DRAFT - RHIC DAP Round Table #5

02/13/2025

https://indico.bnl.gov/event/26435/

1. Introduction

- A high-level comparison of the two CERN platforms potentially relevant to RHIC DAP was presented. Refer to slides 4 and 5 for details.
- The CERN Analysis Preservation portal (CAP) provides tools and services for documenting, sharing, and reusing analyses.
- The CERN Open Data portal provides datasets, software, and documentation to support education and researchers.
- Both sites are developed on Invenio software. CAP implements access control.
- PHENIX deposited some data along with the associated code to reproduce a photon and pi0 analysis. PHENIX emphasized the usefulness of having such a platform at BNL.
- A discussion about the RHIC public data policy has started; DOE's advice will be required.
- Christine reiterated the importance of also considering Photos and Brahms data for RHIC DAP. Questions about these collaborations being still functional were raised, which is a matter to be followed up on.

2. PHENIX requirements

Maxim presented the requirements for DAP services and infrastructure from PHENIX's perspective and experience.

- Docker images are considered the most optimum for preserving the software environment.
- Support for the various databases used by PHENIX needs to be identified.
- CVS support is required. Repository contains calibration data.
- Support for developing an accurate metadata system is required as no real system exists.
- A knowledge hosting system like OpenData.
- Reana is needed, although not critical, for DAP.

- A discussion on the metadata requirement will be followed up at the next meeting.
- Part of the PHENIX software code related to DAP is currently hosted in a private GitHub repository. A question about the possibility of making the code public was raised.

3. Developing content for analysis preservation

Gabor presented his thoughts about preserving the analysis capabilities, driven from the PHENIX experience.

- Reproducing a published, well-documented analysis is feasible; however, reproducing an analysis from general guidance is much more difficult. Let alone designing a new analysis.
- Properly documenting and organizing analysis software is key for enabling DAP.
- Preserving the capability of analysis is equally important to preserving the data for DAP. This requires the host lab to make a long-term effort to maintain the software.
- A metric for evaluating the success of DAP could be the success of repeating an analysis by someone not from the experiment (a 3rd person).
- A discussion followed regarding the PHENIX DAP experience. How challenging is it to maintain one analysis (so far)? Several factors were mentioned: the organization of software and analysis within the experiment, the lack of documentation, and insufficient sustained effort for DAP. It was suggested that PHENIX prepare a set of recommendations based on their experiences to be shared for current and future experiments, which could be presented at a future meeting.

4. CERN portals

Vincent mentioned that he is in contact with both the CAP and OpenData teams at CERN and is exploring the possibility of porting these solutions to BNL. More details will be shared at the next meeting.

5. Follow up

- Identify contacts for Phobos and Brahms collaborations
- Metadata support requirements from PHENIX
- PHENIX to produce a set of recommendations for organizing software and documentation, based on their experience, to ensure efficient DAP.
- OpenData and CAP porting at BNL status.
- CVS requirements and evolution (STAR, PHENIX)