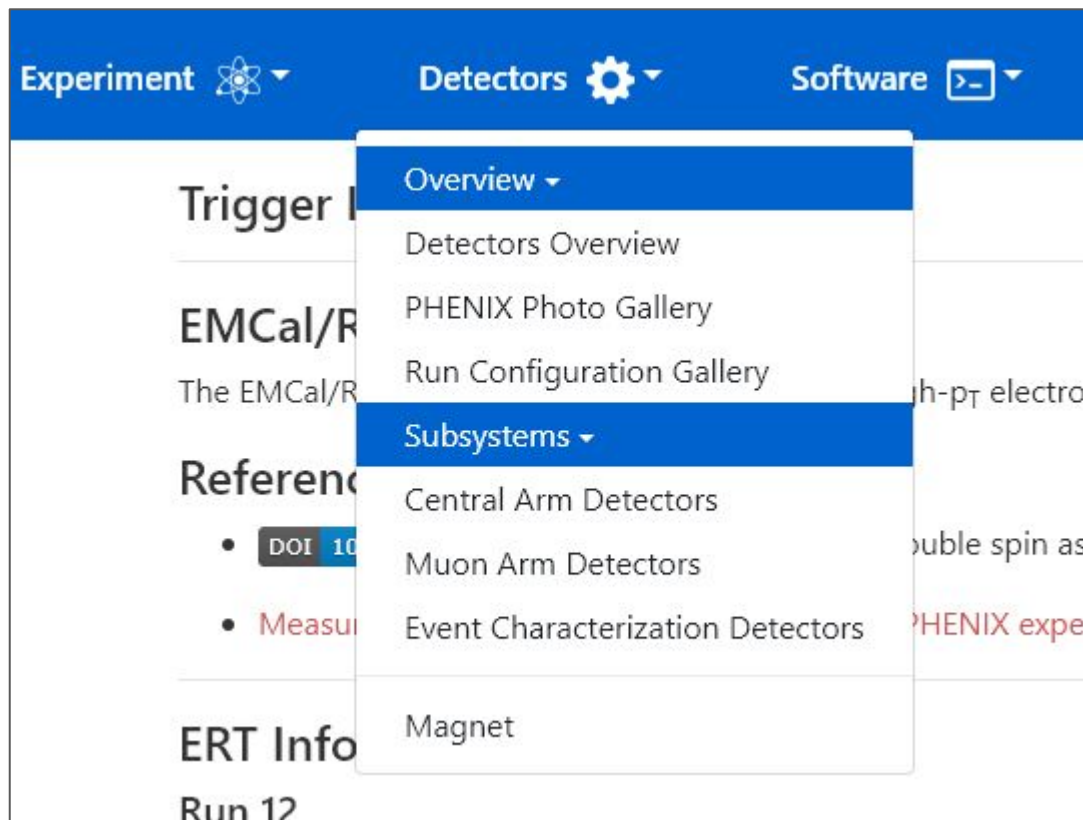
PHENIX Web services migration

The PHENIX Collaboration is currently undertaking an effort to update its Web services pertaining to its publications. As a result, the legacy pages are being migrated to the new platform (this website). Migration of the data component (e.g. the data tables) is also underway but is not currently complete.

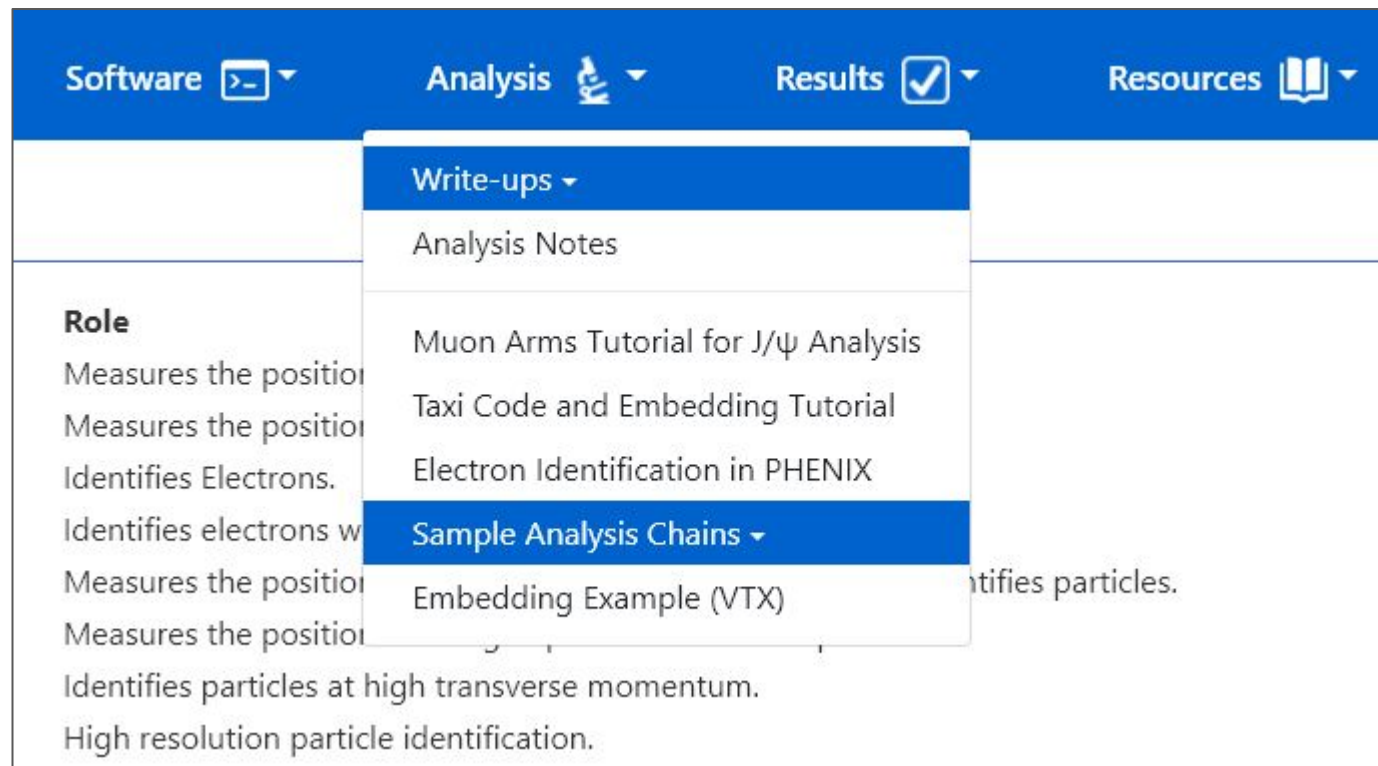
PHENIX is working to upload publication-related data to the [HEPData](#) service at CERN and aims to eventually make all its data available in this manner, replacing the legacy text-based source. Collection of the PHENIX HEPData uploads is available as a [catalog on this web site](#) or directly on the [HEPData site](#).

Legacy “papers.html” links land here

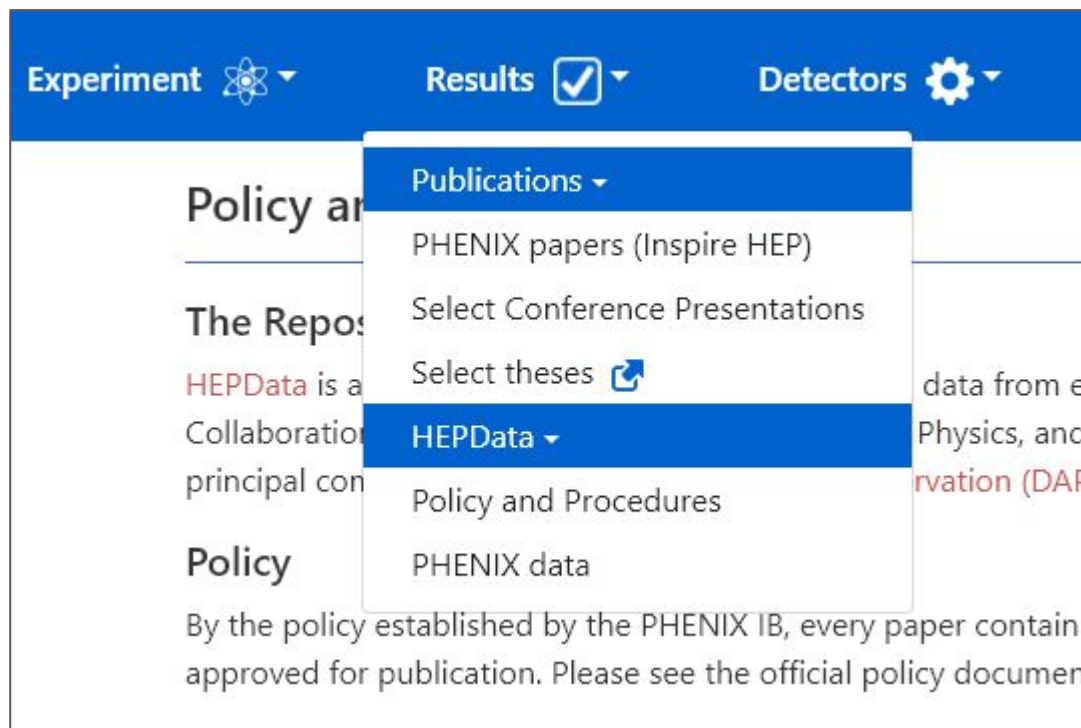
The updated “detectors” menu



The updated “analysis” menu



The updated “results” - waiting to add Open Data



Conference presentations (links to Zenodo)

Experiment

Detectors

Software

Analysis

Results

Resources

Select Conference Presentations

2020

The 36th Winter Workshop on Nuclear Dynamics

PHENIX Presentations

Website

Santa Fe Jets and Heavy Flavor Workshop 2020

PHENIX Presentations

Website

2019

Quark Matter 2019

PHENIX Presentations

Website

Zimányi School 2019

PHENIX Presentations

Website

DNP Fall 2019 Meeting

PHENIX Presentations

Website

International Symposium on Multiparticle Dynamics 2019

PHENIX Presentations

Website

2018

XIII Workshop on Particle Correlations and Femtoscopy

PHENIX Presentations







Website

Hard Probes 2018

PHENIX Presentations

Website

The improved HEPData listing page

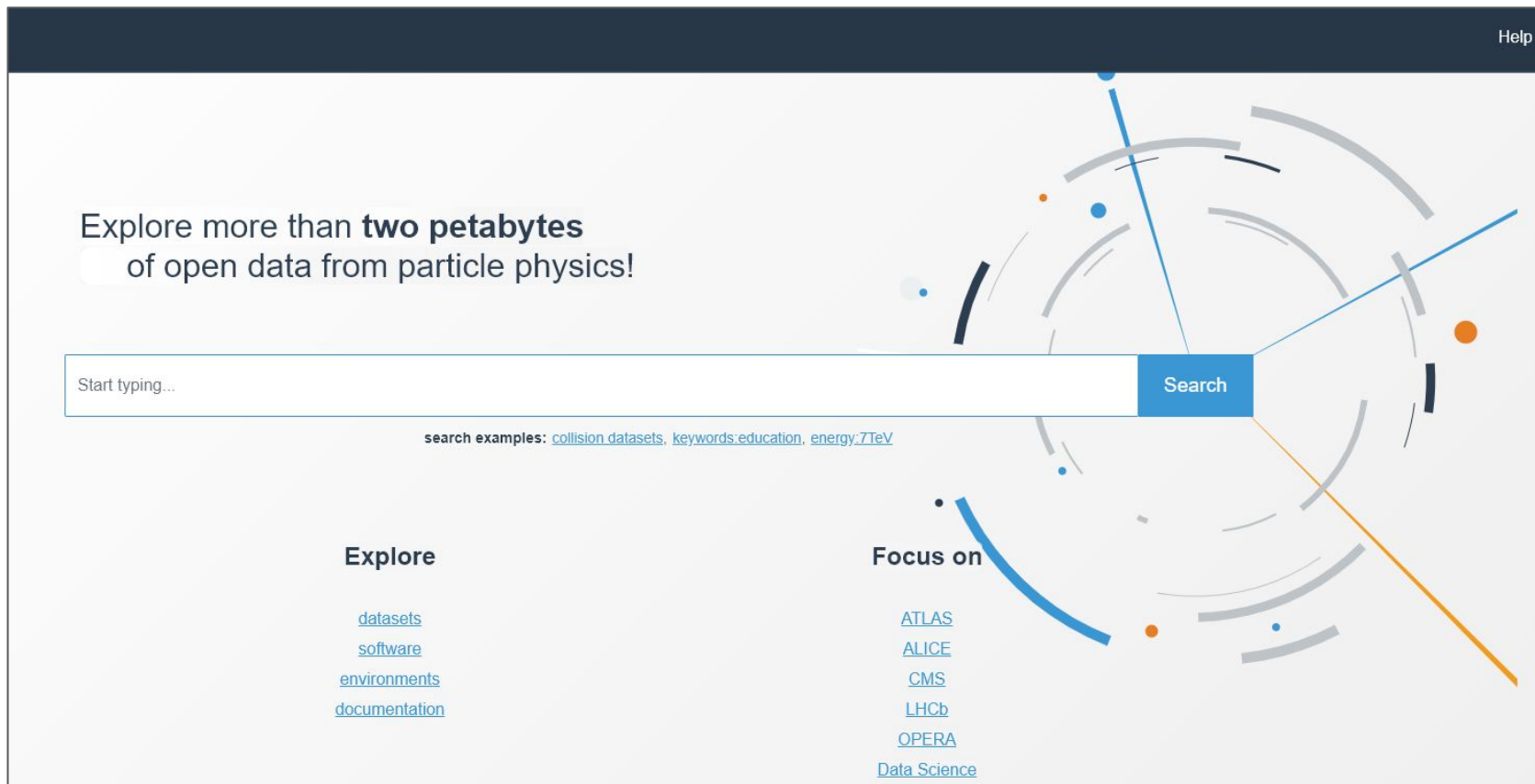
[Experiment](#)  [Results](#)  [Detectors](#)  [Software](#)  [Analysis](#)  [Resources](#) 

PHENIX data

Uploaded HEPData materials (click on HEPData IDs to access entries)

- [98626](#) Measurement of J/ψ at forward and backward rapidity in p+p, p+Al, p+Au, and $^3\text{He}+\text{Au}$ collisions at $\sqrt{s_{\text{NN}}}=200$ GeV
Phys.Rev.C 102 (2020) 014902 (ppg228, inspire 1762446)
- [96764](#) Dihadron azimuthal correlations in Au+Au collisions at $\sqrt{s_{\text{NN}}}=200\text{GeV}$
Phys.Rev.C 78 (2008) 014901 (ppg083, inspire 778396)
- [96601](#) Transverse energy production and charged-particle multiplicity at midrapidity in various systems from $\sqrt{s_{\text{NN}}}=7.7$ to 200GeV
Phys.Rev.C93 (2016) 024901 (ppg174, inspire 1394433)
- [96572](#) Spectra and ratios of identified particles in Au+Au and d+Au collisions at $\sqrt{s_{\text{NN}}}=200\text{GeV}$
Phys.Rev.C88 (2013) 024906 (ppg146, inspire 1227971)
- [96561](#) Neutral pion production with respect to centrality and reaction plane in Au+Au collisions at $\sqrt{s_{\text{NN}}}=200\text{GeV}$
Phys.Rev.C87 (2013) 034911 (ppg133, inspire 1127262)
- [96550](#) Azimuthal anisotropy of π^0 and η mesons in Au + Au collisions at $\sqrt{s_{\text{NN}}}=200\text{GeV}$
Phys.Rev.C88 (2013) 064910 (ppg129, inspire 1254476)
- [96533](#) Evolution of π^0 suppression in Au+Au collisions from $\sqrt{s_{\text{NN}}}=39$ to 200GeV
Phys.Rev.Lett.109 (2012) 152301 (ppg138, inspire 1107625)
- [96510](#) Suppression of away-side jet fragments with respect to the reaction plane in Au+Au collisions at $\sqrt{s_{\text{NN}}}=200\text{GeV}$
Phys.Rev.C84 (2011) 024904 (ppg116, inspire 872172)

Open Data: <http://opendata.cern.ch/>



The screenshot shows the Open Data portal interface. At the top right is a 'Help' link. The main heading reads 'Explore more than **two petabytes** of open data from particle physics!'. Below this is a search bar with the placeholder text 'Start typing...' and a blue 'Search' button. Under the search bar, it says 'search examples: [collision datasets](#), [keywords: education](#), [energy: 7TeV](#)'. On the left, under the heading 'Explore', there is a list of links: [datasets](#), [software](#), [environments](#), and [documentation](#). On the right, under the heading 'Focus on', there is a list of links: [ATLAS](#), [ALICE](#), [CMS](#), [LHCb](#), [OPERA](#), and [Data Science](#). The background features a decorative graphic of concentric circles and lines in blue, orange, and grey.

Help

Explore more than **two petabytes** of open data from particle physics!

Start typing...

Search

search examples: [collision datasets](#), [keywords: education](#), [energy: 7TeV](#)

Explore

- [datasets](#)
- [software](#)
- [environments](#)
- [documentation](#)

Focus on

- [ATLAS](#)
- [ALICE](#)
- [CMS](#)
- [LHCb](#)
- [OPERA](#)
- [Data Science](#)

Open Data: an example of content

ATLAS ×

☐ include on-demand datasets

Filter by type

☒ Dataset109

☐ Derived67

☐ Simulated42

☒ Documentation5

☐ About1

☐ Activities3

☐ Policy1

☒ Environment5

☐ VM5

☐ News3

☒ Software3

☐ Analysis1

☐ Framework1

☐ Tool1

Filter by experiment

☐ ALICE26

☒ ATLAS125

☐ CMS3964

☐ LHCb12

Sort by: Best match asc. Display: detailed 20 results

Found 125 results.

ATLAS WPath 2015 Masterclass MINERVA Software

This java-based program can be used to view the event displays prepared for the Masterclasses W-Path. The package also includes dataset 2 of the 2015 Masterclasses W-Path datasets. (for Minerva versio...

Software Tool ATLAS

ATLAS Virtual Machine Version μ for 2016 open data release

This set of Virtual Machines (VM) allows you to obtain a computer environment that emulate as much as possible that one use for a particle physicist. Each VM contain ROOT as base software pl...

Environment VM ATLAS

$Z\mu\mu$ + Jets

The ATLAS open data dataset is comprised of real data recorded with the ATLAS detector in 2012 and matching simulated data. Both real and simulated data is subjected to a loose event preselec...

Open Data vs Zenodo

- Zenodo can be used to host any sort of digital data so by definition it's generic
- We started considering the Open Data Portal for placement of analysis-related data files after Gabor provided a writeup and real Ntuples for an analysis tutorial
- While we can use Zenodo (the generic system) to host such materials, the Open Data is designed specifically for this kind of data and has many useful features e.g.
 - Dedicated web pages for the analyses descriptions
 - **Many** production and analysis-specific keywords and predefined categories
 - Integrated hosting of software examples
 - VMs
- Could be the ideal place to put our analysis descriptions based on actual data (with the data and software included)

Open Data: PHENIX prospects

- Originally the system was developed exclusively for the LHC experiments, with OPERA being the only exception (i.e. we can't just start using it like we did with Zenodo)
- It's a great tool, great practice and good PR
- An inquiry placed with the Open Data administration regarding PHENIX joining the portal was favorably received and provided input for their budget request
- Work in progress, hope to finalize this in a week or two

Theses and other materials on Zenodo

- Stacyann continues to upload theses
- The list of keywords has been updated
- Please review latest uploads and comment on keywords
- Need to find a good balance between complex phrases and simple keywords, what works best here?
- Let's stick with lowercase everywhere
 - Can the recent uploads be corrected for this?
- We have 100+ entries on Zenodo by now, their usefulness depends on proper indexing i.e. keywords - help needed.
 - Please review/add/suggest/correct keywords
 - The official list is on GitHub and can be version controlled as usual provided it is in sync with the Zenodo contents
 - <https://phenixcollaboration.github.io/web/resources/keywords.html>







Conference materials

- Can someone take over the upload of the conference materials?
- Upload, tag with keywords
- Let me know the conference name etc
- I'll make sure it shows up on the web page

Redux of a HEPData issue in the previous meetings

- Would it make sense to ask PPGs to use GitHub to store TXT files? This could make the transition smoother since we already handle HEPData submission packages on GitHub.
 - Even the legacy page could point to GitHub
- A predefined folder structure will make this transparent (e.g. ppg0123)

HEPData: policy and procedures page updated

[Experiment](#)  [Results](#)  [Detectors](#)  [Software](#)  [Analysis](#)  [Resources](#) 

Policy and Procedures

The Repository

HEPData is an open-access repository for scattering data from experimental particle physics. *It includes data points* from several thousand publications produced by multiple Collaborations working in High Energy and Nuclear Physics, and is hosted by CERN as a part of its Open Data initiative. The PHENIX Collaboration is using this platform as one of the principal components of its [Data and Analysis Preservation \(DAP\)](#) effort and manages a growing [collection of HEPData entries](#).

Policy

By the policy established by the PHENIX IB, every paper containing tables and/or plots must be accompanied by a data package containing the tables and/or plots data before it is approved for publication. Please see the official policy document (sec. IV.iv):

- [PHENIX Policies for Publication and for Other Dissemination of Results](#)

The data package must

- Conform to the specific format required by the HEPData portal (please see the [documentation](#))
- Be certified by the IRC for each publication
- Have an Inspire ID associated with it

General Comments

In order for data to be successfully uploaded to the HEPData portal, it must conform to a specific format (please check the HEPData site for documentation). Data contents of each item (e.g. a plot) are described in a corresponding file formatted in YAML (i.e. if there are 5 plots in the paper you are expected to provide 5 YAML data files). In addition, a special file *submission.yml* describes the submission as a whole and in particular contains the abstract. The abstract should be formatted as LaTeX and it will be rendered on the HEPData site however not every LaTeX feature will work correctly and this needs to be checked (see the “sandbox” feature below). Since YAML allows comments - lines starting with a “#” sign - it is easy to add any sort of helpful extra information to *submission.yml*.

Existing text files can be converted to the HEPData format with some effort. The DAP team is looking at technical solutions to facilitate this process. For example, if plots are generated using ROOT macros the code can be instrumented to output same data in a format compatible with HEPData. There is a helpful write-up about preparing data for upload:

HEPData: other work items

- The HEPData policy page further updated
- Did a thorough exercise with PPG228 - many thanks to Krista
- We have tested the data checking/approval workflow rather extensively and in a realistic setting for the first time (i.e. the reviewer is in the working group and there is an iterative loop with GitHub for version control).
- We are making use of YAML/HEPData validation tools and will share experience with future submitters.
- Please review the page, and we'll announce it soon
- Is anyone keeping track of what recent PPG items need to address the HEPData aspect? Plenty of e-mail threads and updates lately. Need to start preparing HEPData packages before time pressure develops.

Summary

- Progress with many aspects of the website
- Steady uploads to Zenodo including BUPs and theses
- The CERN Open Data portal access is being negotiated and we are optimistic it will work for us
- Tested the newly created HEPData procedures while uploading and validating ppg228 - should we advertise this to the Collaboration?
- There is a recent crop of publications in the pipeline which should be HEPDatified