






The PHENIX DAP update

- Zenodo
- Keywords and Conferences
- HEPData/GitHub

Maxim Potekhin
(BNL, NPPS)
11/12/2020

Zenodo: completed BUP uploads

[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)
- DOI [10.5281/zenodo.4264805](https://doi.org/10.5281/zenodo.4264805) PHENIX Beam Use Proposal for RHIC Runs 5-9 (W. Zajc)
- DOI [10.5281/zenodo.4264836](https://doi.org/10.5281/zenodo.4264836) PHENIX Beam Use Proposal for RHIC Runs 4-8 (W. Zajc)

Added DNP2020, ready for uploads to Zenodo

Experiment 	Results 	Detectors 	Software 	Analysis 
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
DNP Fall 2020 Meeting			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

Growing list of keywords, correcting some on Zenodo

Experiment

Results

Detectors

Software

Analysis

Conferences (9 items)

Keyword	Description
dnp19	DNP (2019)
dnp20	DNP (2020)
hp18	Hard Probes 2018
ismd19	International Symposium on Multiparticle Dynamics (2019)
qm2019	Quark Matter (2019)
sfjhf20	Santa Fe Jets and Heavy Flavor Workshop (2020)
wpcf2018	XIII Workshop on Particle Correlations and Femtoscopy
wwnd2020	The 36th Winter Workshop on Nuclear Dynamics (2020)
zs19	Zimányi School (2019)

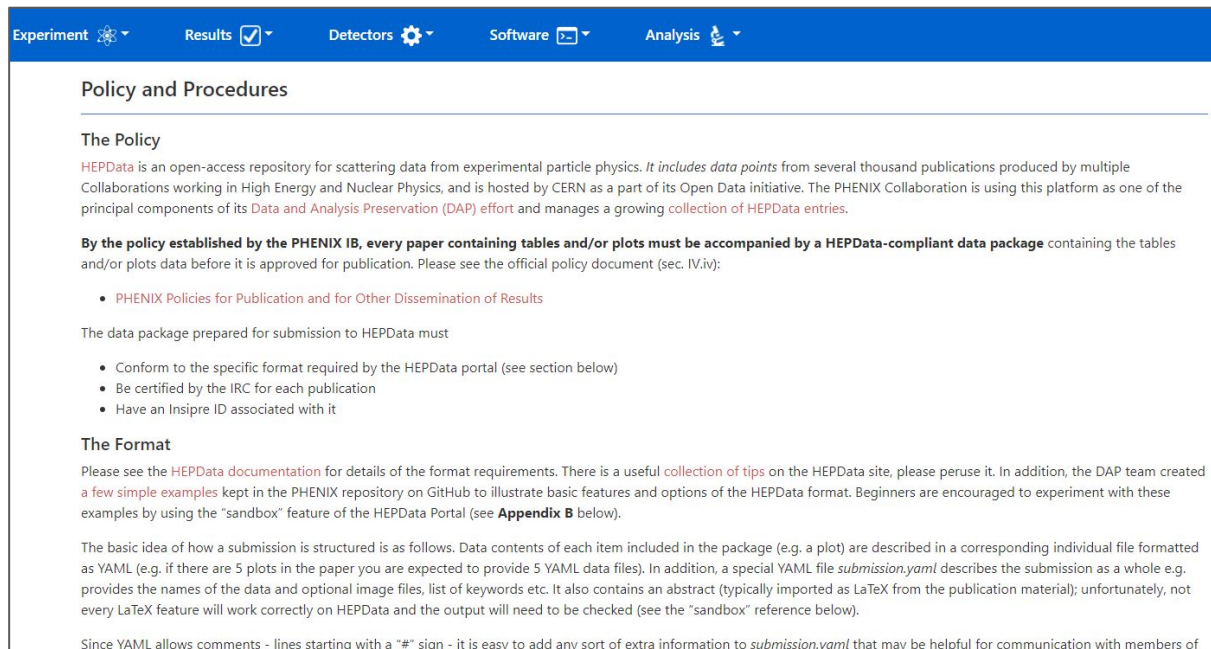
Physics (69 items)

Keyword	Description
3he+au	Helium3-on-gold collisions
anisotropy	Anisotropy
asymmetry	Asymmetry
au+au	Gold-on-gold collisions
azimuthal	Azimuthal
b-meson	B meson
binary scaling	Binary scaling
bose-einstein	Bose-Einstein statistics
bottom	Particles containing the b-quark
centrality	Centrality characteristic of the collision

HEPData instructions page: approve and publish?

A substantial rewrite of the HEPData page, more details added, appendix section added to reduce clutter in the main portion of the instructions

<https://phenixcollaboration.github.io/web/results/hepdata.html>



The screenshot shows the HEPData website interface. At the top is a blue navigation bar with tabs: Experiment, Results (selected), Detectors, Software, and Analysis. Below the navigation bar is the 'Policy and Procedures' section. It includes a sub-section 'The Policy' with a paragraph about HEPData's open-access nature and its role in the PHENIX Collaboration's Data and Analysis Preservation (DAP) effort. A bolded policy statement follows, stating that every paper containing tables and/or plots must be accompanied by a HEPData-compliant data package. A bullet point links to 'PHENIX Policies for Publication and for Other Dissemination of Results'. Below this, it states that the data package must conform to specific requirements, listed in another bullet point: conform to the HEPData portal format, be certified by the IRC, and have an Inspire ID. The 'The Format' section follows, explaining that users should refer to the HEPData documentation for format requirements and that the DAP team has provided simple examples on GitHub. It also mentions that the submission file should include an abstract and keywords. A final paragraph explains the basic structure of a submission, including individual data files in YAML format and a special 'submission.yaml' file. At the bottom, a note mentions that YAML allows comments for extra information.

Experiment Results Detectors Software Analysis

Policy and Procedures

The Policy

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*.

By the policy established by the PHENIX IB, every paper containing tables and/or plots must be accompanied by a HEPData-compliant 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 prepared for submission to HEPData must

- Conform to the specific format required by the HEPData portal (see section below)
- Be certified by the IRC for each publication
- Have an Inspire ID associated with it

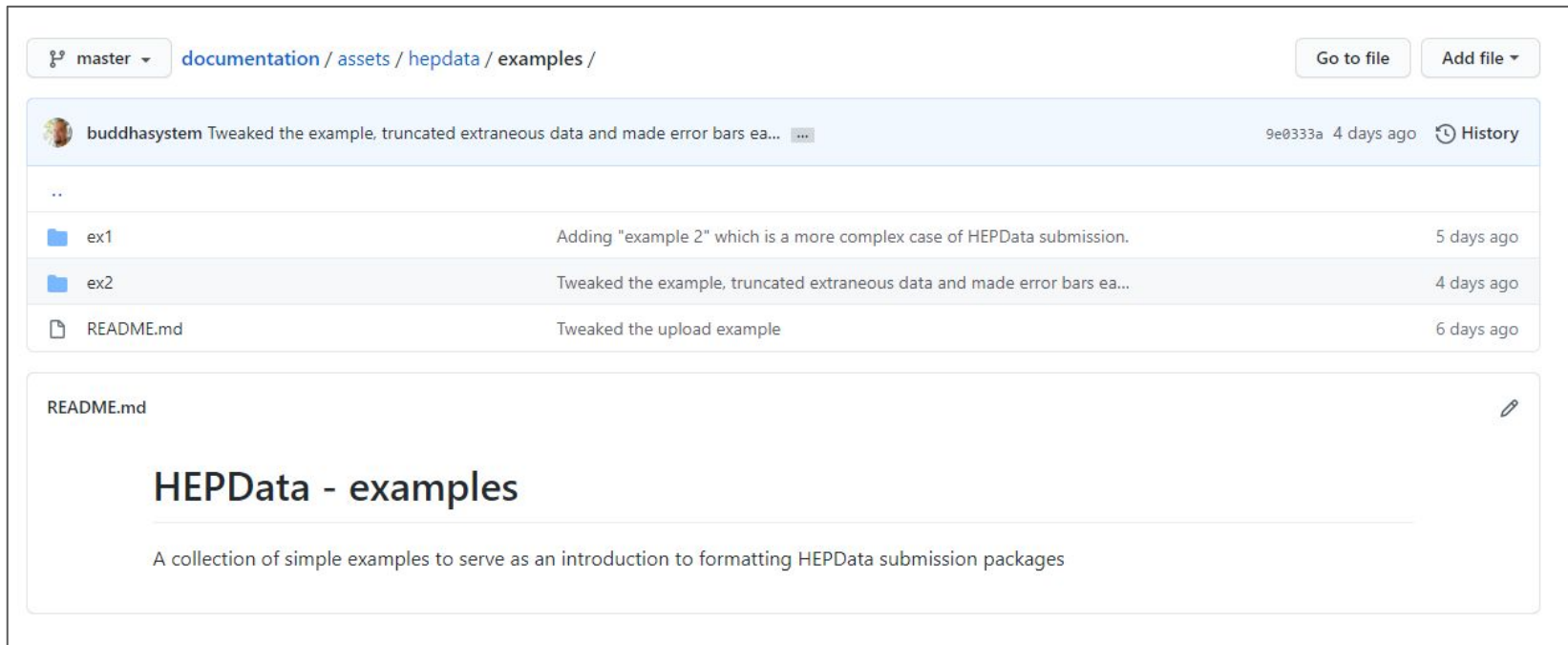
The Format

Please see the [HEPData documentation](#) for details of the format requirements. There is a useful *collection of tips* on the HEPData site, please peruse it. In addition, the DAP team created a few simple examples kept in the PHENIX repository on GitHub to illustrate basic features and options of the HEPData format. Beginners are encouraged to experiment with these examples by using the "sandbox" feature of the HEPData Portal (see **Appendix B** below).

The basic idea of how a submission is structured is as follows. Data contents of each item included in the package (e.g. a plot) are described in a corresponding individual file formatted as YAML (e.g. if there are 5 plots in the paper you are expected to provide 5 YAML data files). In addition, a special YAML file *submission.yaml* describes the submission as a whole e.g. provides the names of the data and optional image files, list of keywords etc. It also contains an abstract (typically imported as LaTeX from the publication material); unfortunately, not every LaTeX feature will work correctly on HEPData and the output will need to be checked (see the "sandbox" reference below).

Since YAML allows comments - lines starting with a "#" sign - it is easy to add any sort of extra information to *submission.yaml* that may be helpful for communication with members of

Created original HEPData examples (on GitHub), links added to the HEPData page on site



The screenshot shows a GitHub repository interface. At the top, the breadcrumb navigation is 'documentation / assets / hepdata / examples /'. To the right are buttons for 'Go to file' and 'Add file'. Below this is a commit history table with columns for the commit message, commit hash, time ago, and a 'History' link. The table lists three commits: a '..' commit, a commit adding 'example 2' (ex2), and a commit tweaking the upload example (README.md). Below the table, the 'README.md' file is selected and its content is displayed. The content includes the title 'HEPData - examples' and a description: 'A collection of simple examples to serve as an introduction to formatting HEPData submission packages'.

master documentation / assets / hepdata / examples / Go to file Add file

Commit Message	Commit Hash	Time Ago	History
..			
Adding "example 2" which is a more complex case of HEPData submission.		5 days ago	
Tweaked the example, truncated extraneous data and made error bars ea...		4 days ago	
Tweaked the upload example		6 days ago	

README.md

HEPData - examples

A collection of simple examples to serve as an introduction to formatting HEPData submission packages

Other HEPData items

- HEPData workshop on 11/10/2020, one more to follow (C.Nattrass)
- Need to add functionality to the keywords page to make it easier to refer to **Zenodo and HEPData** as needed, by same keywords
- “Fixing” older HEPData submissions by adding more keywords is doable, it’s easier for more recent ones where we have complete packages on GitHub
- Unlike Zenodo, there must be a complete new submission for such a correction to take effect
- We have a list of PPGs in the HEPData pipeline or ones which need guidance. This activity is picking up, with the momentum generated by the workshop etc
 - All agree extra DB work is not justified
 - Christine’s spreadsheet is a good starting point, where/how to keep it?

Misc items

- PHENIX hosting on Open Data is approved in principle, but they still need to formally allocate and configure storage which is done by the facility people, and unfortunately this takes time (they are really busy these days)
- PHENIX Phone Book - currently doing DB dump for the site, but
 - Tables are not normalized
 - MGS not in sync with Institutional Affiliation (years don't match), this is currently fixed in a custom script created for the site, not a clean solution
 - In some cases people on MGS are not in the main “people” table
 - We can't count on active support of these tables
- **Proposal: switch to to the CSV file kept on GitHub (as it is now) as the definitive list of current members.** Not hard to adjust the MGS procedure.
- Easy to edit directly on GitHub or Excel or any editor. See next page.

The Phone Book (current CSV on GitHub)



80 lines (80 sloc) | 9 KB

Raw Blame

Search this file...

	family_name	first_name	email	inst_name
2	Li	Xuan	xuanlipx@rcf.rhic.bnl.gov	Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA
3	Aidala	Christine	caidala@bnl.gov	Department of Physics, University of Michigan, Ann Arbor, Michigan 48109-1040, USA
4	Akiba	Yasuyuki	akiba@bnl.gov	RIKEN Nishina Center for Accelerator-Based Science, Wako, Saitama 351-0198, Japan
5	Bazilevsky	Alexander	shura@bnl.gov	Physics Department, Brookhaven National Laboratory (BNL), Upton, New York 11973-5000, USA
6	Berdnikov	Yaroslav	berdnikov@spbstu.ru	Saint Petersburg State Polytechnic University, St. Petersburg, 195251 Russia
7	Chiu	Mickey	chiu@bnl.gov	Physics Department, Brookhaven National Laboratory (BNL), Upton, New York 11973-5000, USA
8	David	Gabor	david@bnl.gov	Department of Physics and Astronomy, Stony Brook University, SUNY, Stony Brook, New York 11794-3800, USA
9	Frawley	Anthony	afrawley@fsu.edu	Florida State University, Tallahassee, Florida 32306, USA
10	Grau	Nathan	ngrau@augie.edu	Department of Physics, Augustana University, Sioux Falls, South Dakota 57197, USA
11	Hong	Byungsik	bhong@korea.ac.kr	Korea University, Seoul 02841, Korea
12	Imrek	Jozsef	mazsi@bnl.gov	Debrecen University, H-4010 Debrecen, Egyetem tér 1, Hungary
13	Johnson	Brant	brant@bnl.gov	Georgia State University, Atlanta, Georgia 30303, USA
14	Lajoie	John	lajoie@iastate.edu	Iowa State University, Ames, Iowa 50011, USA
15	Lebedev	Alexandre	lebedev@iastate.edu	Iowa State University, Ames, Iowa 50011, USA
16	Liu	Ming	ming@bnl.gov	Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

Misc items cont'd

- Private repository on GitHub? A relatively new feature
 - Analysis notes etc?
 - Manageable due to a modest number of participants
- Do we want to reduce dependence on custom DB/PHP applications e.g. such as they exist on the internal site? It will be harder to do in the future.