Data and Analysis Preservation tools for PHENIX

M.Potekhin

Nuclear and Particle Physics Software Group

PHENIX PWG Conveners' Meeting

BNL

March 11th, 2020

Data and Analysis Preservation

- We'll use the term *DAP* to refer to data **and** analysis preservation
  - cf. data which can't be analyzed is useless
- DAP is commonly described as a union of
  - Bit preservation
  - Software Infrastructure and application code
  - Analysis know-how (knowledge management)
- The goals
  - to have a reproducible analysis capability over a **long period of time**
  - implied: a **modified** or new analysis within the same framework
- There are obvious synergies with maintaining short-term reproducibility - which is required to ensure adequate quality of the analyses
- Accordingly, we are looking at solutions to cover all these items

DAP: recent history in PHENIX

- Funding agencies in the US increasingly require both new and **existing** experiments to plan/to have DAP capabilities
  - the problem: nobody in the research community has “spare cycles” while extra funding is scarce or non-existent
  - but we should still make an effort
- PHENIX has participated in the DAP workshop at CERN in late 2019
- A Lab review took place where DAP was one of the focus items
  - PHENIX DAP effort is considered a pilot project for other experiments
- A DAP Task Force has been created last year (although available effort is limited)
- In the past few months we have identified a few tools and approaches to help PHENIX address some of the DAP-related work items

DAP assets in PHENIX

- Bit preservation is taken care of (SDCC) with one tape copy of the data
- Calibration/Conditions type data reside on disk and in some cases under version control
- Production is ongoing
- Many active analyses under way, with O(10) papers published annually
- Well maintained general software infrastructure
  - build, validation and release procedures in place (thanks Chris!)
  - production jobs run in containers etc

DAP challenges in PHENIX

- But there are also challenges -
  - Software infrastructure documentation becoming (or already being) obsolete
  - Personnel (leaving or migrating to other projects)
  - Continuity of knowledge therefore is a problem (and the software effort underfunded)
  - Lack of comprehensive conditions DB in PHENIX (some data are in DB, some not) requires extra expertise from the researchers (cf. the use of custom “data artefacts” created in individual analyses)
  - Less than strict adherence to the analysis note guidance and template
    - historically, not every note contains 100% of information on reproducible steps/recipes/data products

DAP = Knowledge Management

- (Long-term) knowledge management is the key to DAP
- Keeping record of minute (but still crucial) details of the analysis procedures requires a non-trivial extra burden on the researcher
  - while getting a quality paper published is the top priority, this usually takes the back seat
- “Foundation software” e.g. PISA and fun4all need an updated and reliable documentation with working tutorials so researchers are in control of their software both in the short and in the long term
  - a review of the tutorials has started
- A lot of the “know how” is folded into the analysis software which needs to be reviewed and documented
  - workflows can be relatively complex
- There is a fairly narrow window of opportunity to address at least some of the challenges

DAP topics for today

- Focus on options and existing solutions suitable for DAP in PHENIX:
  - Curation and consolidation of mostly static documentation on a newly designed website based on a proven technology
  - A possibility to leverage an existing and powerful platform for managing papers and similar documents

- We'll leave specific current software issues for a discussion at a later date

PHENIX Web documentation

- Currently the information describing the detector, run conditions, foundation software and other topics is often fragmented and sometimes outdated (and sometime woefully obsolete)
  - e.g. kept on separate servers
  - Broken links can be really annoying
  - Mix of HTML pages, Wikis and other platforms
  - A lot of information is no longer relevant for DAP so it dilutes the content
- A large fraction of the documents exist as HTML which is difficult to edit and makes keeping consistency of references a challenge
- A large part is outside of version control
- Fixing existing websites is not practical for a number of reasons

The New PHENIX Website

- To better serve the short and long-term goals of PHENIX 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 possible
- ...but it is not optimal:
  - the material is static anyway so a database server is an overkill as a resource
  - it would require DB and other support, security updates, admin expertise etc
  - handling structured data is possible but not straightforward (see below)
  - version control is less than ideal
  - backup is possible but not straightforward
- Ideally, the long-term Web platform would combine the following:
  - static content (to simplify long-term maintenance and for better security)
  - structured data kept separately from layout (for referential integrity)
- Luckily, ready solutions do exist. Examples are in the following slides.

Example 1: HEP Software Foundation

- https://hepsoftwarefoundation.org/

HEP Software Foundation | Activities | Working Groups | Communication | Projects & Support | About

HSF
HEP Software Foundation

The HEP Software Foundation facilitates cooperation and common efforts in High Energy Physics software and computing internationally.

HSF and WLCG Workshop, Lund 11-15 May 2020 (more info)

PyHEP 2020 "Python in HEP" Workshop, Austin, Texas, USA, 11-13 July 2020 (more info)

Meetings

The HSF holds regular meetings in its activity areas and has weekly coordination meetings as well. All of our meetings are open for everyone to join.

- HSF Weekly Meeting #181, 27 February, 2020
- HSF Software Tools and Packaging Working Group Meeting #1, 26th February 2020
- HSF Weekly Meeting #180, 13 February, 2020

Upcoming HSF events »

Full list of past meetings »

From JLab to Lund

The last HSF Workshop, which we held in JLab, was a great success. Now we continue the series, joining with WLCG again for the next workshop hosted by our colleagues in Lund.

We have the pleasure of being in the beautiful main building of the university and we look forward to another very productive workshop.

Registration now open!

Activities

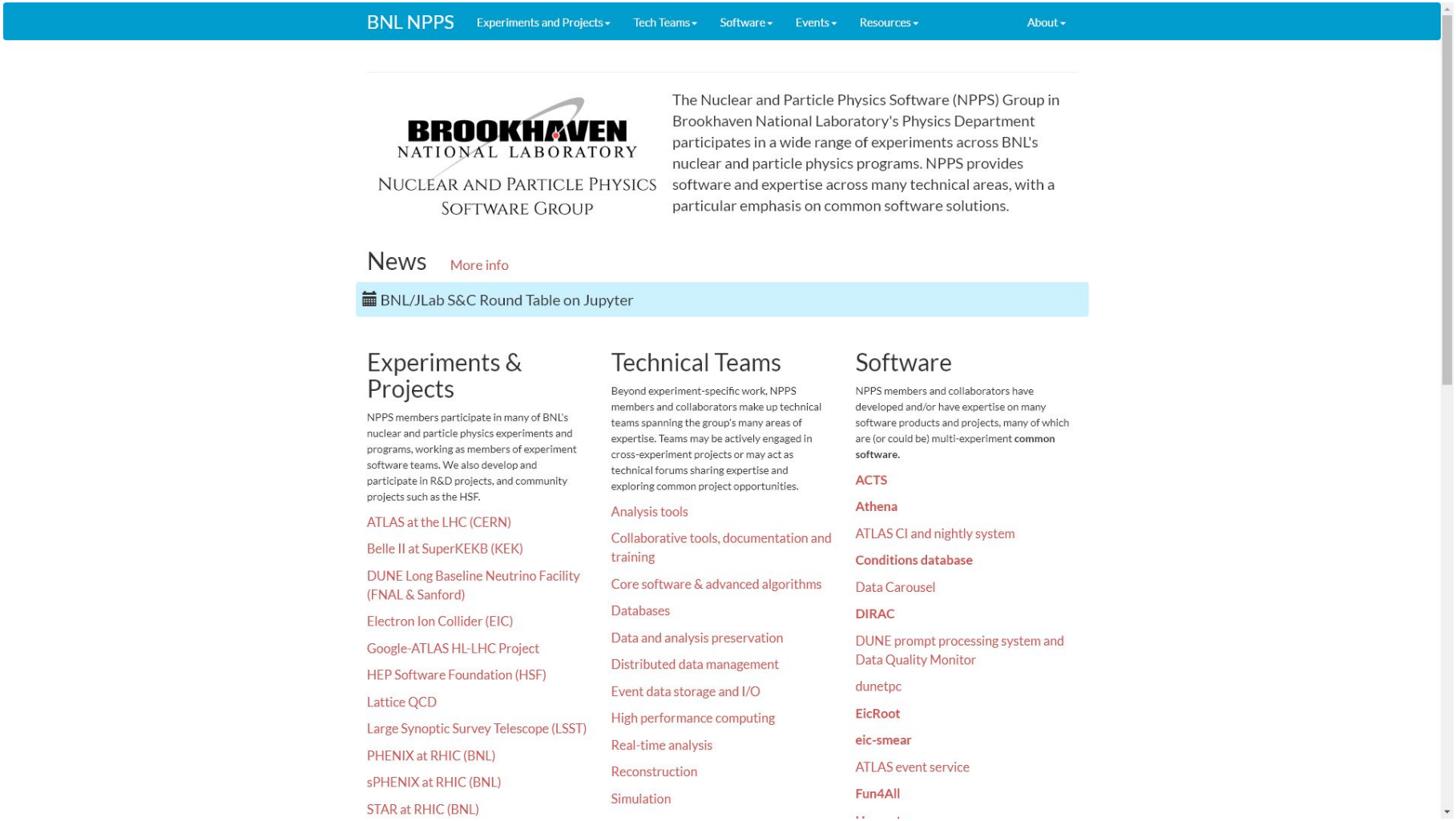
We organise many activities, from our working groups, to organising events, to supporting projects as HSF projects, and helping communication within the community through our discussion forums, technical notes and a knowledge base.

How to get involved »

Thanks to GitHub Pages, Jekyll and Bootstrap

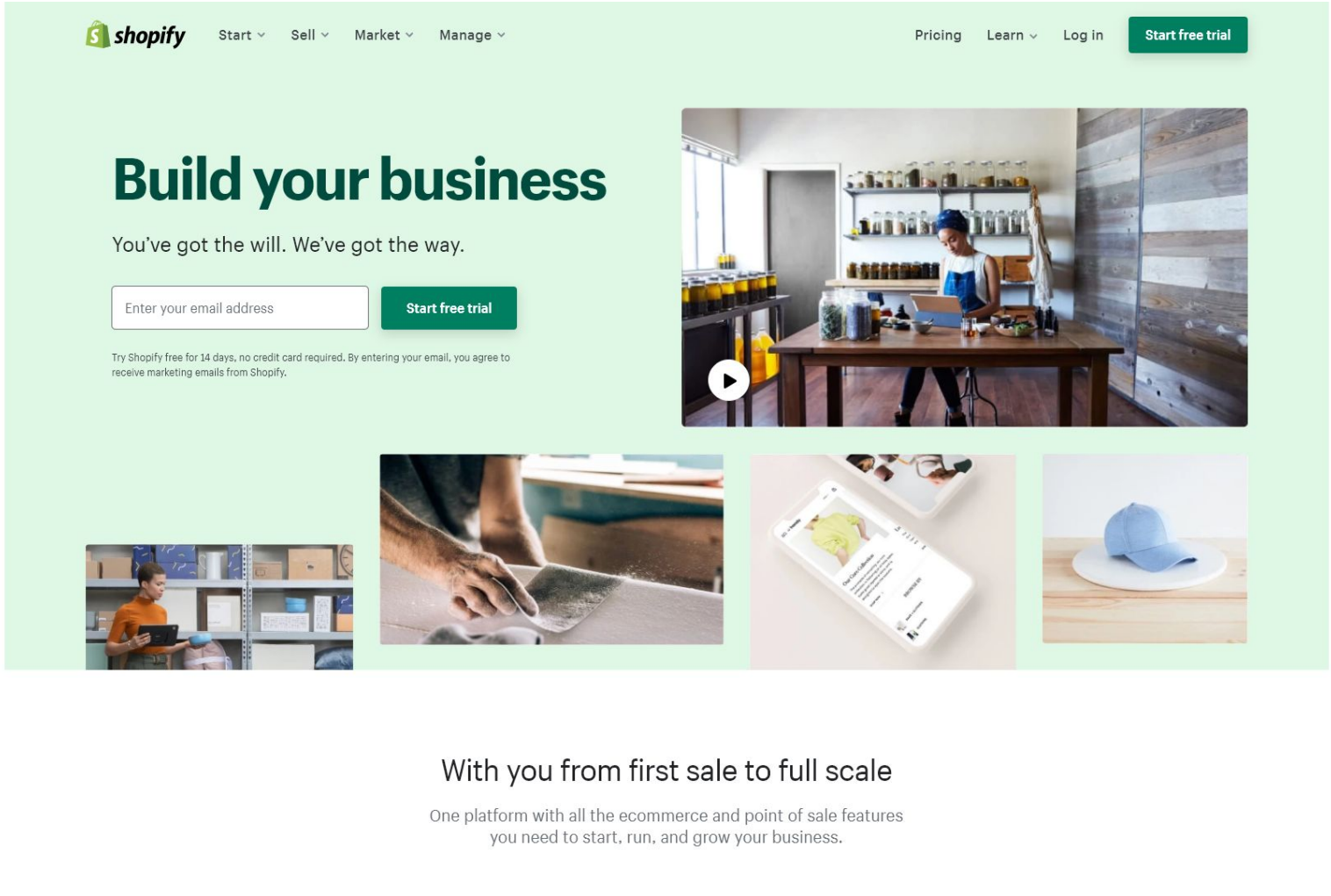
Example 2: NPPS Group at BNL

- https://npps.bnl.gov/



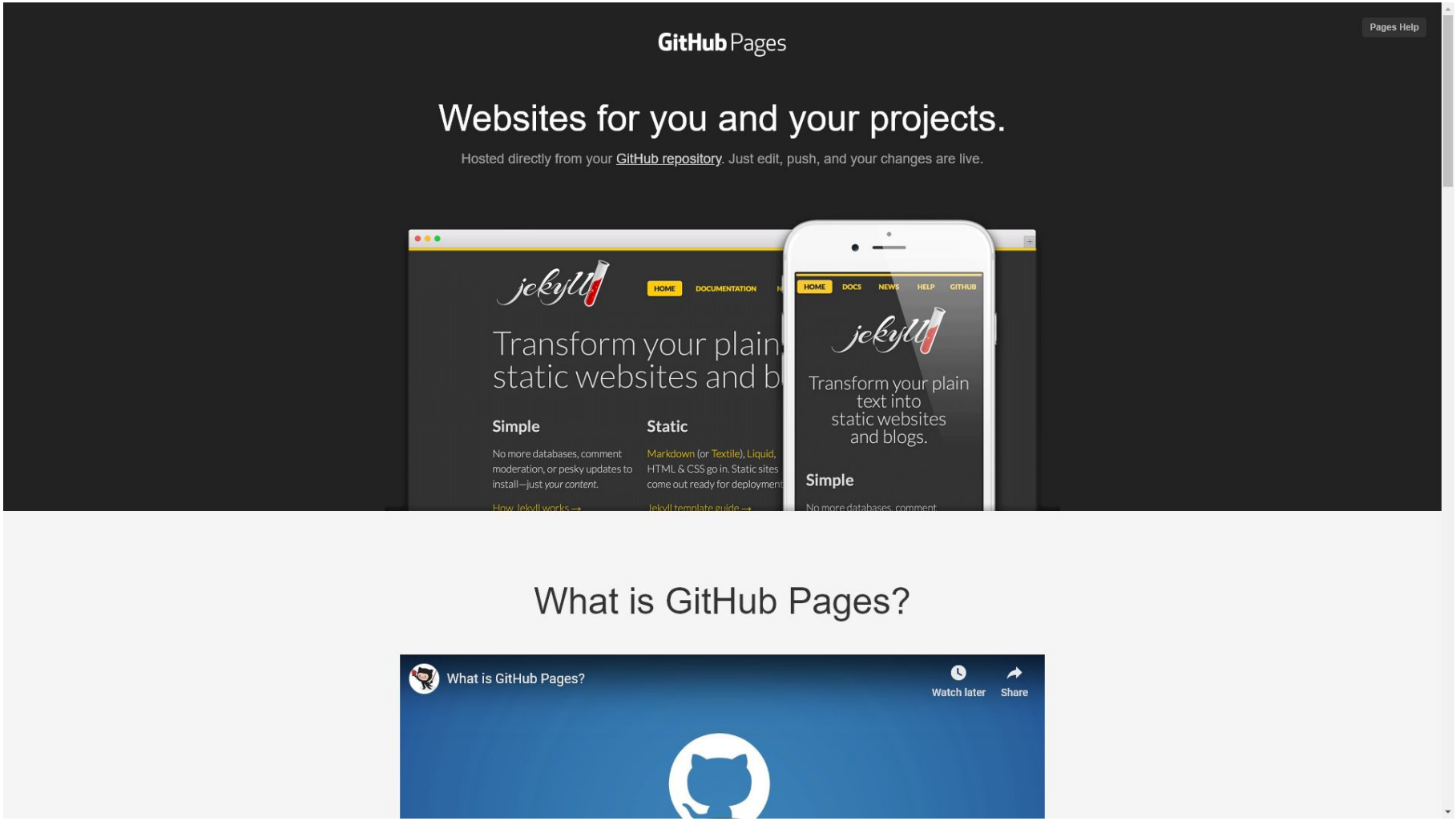
Example 3: Shopify (e-commerce)

- https://www.shopify.com/



Example 4: GitHub pages

- https://pages.github.com/



The platform

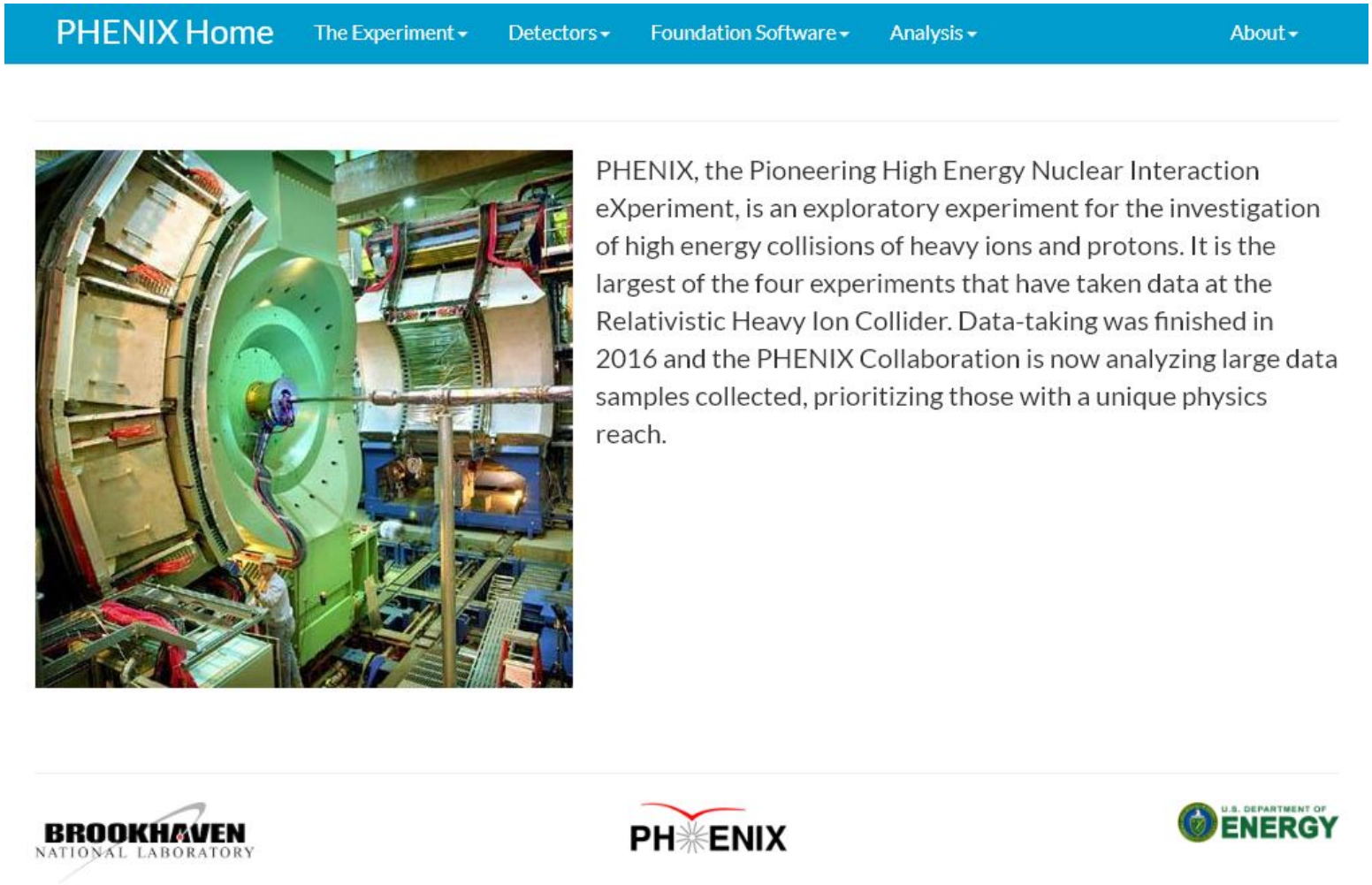
- Common in all the examples given above is the static website generator called “Jekyll” which utilizes a standard template language “Liquid” for layouts, parses Markdown syntax for content and YAML for the data component
  - YAML is essentially a human-readable variant of JSON
  - Likewise, Markdown is a lot more readable and easier to format than HTML
  - “Liquid” is quite similar to other popular template languages and easy to learn
- YAML can be embedded in the individual pages as needed
- A lot of flexibility in how the data is referenced
- 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
  - ideal 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 USB stick
- At the same time, 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 hopefully analysis know-how
  - NB. theses and analysis notes will be kept in a dedicated document management system as it is a different domain - see the second part of this presentation
- 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
- Create a basic layout for the pages and also for the data on disk (almost done)
- The most difficult part: survey, curate, improve and organize various materials and make them ready for inclusion on the new site
- Status:
  - a prototype website has been stubbed out recently
  - very little content right now but keep growing continuously
  - thanks to Gabor for collecting materials
  - your input is needed and appreciated
  - please help in gathering and reviewing materials
- This is the current location: https://phenixcollaboration.github.io/web/

The landing page

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



Summary of runs

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

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

Summary of Runs

Run	Period	Coordinator(s)
Run 01	1999-2001	Achim Franz, BNL.
Run 02	2001-2002	Anthony D. Frawley, FSU.
Run 03	2002-2003	Matthias Grosse Perdekamp, UIUC and RBRC.
Run 04	2003-2004	Terry C. Awes, ORNL.
Run 05	2004-2005	John Lajoie, ISU.
Run 06	2005-2006	Abhay Deshpande, SUNYSB and RBRC.
Run 07	2006-2007	Michael J. Leitch, LANL.
Run 08	2007-2008	Michael J. Leitch, LANL.
Run 09	2008-2009	John Haggerty, BNL.
Run 10	2009-2010	Stefan Bathe, RIKEN-BNL.
Run 11	2010-2011	Takao Sakaguchi, BNL.
Run 12	2011-2012	Xiaochun He, GSU.
Run 13	2013	Hubert van Hecke, LANL.
Run 14	2014	Klaus Dehmelt, SUNYSB.
Run 15	2015	Douglas Fields, UNM; Itaru Nakagawa, Riken/RBRC (spin).
Run 16	2016	Denis Jouan, IPNO.

Summary tables (click for larger image)

Runs 01-08 summary table

TABLE 1.1. Summary of RHIC Luminosity 2001-2008

Run	Year	Species	$\sqrt{s_{NN}}$ (GeV)	Integrated Luminosity	N_{events}	Size *of* Data Set
01	2000	Au-Au	130	$1b^{-1}$	10M	3TB
02	2001/02	Au-Au	200	$24b^{-1}$	10M	10 TB
		Au-Au	19		<M	
		p-p	200	$0.15pb^{-1}$	3.7B	20TB
03	2002/03	d-Au	200	$2.74nb^{-1}$	5.5B	46 TB
		p-p	200	$0.35pb^{-1}$	6.6B	35 TB
04	2003/04	Au-Au	200	$241\mu b^{-1}$	1.5B	270 TB
		Au-Au	62.4	$9\mu b^{-1}$	58M	10 TB
05	2005	Cu-Cu	200	$3nb^{-1}$	8.6 B	173 TB
		Cu-Cu	62.4	$0.19nb^{-1}$	0.5 B	48 TB
		Cu-Cu	22.4	$2.7\mu b^{-1}$	3.7B	20TB
		p-p	200	$3.8pb^{-1}$	85 B	262TB
06	2006	p-p	200	$10.7pb^{-1}$	233B	310 TB
		p-p	62.4	$0.1pb^{-1}$	28B	25 TB
07	2007	Au-Au	200	$813\mu b^{-1}$	5.1B	650 TB
08	2007/08	d-Au	200	$80nb^{-1}$	160 B	437 TB
		p-p	200	$5.2pb^{-1}$	115 B	118 TB
		Au-Au	9.2		< 5k	

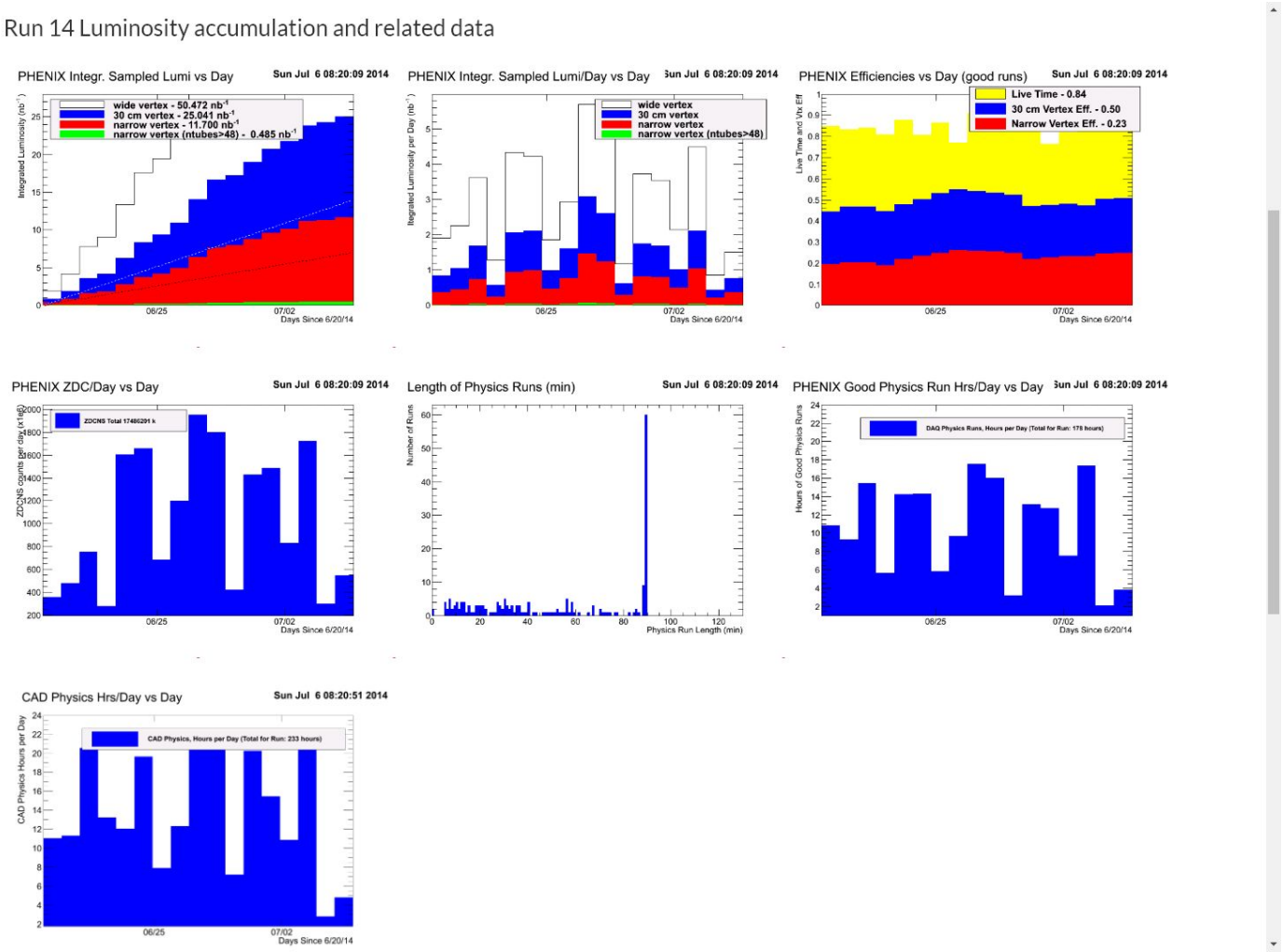
Runs 09-16 summary table

Run configuration gallery

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

Example of luminosity plots

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



Example of a page layout

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

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

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.

...and the code

Files for individual detectors are picked up automatically (according to the category) from a folder and sorted according to the weight assigned, then the page is automatically formed.

Same code is used for all detector categories.

```
# {{ page.title }}
<table WIDTH="100%">
<tr><th>Name</th><th>Role</th></tr>
{% assign items = site.detectors | sort: 'weight' %}
{% for detector in items %}
{% if detector.category == page.detector_category %}
{% include detector_category.md content=detector %}
{% endif %}
{% endfor %}
</table>
```

YAML

Example 1:

```
### RUN SUMMARY (will include reference to luminosity data etc)
- run: 01
  period: 1999-2001
  coordinator: Achim Franz, BNL.
- run: 02
  period: 2001-2002
  coordinator: Anthony D. Frawley, FSU.
```

Example 2:

```
### MAIN PHENIX PHOTO GALLERY
- path: /images/photos/3177203852_d3d30ae925_k.jpg
  title:
  gallery: main
  type: photo
- path: /images/photos/3253561842_6960470407_k.jpg
  title:
  gallery: main
  type: photo
```

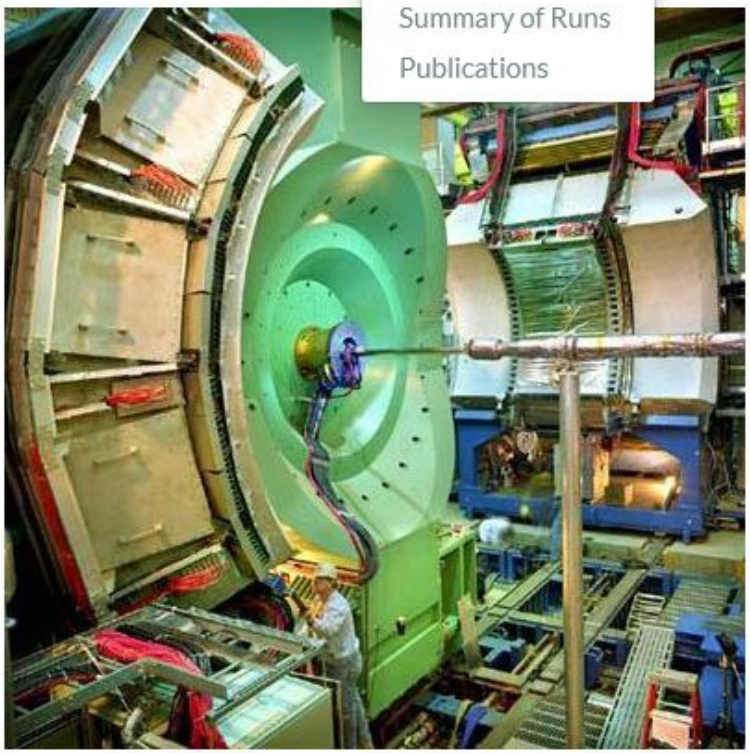
The core issue

- It is up to the conveners and other members of the PHENIX leadership to define the following:
  - WHAT information should be stored and to what depth
  - HOW it should be organized and cross-referenced
  - WHO is willing to contribute the content and perform quality assurance of information published on the site
- This is the core issue and all and any technology decisions are less important.
- Please visit the prototype site. Let us take a look at a few more screenshots in the following slides.

...back to the site: the “experiment” menu

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

- Concept
- History
- Summary of Runs
- Publications



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.

The detectors menu

PHENIX Home | The Experiment | Detectors | Foundation Software | Analysis | 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.

Another example of a detector category menu

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

Event Characterization Detectors

Name	Role
Beam-Beam Counters	Measures collision location and centrality. Starts the stopwatch for an event.
Zero Degree Calorimeters and Shower Max Detectors	Measures collision location and centrality.
Forward Calorimeters	For d+Au collisions, it measures surviving neutrons and protons from the original deuteron.
Multiplicity Vertex Detector	Measures collision location and charged particle distributions.
Reaction Plane Detector	Accurately measures the reaction plane of each collision.

The new site: issues

- The site is work in progress
- The data immediately available is incomplete e.g. configuration diagrams for some of the runs are missing; compiling various detector parameters is labor intensive
  - can the conveners help to get this effort going?
- Creating fully functional documentation and tutorial pages for PISA and fun4all will take effort
- Likewise, summarizing common patterns in the analysis techniques used in PHENIX (e.g. embedding)

Document management

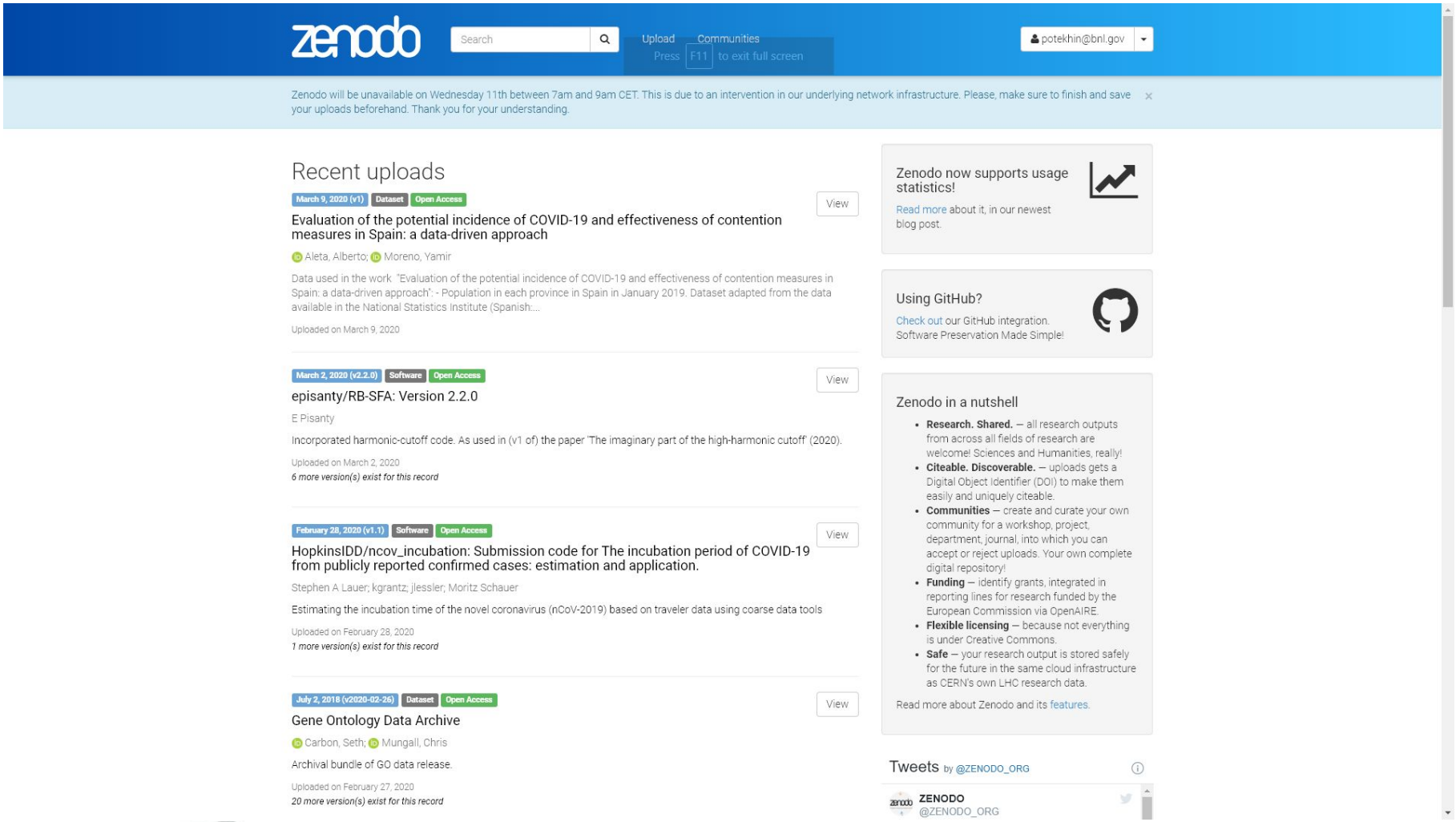
- Historically PHENIX relies on custom PHP applications to manage its analysis notes and theses materials
- These are workable solutions but we need to consider this in the longer run (e.g. maintenance, updates etc)
- Query functions can be better
- Need to revisit what information needs to be "internal" for correct classification going forward
- Alternatives include the "DocDB" system widely used at FNAL and elsewhere and more recently at BNL
  - however that system also shows its age
- All things considered, it is better to use a modern solution with wide community acceptance and which is expected to be supported for a long period of time
- We now look at the (substantial) experience with document management at CERN and in a wider community
  - BNL is an active partner in this activity
  - An integral part of the wider domain of "research data management"

Document management at CERN

- Many information systems at CERN are based on an array of services rolled out in 2002 and then called CDS, more recently rebranded as “Invenio”
  - a proven platform
  - a recent version of Invenio-based app is actively evaluated in sPHENIX
- The focus of the current development is to transform the system from a service create on site and hosted at CERN to a set of portable products
  - i.e. such that it can be installed at any site
  - due to complexity this takes time
- Next generation: Invenio RDM stands for “Research Data Management” and promises to be a “turnkey” system adaptable to many data and document management needs in many areas including those outside HEP/nuclear physics
  - while LHC is still the main driver and a leading source of funding
- Essentially a digital repository
- BNL is a partner in this development (as a test site)
  - BNL maintains a working relationship with DAP experts at CERN
- For Invenio RDM delivery is expected in late 2020 - and perhaps it will take longer for the system to become stable and ready for prime time
  - may not fit well with the PHENIX timeline
- What can we use now?
  - let’s take a look at the current Invenio-based service at CERN called “Zenodo”

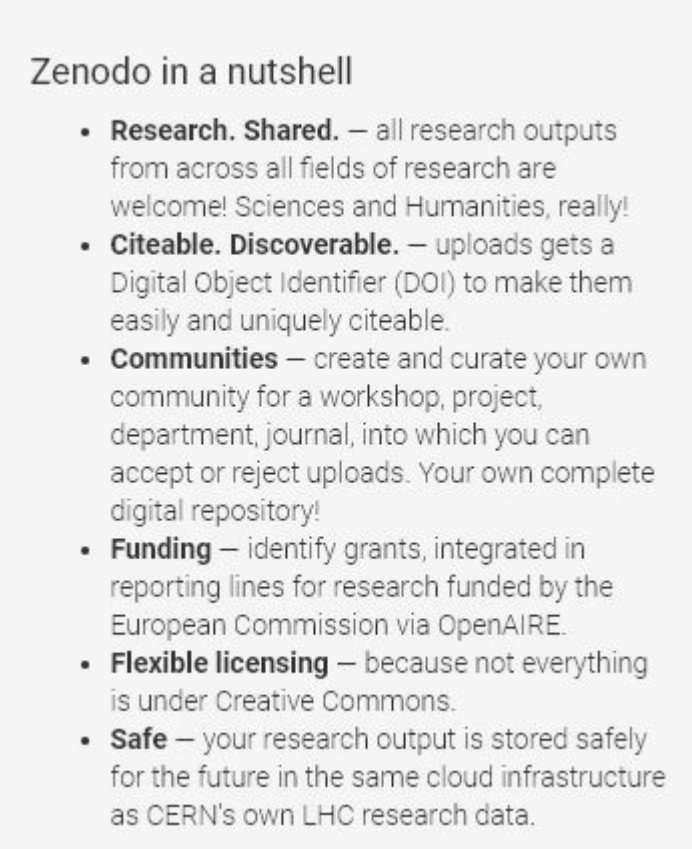
Zenodo

- https://zenodo.org/
- The name derived from Ζηνόδοτος, a Greek scholar and inventor of metadata in 280 BC



Zenodo in a nutshell

- https://zenodo.org/

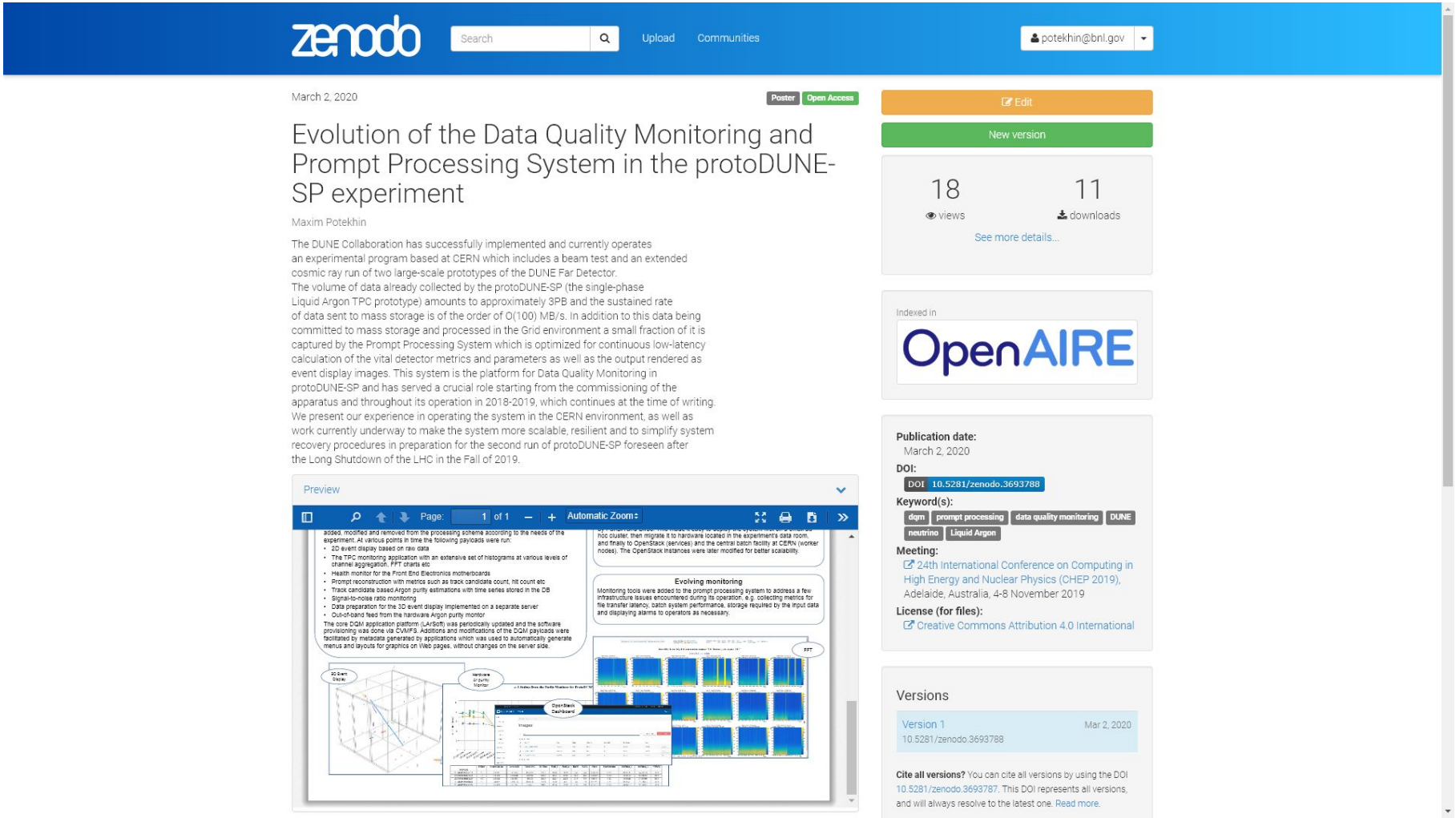


Zenodo: how do I use it?

- Pretty straightforward - upload any file and dress it up with metadata e.g. tags
  - in that regard similar to the current PHENIX notes system, sPHENIX Invenio app and DocDB (but read below about search capabilities)
- Importantly, access is controlled
  - closed (just the uploader can see)
  - restricted (a user can request access which you then explicitly grant)
  - public (self explanatory)
- Metadata can be edited at any time later but not the source file w/o changing the version
- Version support is important
- Example: in the next stilde I show how I uploaded my poster for the CHEP'19 conference and how it is managed

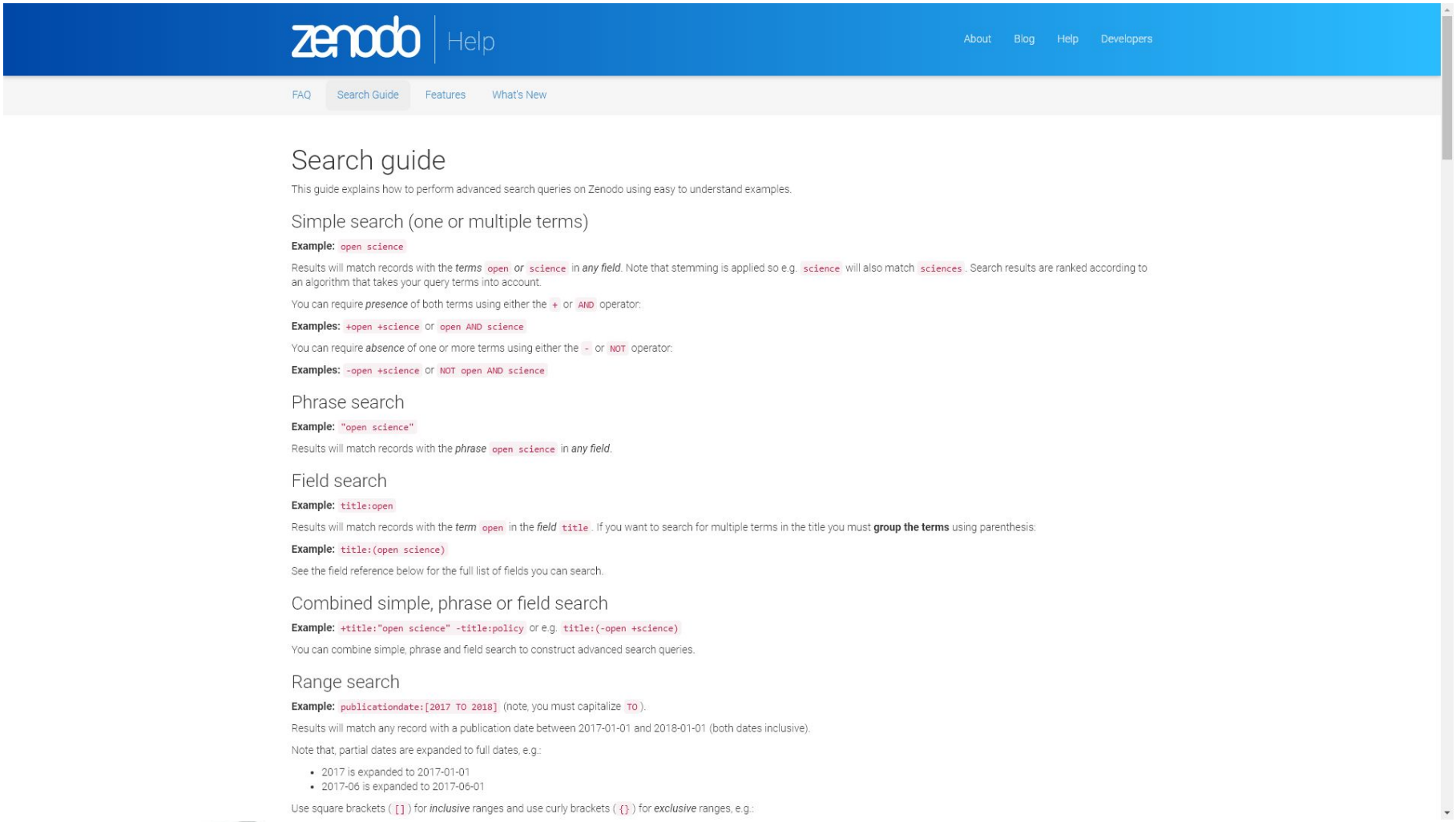
Zenodo document example

- Note tags, DOI (automatically generated), the conference link, the edit button and the version button



Zenodo search capabilities

- https://help.zenodo.org/guides/search/
- pretty advanced… elastic search (full text) is work in progress for Invenio RDM...



zenodo | Help

About Blog Help Developers

FAQ Search Guide Features What's New

Search guide

This guide explains how to perform advanced search queries on Zenodo using easy to understand examples.

Simple search (one or multiple terms)

Example: `open science`

Results will match records with the *terms* `open` *or* `science` in *any field*. Note that stemming is applied so e.g. `science` will also match `sciences`. Search results are ranked according to an algorithm that takes your query terms into account.

You can require *presence* of both terms using either the `+` or `AND` operator:

Examples: `+open +science` or `open AND science`

You can require *absence* of one or more terms using either the `-` or `NOT` operator:

Examples: `-open +science` or `NOT open AND science`

Phrase search

Example: `"open science"`

Results will match records with the *phrase* `open science` in *any field*.

Field search

Example: `title:open`

Results will match records with the *term* `open` in the *field* `title`. If you want to search for multiple terms in the title you must **group the terms** using parenthesis:

Example: `title:(open science)`

See the field reference below for the full list of fields you can search.

Combined simple, phrase or field search

Example: `+title:"open science" -title:policy` or e.g. `title:(-open +science)`

You can combine simple, phrase and field search to construct advanced search queries.

Range search

Example: `publicationdate:[2017 TO 2018]` (note, you must capitalize `TO`).

Results will match any record with a publication date between 2017-01-01 and 2018-01-01 (both dates inclusive).

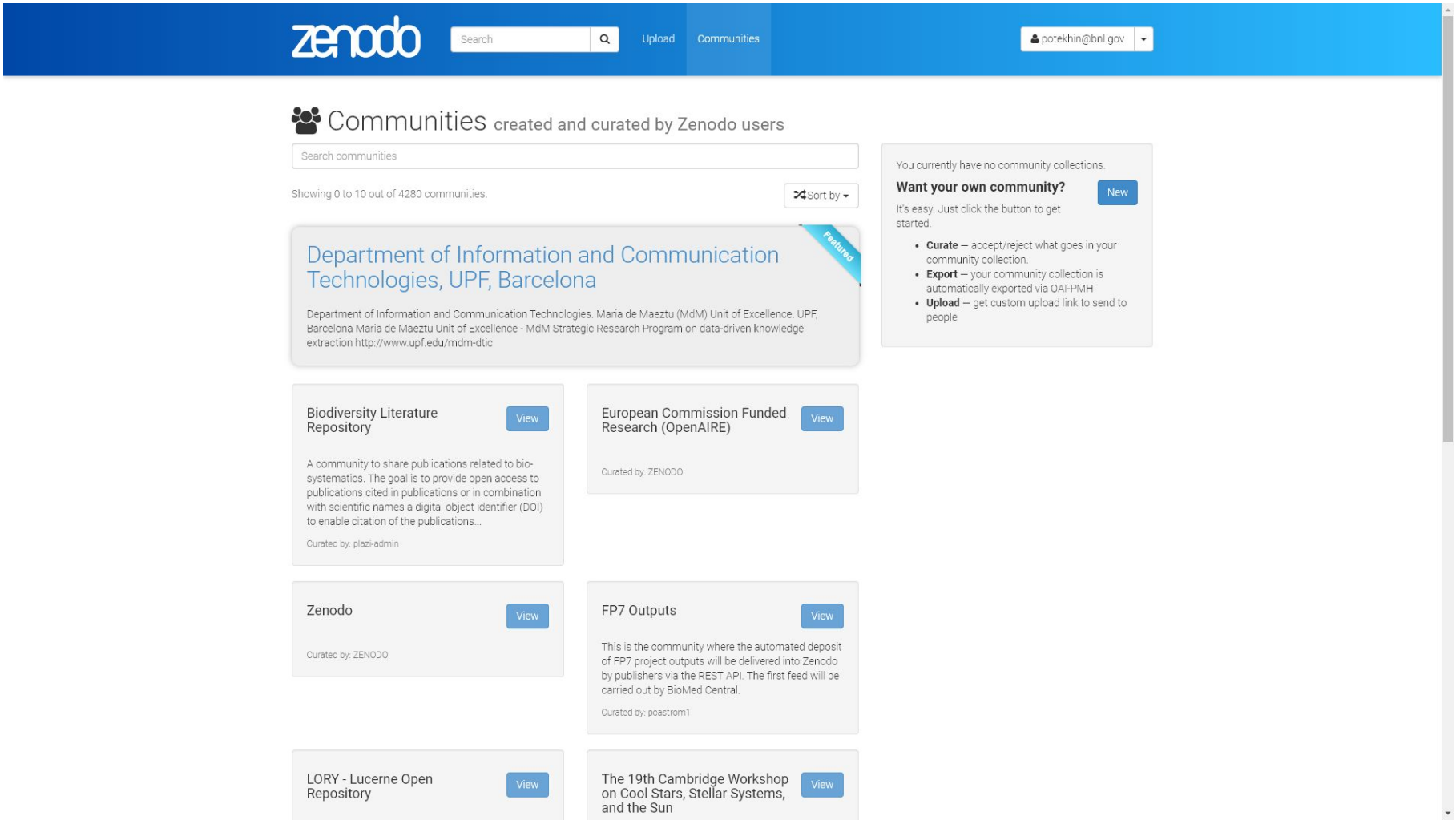
Note that, partial dates are expanded to full dates, e.g.:

- 2017 is expanded to 2017-01-01
- 2017-06 is expanded to 2017-06-01

Use square brackets (`[]`) for *inclusive* ranges and use curly brackets (`{}`) for *exclusive* ranges, e.g.:

Zenodo communities

- Users can create communities (e.g. “PHENIX”)
- Not mandatory but might be helpful to maintain a collection of documents specific to PHENIX



Zenodo: other considerations

- Integration with GitHub which make software code citeable (DOI)
- Service is free
- Information from the developers is such that it should be possible to migrate to a new repository from the same product line later (e.g. when Invenio RDM becomes available at BNL)
- The most important thing - Zenodo presents a superset of functionality implemented in DocDB (the FNAL document management system also adopted at BNL)
  - from experience, this is a good thing
- Community + Access control = Ownership
- Any type of data can be archived (datasets, ntuples) - with some volume limitations
  - this can be quite useful for some analyses
  - well aligned with common practices of CERN experiments
- Solid developer base and community involvement

Zenodo: should we use it?

- Please take a look at the system
- It's a policy decision - the capability is there
- Appears to have enough long-term viability to meet the DAP needs of PHENIX
- We don't lose control over original documents since the original files are still stored at BNL
- Appears to meet all reasonable requirements (versioning, tagging, extensive search capabilities, access control etc)
- A good pilot project would be to create a collection of PhD theses with proper tags
- Suggestions?

Summary

- We (the DAP task force) have identified a suitable Web platform in which keep long-term, static documentation such as detector and run information, software manuals and other such materials. Participation 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.
- There is a platform for managing documents which is readily available and feature-rich. It is hosted at CERN, is free and is expected to remain in operation for a long time. Using this platform (Zenodo) is a decision of policy rather than capability. Hosting PhD theses on Zenodo could be a pilot project and will create reliable references via DOI.
- Common features of the solutions above
  - they are state-of-the-art
  - they offer free and reliable hosting in foreseeable future.
- Both items have the potential to aid analyses in short and medium term as well as help PHENIX meet its long-term data preservation mandate.

Backup slides

PHENIX Analysis Notes Archive - the query page

PHENIX

Analysis/Technical Notes query form

Search Form

Download Analysis note template here. Use this template to write new analysis note for preliminary requests and final journal publiation
Use this form to search for technical and analysis notes.
Select first the desired note type either analysis notes or techinical notes. Searches may be made by author, title, note number, submission date (year) or by keyword. The default is "All" and it displays the entire list of the selected note type. Either first name or last name can be used for author. The search string for "Search by Submission Date (Year)" should be of the form "yyyy" (e.g. 2012).

Use AN Submission Form to add a new analysis note and TN Submission Form for a technical note.

Type of Note: Analysis Notes

Search by: Select from the list **Search String:**

Search by Author:
(Start typing author's name (last or first) until the desired name appears in the list and then select.)

Run Number: Select from the list

Collision Species/Energy: Select from the list

Physics Working Group: Select from the list

Analysis Type: Select from the list

Search Now Clear Form

PHENIX Analysis Notes - the archive contents

Number	Date	Title	Authors	Key Word	Links
an1425	2019-10-08	Run15 pAu identified pion and anti-proton spectra	Weizhuang Peng, Julia Velkovska	PLHF, p+Al_200GeV, Run-15, identified particles	an1425 draft an1282
an1424	2019-10-08	Run15 pAu identified pion and anti-proton spectra			an1424 draft
an1423	2019-10-08	Run15 pAu identified pion and anti-proton spectra			an1423 draft
an1422	2019-10-08	Run15 pAu identified pion and anti-proton spectra			an1422 draft
an1421	2019-10-08	Run15 pAu identified pion and anti-proton spectra			an1421 draft
an1420	2019-10-08	Jet Analysis in Run 15 p+p Collisions	John Lajoie, Milap Patel, Marzia Rosati, Jonathan Runchey	HHJ, p+p_200GeV, Run-15, jets	an1420 draft
an1419	2019-10-14	Neutral pion R_AA in p+Al, p+Au, d+Au and 3He+Au using combined Run-5, Run-8 and Run-15 \pp reference (PPG202)	Gabor David, Axel Drees, Norbert Novitzky	Heavy Ion, He3+Au_200GeV, p+Au_200GeV, p+Al_200GeV, d+Au_62GeV, d+Au_200GeV, Run-5, Run-8, Run-15, Run-14, single high pT particles, identified particles	an1419.01 draft an1152 an1269 an1270
an1418	2019-10-25	Model calculation of nuclear absorption in J/psi production at backward rapidity in PHENIX	Anthony Frawley	Heavy Ion, p+Al_200GeV, p+p_200GeV, d+Au_200GeV, Run-15, Run-14, quarkonia	an1418.02 draft
an1417	2019-10-22	Low pT Direct Photon Production in Au+Au Collisions at 200 GeV Beam Energy	Gabor David, Axel Drees, Roli Esha, Wenqing Fan, Norbert Novitzky	Photon, PLHF, Heavy Ion, Au+Au_200GeV, Run-14, direct photons	an1417.01 draft
an1414	2019-09-18	Template for PHENIX Analysis Notes	Yasuyuki Akiba, Gabor David		an1414 draft
an1413	2019-08-26	PHENIX Run14, Run15, Run16 PC2/PC3 track matching recalibration	Qiao Xu	PLHF, He3+Au_200GeV, Au+Au_14.6GeV, p+Au_200GeV, p+Al_200GeV, d+Au_62GeV, d+Au_39GeV, d+Au_20GeV, p+p_200GeV, d+Au_200GeV, Au+Au_200GeV, Run-15, Run-14, Run-16	an1413.01 draft
an1412	2019-08-02	$K*$ production in U+U at $\sqrt{sNN}$ = 192 GeV in Run12	Alexander Berdnikov, Yaroslav Berdnikov, Vladislav Borisov, Dmitry Kotov, Daria Larionova, Iurii Mitrankov	HHJ, PLHF, U+U_193GeV, Run-12, single high pT particles	an1412 draft an065 an1010 an1374 an1401 an1402 an770 an911 an964 PPG148
an1411	2019-08-02	Protons production in Run12 Cu+Au at sqrt(s_NN) = 200 GeV	Alexander Berdnikov, Yaroslav Berdnikov, Dmitry Kotov, Mariia Larionova, Iurii Mitrankov	HHJ, PLHF, Cu+Au_200GeV, Run-12, identified particles	an1411 draft an1074 an1231 an1260 an1374 an683 an814 PPG146
an1410	2019-10-24	Direct, elliptic and triangular flow of π0 → γγ in d+Au collisions at 200 and 62 GeV	Veronica CanoaRoman, Gabor David, Abhay Deshpande, Jaehyeon Do, Axel Drees, Tom Hemmick, Carlos PerezLara	Hadron, Light, PLHF, d+Au_62GeV, d+Au_200GeV, Run-16, correlations, identified particles	an1410.02 draft an1367 an1406 an1407
an1409	2019-06-24	Run14 AuAu EMCal Geometry Tuning	Gabor David, Axel Drees, Wenqing Fan	PLHF, Au+Au_200GeV, Run-14, direct photons	an1409 draft
an1408	2019-06-20	Measurement of subevent cumulant flow in Run15 p+Au and Run 16 d+Au collisions	Ronald Belmont, Qiao Xu	PLHF, p+Au_200GeV, d+Au_62GeV, d+Au_39GeV, d+Au_20GeV, d+Au_200GeV, Run-15, Run-16, correlations	an1408.03 draft an1273 PPG206 PPG221
an1407	2019-05-30	PileUp Rejection Criteria based on BBC	Veronica CanoaRoman, Jaehyeon Do, Carlos PerezLara	Global, Heavy Ion, p+p_200GeV, d+Au_200GeV, Run-15, Run-16, correlations	an1407 draft an1304
an1406	2019-05-27	Q vector calibration	Veronica CanoaRoman, Jaehyeon Do, Carlos PerezLara	Heavy Ion, d+Au_200GeV, Run-16, correlations	an1406 draft
an1405	2019-10-20	Final Results on Double Helicity Asymmetries in Charged Pion Production in Longitudinally Polarized Proton-Proton Collisions at sqrt(s) = 510 GeV	Yuji Goto, Byungsik Hong, Ju Hwan Kang, Sook Hyun Lee, TaeBong Moon, Ralf Seidl, Inseok Yoon	Spin, p+p_510GeV, Run-13, single high pT particles, identified particles, A_LL	an1405.04 draft
an1404	2019-05-08	Measurement and analysis of three-pion HBT correlations for 0-30% Centrality in 200 GeV AuAu collisions	Mate Csanad, Bálint Kurgyis	PLHF, Au+Au_200GeV, Run-10, correlations	an1404.01 draft an1187 an1244 an1288 an920 PPG194
an1403	2019-08-15	J/ψ as a function of p_T in small systems with Yue Hang Leung's Correlated Background: Run15pp and Run15pAu, Run15pAl, Run14^{3}HeAu Centrality	Matthew Durham, Anthony Frawley, Sanghoon Lim, Krista Smith	HHJ, He3+Au_200GeV, p+Al_200GeV, p+p_200GeV, Run-8, Run-15, Run-14, J/psi, quarkonia, lepton pairs	an1403.06 draft an1306 an1354 an1369 an1391
an1402	2019-04-04	$K*$ production in Cu+Au at $\sqrt{sNN}$ = 200 GeV in Run12	Alexander Berdnikov, Yaroslav Berdnikov, Vladislav Borisov, Dmitry Kotov, Iurii Mitrankov	HHJ, PLHF, Cu+Cu_200GeV, Cu+Au_200GeV, Run-5, Run-12, single high pT particles, phi	an1402.01 draft an065 an1010 an1374 an770 an911 an964 PPG148
				HHJ, PLHF, Au+Au_200GeV,	an1401.01 draft

Invenio@BNL

- Test deployment of Invenio v3 at BNL by SDCC experts (C. Gamboa)
- An application developed by BNL SDCC for one of the BNL departments
  - some specific requirements to the document workflow, not necessarily applicable to other projects
  - work in progress now to make this work for sPHENIX (custom app), discussions ongoing
- Evaluation for the BNL EIC working group is yet to happen but is planned

The role of BNL SDCC

BROOKHAVEN NATIONAL LABORATORY | Scientific Data and Computing Center

- BNL SDCC will be (already effectively is) the “bit preservation” site for PHENIX
- In recent years BNL SDCC serves as the single fully functional center of the PHENIX computing infrastructure and provides critical services to this and other experiments. At present virtually all of the stages of data processing and analysis (including simulation) take place at BNL’s SDCC
- A variety of data products necessary for operation of PHENIX software (databases, files containing conditions/calibrations data etc) are stored and managed by SDCC
  - complete removal of dependencies would be very costly
- For that reason, any DAP plan in PHENIX needs to assume that future reuse and analysis of the data will take place at SDCC
- SDCC has a versatile team of experts that help solve problems outside of core production activities
  - a recent example is installation and evaluation of Invenio-based systems at BNL