An overview of the PHENIX DAP website

M.Potekhin

Nuclear and Particle Physics Software Group

PHENIX EC Meeting

BNL

March 19th, 2020

Intro

- The term *DAP* refers to data **and** analysis preservation, commonly described as a union of
  - Bit preservation
  - Software Infrastructure and application code long-term preservation and management
  - Analysis know-how (and other knowledge management)
- The PHENIX DAP effort started in 2019 (task force created)
- Please see the recent slides for more detail, general DAP situation in PHENIX and overview of the tools being introduced, here's the link:
  - https://docs.google.com/presentation/d/1LF-0_qOs7RlmWUka549PtJfbFEDeeZVFciLaEgRfuM8/edit?usp=sharing
  - some of this info presented here in a condensed format
- Long-term DAP goals have synergies with maintaining short-term reproducibility
- The focus of this presentation: improving knowledge management in PHENIX by creating a durable and curated Web resource

DAP needs Knowledge Management

- Long-term knowledge management is the key to DAP - some challenges here
  - Software documentation becoming (or some already being) obsolete
  - Personnel migrating to other projects - continuity of knowledge is a problem
  - A lot of the "know how" is folded into the analysis software which needs to be reviewed and documented (and the software effort underfunded)
- Keeping record of details of the analysis procedures creates a non-trivial extra burden on the researcher
  - this sometimes takes the backseat
- There is a fairly narrow window of opportunity to address at least some of the challenges

The PHENIX Web documentation

- Currently the information describing the detector, run conditions, “foundation software” and many other topics is fragmented and sometimes outdated or obsolete
  - Information spread over multiple servers
  - Broken links are fairly commonplace
  - Info which is needed for DAP is often mixed with info which is not - thus diluting the content - hard to find info one is looking for
  - Mix of HTML pages, images, Wikis and other platforms
- Most of it is outside of version control
- Fixing existing websites is not practical for these and other reasons

The New PHENIX Website

- To better serve the short and long-term DAP goals the Collaboration needs a consolidated and curated Web resource
- Running a content management system (such as Wiki, Drupal etc) over a long period of time is an option but it's not optimal:
  - the material is static anyway (but a DB server needs maintenance)
  - requires PHP and other support, security updates, admin expertise etc
  - version control suboptimal
  - backup not straightforward and requires infrastructure
- Ideally, the long-term Web platform would combine the following:
  - static content (to simplify deployment, long-term maintenance and security)
  - structured data kept separately from the layout
- Solutions do exist. Please see the link on slide 2 for technology survey and details.

The platform

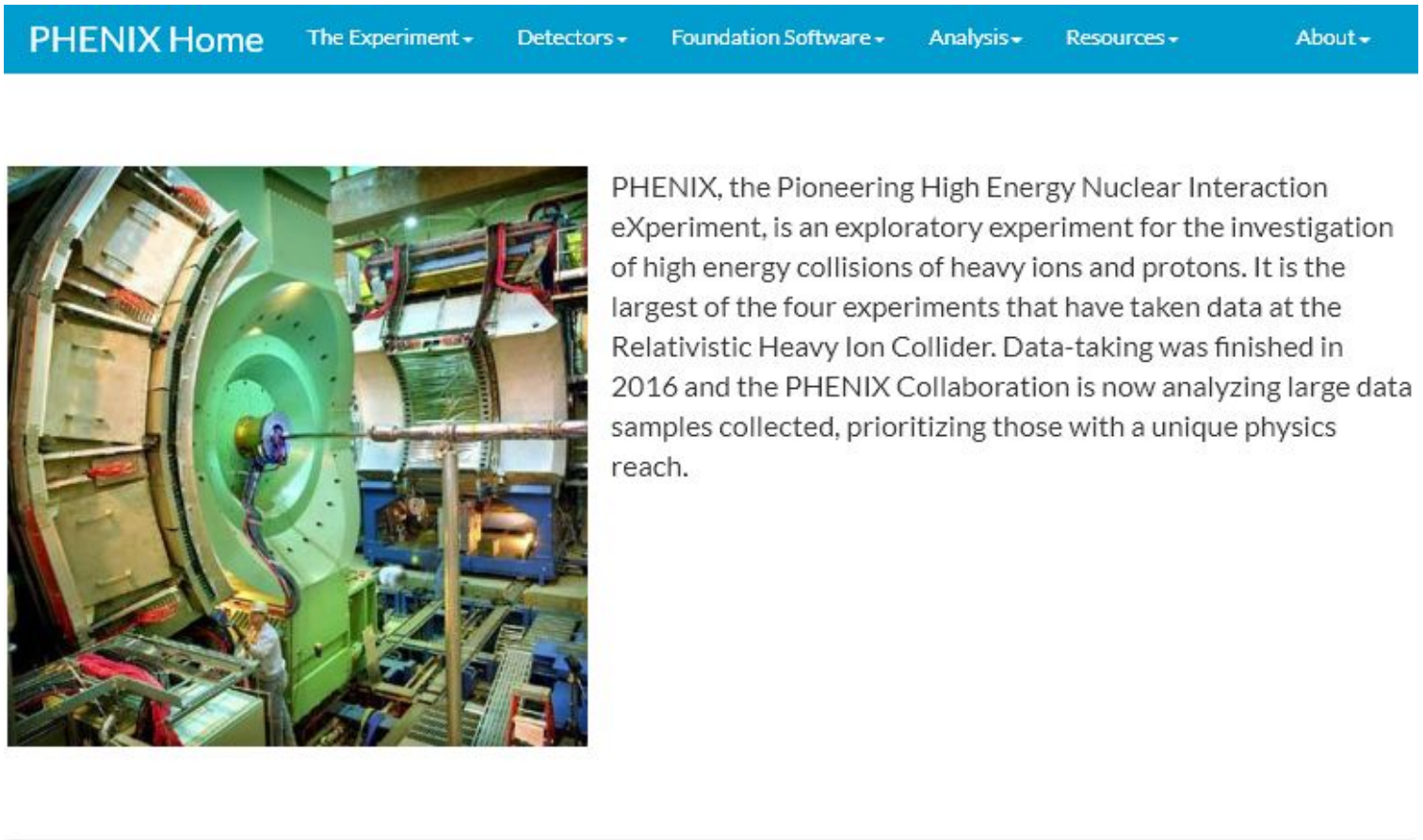
- The static website generator called “Jekyll” which utilizes a standard template language “Liquid” for layouts, parses Markdown syntax for content and YAML for the data component (i.e. zero data in the pages themselves)
  - YAML is essentially a human-readable variant of JSON
  - Likewise, Markdown is a lot more readable and easier to format than HTML
  - “Liquid” is quite similar to other popular template languages and easy to learn
- Data stored YAML can be embedded in individual pages as needed (very flexible)
  - can reference and cross-reference data in many ways
- Site skin/appearance can be change w/o modifying the content
- This site generator is fully integrated into GitHub
  - one just has to configure a repository in a compatible way
  - site is rebuilt with every push to GitHub (free hosting!)
  - good platform for collaborative development with state-of-the-art version control
- Once generated, the site is a hierarchy of a few folders and is portable to any web server and even accessible through a “file://” handle in your local browser
  - i.e. you don’t even need a web server at all, the whole website can be shipped through e-mail, Dropbox or even a USB stick
- NB. GitHub provides hosting of such sites for free and we expect it to remain in business on a time scale relevant for the PHENIX data preservation mandate
  - also, fairly low effort required to deploy such a site at BNL or elsewhere

The plan

- The goal is to create an optimal experience for the user looking for information about the experiment, its detector subsystems, foundation software and going forward, analysis know-how
  - NB. theses and analysis notes will be kept in a dedicated document management system as it is a different domain - to be addressed separately
- Look at best practices and borrow design elements from other projects
- Keep Web development under the official PHENIX Collaboration umbrella on GitHub
  - https://github.com/PhenixCollaboration
- The difficult part: find, curate, and organize various materials to make them ready for the new site
- Status:
  - a prototype website has been recently created
  - not much content right now but it keeps growing continuously
  - your input is needed and appreciated
  - please help in gathering and reviewing materials
- This is the current location: https://phenixcollaboration.github.io/web/
  - please take a look, your feedback is important

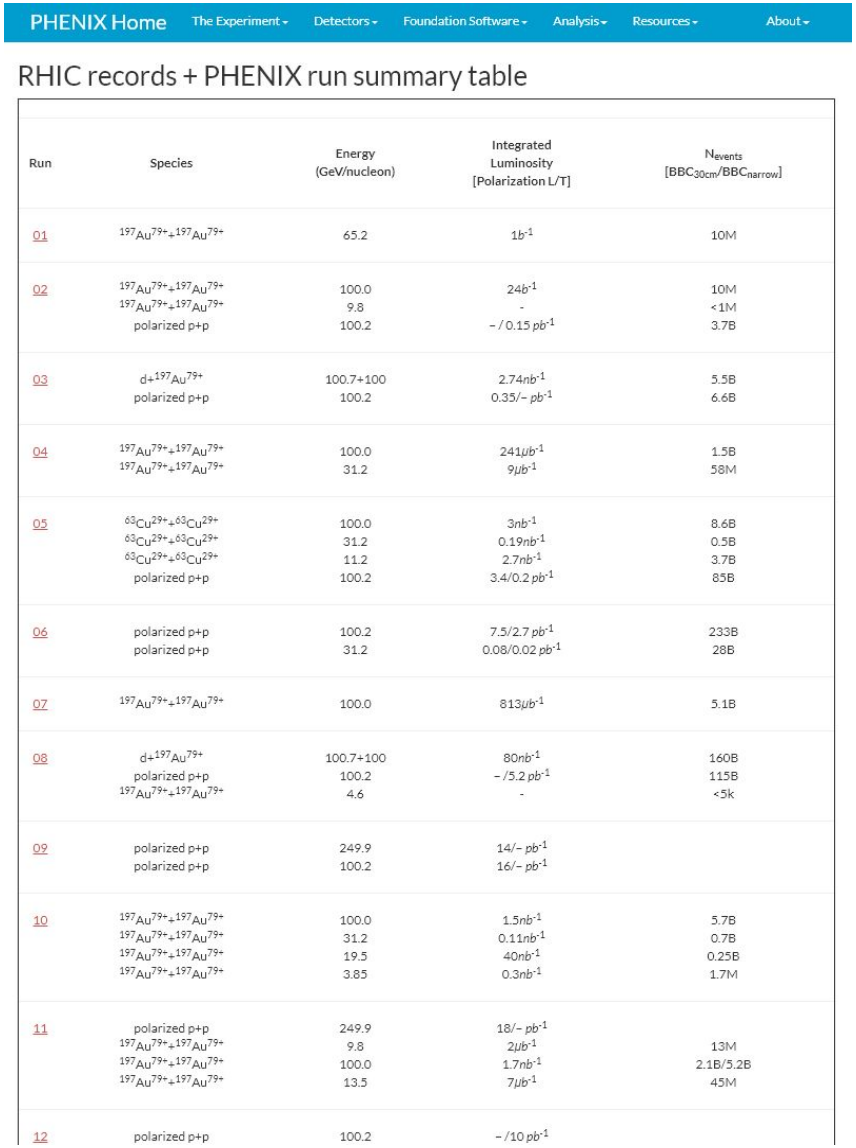
The PHENIX DAP website (prototype)

- https://phenixcollaboration.github.io/web/



Run Summary Table (new)

- https://phenixcollaboration.github.io/web/experiment/runs.html



PHENIX Home | The Experiment | Detectors | Foundation Software | Analysis | Resources | About

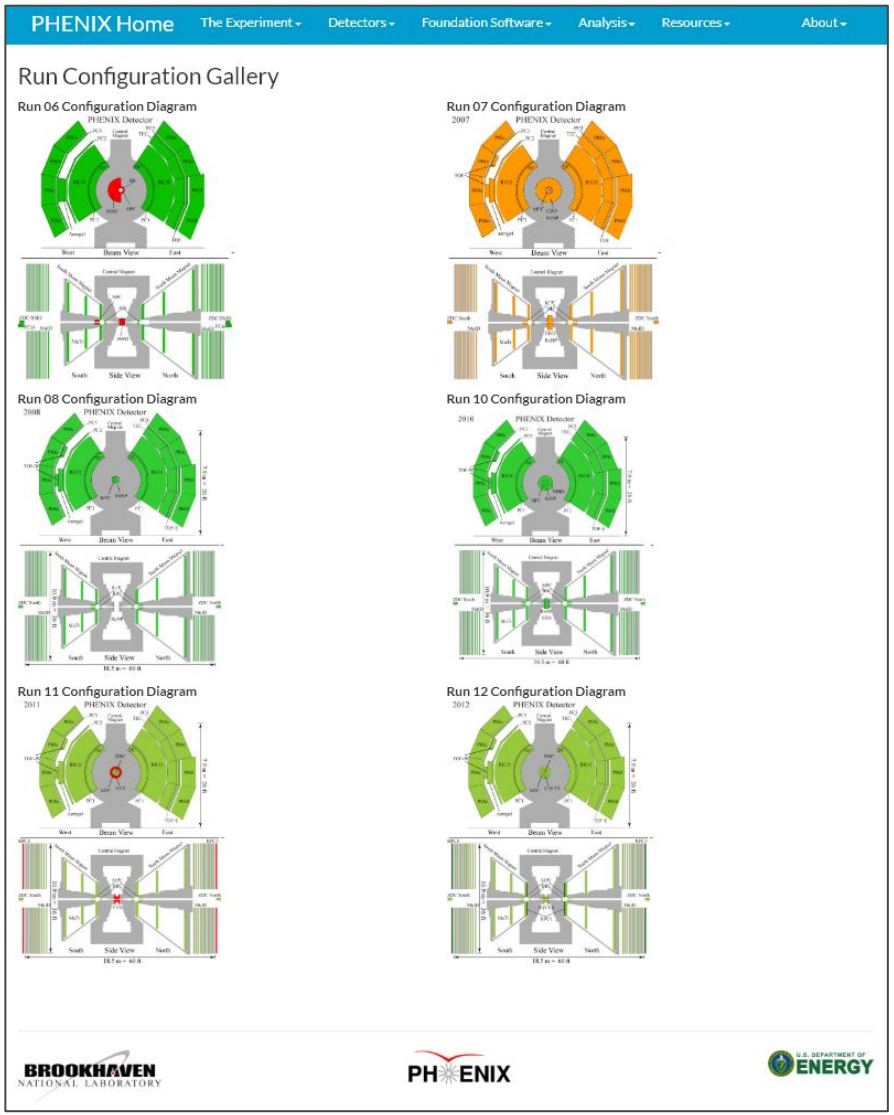
RHIC records + PHENIX run summary table

Run	Species	Energy (GeV/nucleon)	Integrated Luminosity [Polarization L/T]	N_{events} [BBC_{30cm}/BBC_{narrow}]
01	$^{197}Au^{79+}$+$^{197}Au^{79+}$	65.2	$1b^{-1}$	10M
02	$^{197}Au^{79+}$+$^{197}Au^{79+}$ $^{197}Au^{79+}$+$^{197}Au^{79+}$ polarized p+p	100.0 9.8 100.2	$24b^{-1}$ - – / 0.15 pb^{-1}	10M <1M 3.7B
03	d+$^{197}Au^{79+}$ polarized p+p	100.7+100 100.2	$2.74nb^{-1}$ 0.35/– pb^{-1}	5.5B 6.6B
04	$^{197}Au^{79+}$+$^{197}Au^{79+}$ $^{197}Au^{79+}$+$^{197}Au^{79+}$	100.0 31.2	$241\mu b^{-1}$ $9\mu b^{-1}$	1.5B 58M
05	$^{63}Cu^{29+}$+$^{63}Cu^{29+}$ $^{63}Cu^{29+}$+$^{63}Cu^{29+}$ $^{63}Cu^{29+}$+$^{63}Cu^{29+}$ polarized p+p	100.0 31.2 11.2 100.2	$3nb^{-1}$ $0.19nb^{-1}$ $2.7nb^{-1}$ 3.4/0.2 pb^{-1}	8.6B 0.5B 3.7B 85B
06	polarized p+p polarized p+p	100.2 31.2	7.5/2.7 pb^{-1} 0.08/0.02 pb^{-1}	233B 28B
07	$^{197}Au^{79+}$+$^{197}Au^{79+}$	100.0	$813\mu b^{-1}$	5.1B
08	d+$^{197}Au^{79+}$ polarized p+p $^{197}Au^{79+}$+$^{197}Au^{79+}$	100.7+100 100.2 4.6	$80nb^{-1}$ – /5.2 pb^{-1} -	160B 115B <5k
09	polarized p+p polarized p+p	249.9 100.2	14/– pb^{-1} 16/– pb^{-1}	
10	$^{197}Au^{79+}$+$^{197}Au^{79+}$ $^{197}Au^{79+}$+$^{197}Au^{79+}$ $^{197}Au^{79+}$+$^{197}Au^{79+}$ $^{197}Au^{79+}$+$^{197}Au^{79+}$	100.0 31.2 19.5 3.85	$1.5nb^{-1}$ $0.11nb^{-1}$ $40nb^{-1}$ $0.3nb^{-1}$	5.7B 0.7B 0.25B 1.7M
11	polarized p+p $^{197}Au^{79+}$+$^{197}Au^{79+}$ $^{197}Au^{79+}$+$^{197}Au^{79+}$ $^{197}Au^{79+}$+$^{197}Au^{79+}$	249.9 9.8 100.0 13.5	18/– pb^{-1} $2\mu b^{-1}$ $1.7nb^{-1}$ $7\mu b^{-1}$	 13M 2.1B/5.2B 45M
12	polarized p+p	100.2	– /10 pb^{-1}	

Info from a few sources collected on one page, with links to individual run pages

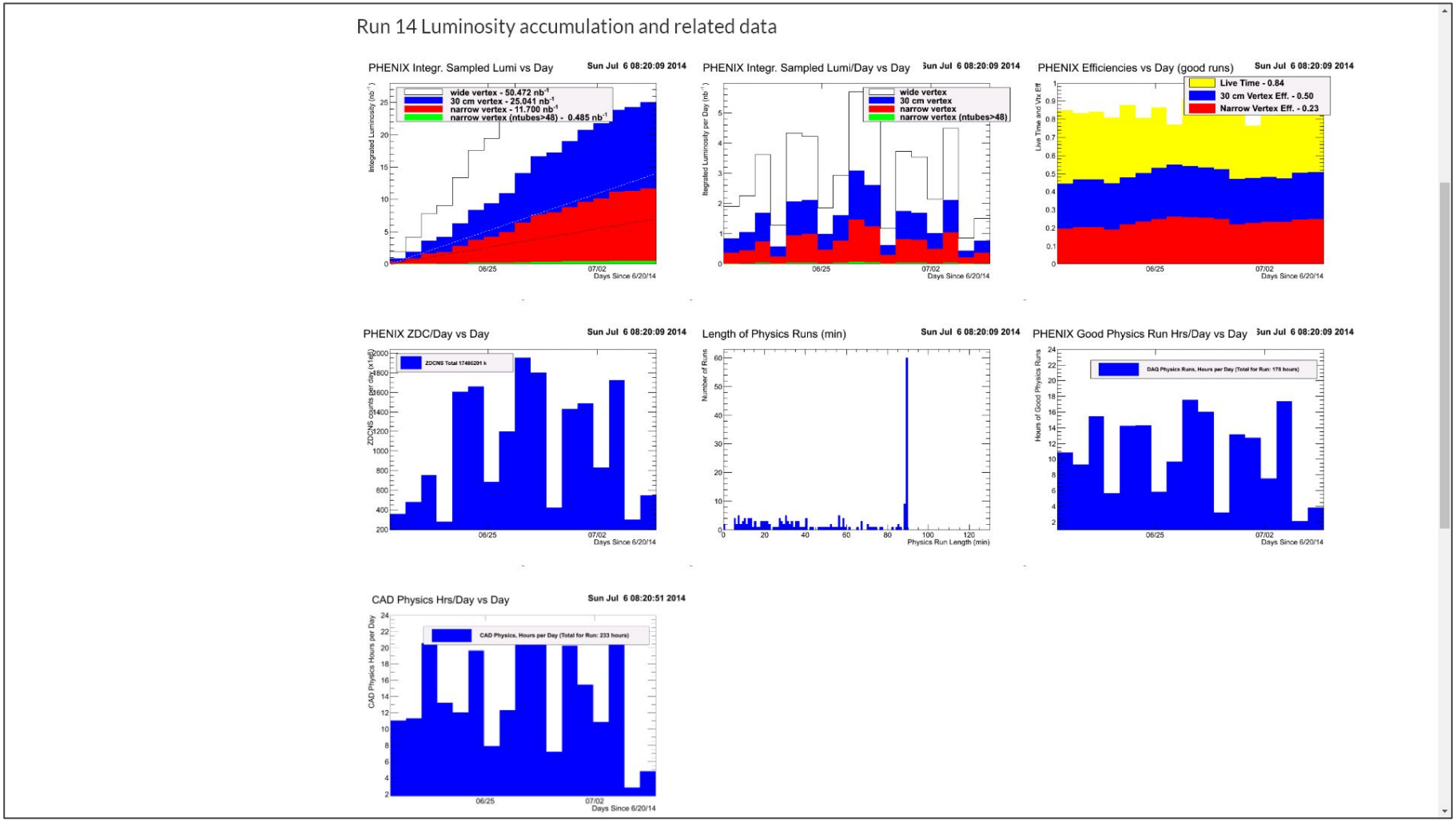
Run configuration gallery

- https://phenixcollaboration.github.io/web/detectors/run_configuration_gallery.html



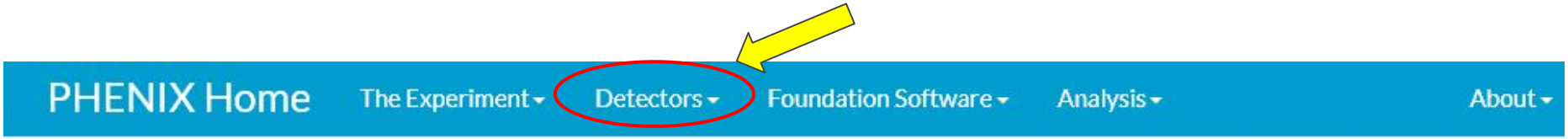
Luminosity plots harvested from run pages

- https://phenixcollaboration.github.io/web/runs/run14.html



Detector Categories

- https://phenixcollaboration.github.io/web/detectors/central_arm.html



Central Arm Detectors

Name	Role
Drift Chamber	Measures the position and momentum of charged particles
Pad Chambers	Measures the position of charged particles with precision
Ring Imaging Cherenkov	Identifies Electrons
Hadron Blind Detector	Identifies electrons without detecting hadrons
Time Expansion Chamber	Measures the position and momentum of charged particles. Identifies particles.
Time-of-Flight	Measures the position of charged particles. Identifies particles.
Aerogel Cerenkov Counter	Identifies particles at high transverse momentum.
Time-of-Flight West	High resolution particle identification.
Electromagnetic Calorimeter	Measures the position and energy of charged and neutral particles. Identifies photons and charged particles.
Silicon Vertex Tracker	Charged particle tracking near the collision vertex.

Core issues

- Potentially the amount of information to be processed for inclusion on the site is very large, so we need to define the scope
- It is up to the members of the PHENIX leadership to decide the following:
  - WHAT information should be stored and to what depth/detail
    - what are the use cases?
  - HOW it should be organized and cross-referenced
  - WHO will contribute the content and perform quality assurance of the information published on the site
  - How to best interface existing databases
    - which DBs will be kept operational long-term?

A few more screenshots

- As already mentioned, this is only a start, not much info in most categories, but some are already developed
  - e.g. the “analysis” part is not there yet etc
- Please take a look at the site and make recommendations

The Experiment Menu

PHENIX Home | The Experiment | Detectors | Foundation Software | Analysis | Resources | About

- Concept
- History
- Run Summary Table
- Auxiliary Run Info

PHENIX, the Pioneering High Energy Nuclear Interaction eXperiment, is an exploratory experiment for the investigation of high energy collisions of heavy ions and protons. It is the largest of the four experiments that have taken data at the Relativistic Heavy Ion Collider. Data-taking was finished in 2016 and the PHENIX Collaboration is now analyzing large data samples collected, prioritizing those with a unique physics reach.

BROOKHAVEN NATIONAL LABORATORY

PHENIX

U.S. DEPARTMENT OF ENERGY

Detectors Menu

PHENIX Home | The Experiment | Detectors | Foundation Software | Analysis | Resources | About

- Detectors Overview
- PHENIX Photo Gallery
- Run Configuration Gallery
- Central Arm Detectors
- Muon Arm Detectors
- Event Characterization Detectors
- Magnet

ing High Energy Nuclear Interaction
loratory experiment for the investigation
ons of heavy ions and protons. It is the
periments that have taken data at the
n Collider. Data-taking was finished in
2016 and the PHENIX Collaboration is now analyzing large data samples collected, prioritizing those with a unique physics reach.

BROOKHAVEN NATIONAL LABORATORY

PHENIX

U.S. DEPARTMENT OF ENERGY

A collection of papers in the “Resources” section

PHENIX Home | The Experiment | Detectors | Foundation Software | Analysis | Resources | About

Documents

General Overviews

- The PHENIX Experiment at RHIC - Decadal Plan 2011–2020
- PHENIX Overview (NIM A 499, 2003)

Detector Subsystems

- Calorimeter (NIM A 499, 2003)
- Inner Detectors (NIM A 499, 2003)
- Magnet System (NIM A 499, 2003)
- Muon Arms (NIM A 499, 2003)
- Central Arm Tracking Detectors (NIM A 499, 2003)
- Central Arm Particle ID Detectors (NIM A 499, 2003)
- A reaction plane detector for PHENIX at RHIC (NIM A 636, 2011)
- Design, construction, operation and performance of a Hadron Blind Detector for the PHENIX experiment (NIM A 646, 2011)
- The PHENIX Forward Silicon Vertex Detector (NIM A 755, 2014)

Data Reconstruction and Analysis

- Event Reconstruction in the PHENIX Central Arm Spectrometers (NIM A 482, 2002)

PHENIX Systems

- Online Systems (NIM A 499, 2003)

Summaries

- Spin Program Summary Table
- Dataset Summary for Runs 9-16

Summary

- We (the DAP task force) have identified a suitable and proven Web platform for Knowledge Management in PHENIX DAP
- Guidance from the PHENIX leadership is important in order to determine the scope and the content on this site
  - what level of detail
  - what use cases to prioritize
  - what data to still be kept in the databases and how to interface it
- Participation and a level of commitment of the PHENIX community in populating the new site with information will be crucial for its success. A collaborative environment (a repo on GitHub) is in place.
- Long-term document management will be a topic of a future presentation.