RHIC Data and Analysis Preservation Round Table

10/02/25

Introduction – Another follow up on review report

Follow up on DAPP review

We'll come back to these topics

While the scientific ambition is clear, the evaluation highlights several critical aspects that will require careful attention. The technical foundations for software and source code preservation, and the maintaining the ability to rebuild the software stack from these sources, appear underdeveloped and must be more robustly addressed. Risk management strategies—especially in light of potential funding fluctuations and the need for prioritization—should be further elaborated. Clear engagement and alignment with the experimental collaborations are essential for the success of the initiative.

Three priorities have been identified

1. Software and source code preservation, ability to rebuild the software from sources
2. Risk management: funding fluctuations and prioritization
3. Engagement and alignment with collaborations

Gabor reminded us about these last week,
and proposed scheduling a follow-up discussion

Other topics to follow-up

- Physics case

> The proposal should more clearly and forcefully present the scientific opportunities—the "great physics case"—that justify the effort and investment. This includes demonstrating how preserved data can support innovative or high-priority physics questions.

- Connection with the EIC

> The Electron-Ion Collider (EIC) should be explicitly considered as a future stakeholder and potential user of the preserved data and tools. The proposal would benefit from identifying how it can serve EIC-specific physics goals and build early connections with the EIC community to maximize relevance and uptake.

Metadata

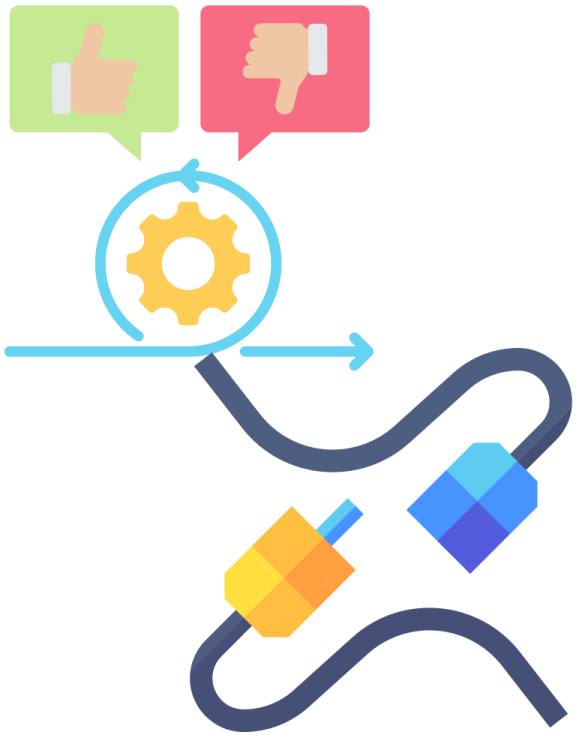
(Meta)data curation must begin as early as possible within the experiments themselves. This is a critical prerequisite for the success of the project and cannot be outsourced or delayed, as well-structured metadata is essential for both reproducibility and effective AI integration.

Vincent’s presentation later today to start the discussion

Goal for 2026: DAP integration

Analysis

Review process



Publication

DAP

The 1,000 Questions Challenge

- The chatbot was developed to help future users access experimental knowledge.
- The ChatBot has ingested thousands of documents
- Last week, Sasha presented the developments and methodology for validating chatbot performance based on expert Q&A pairs.

- **Next steps:**
  - Input from experiments: a large set of relevant Q&A is needed to build and validate the chatbot implementation

Challenge for the best set of Q&A

Today

1. About metadata – Vincent Garonne
2. Follow-up on risk analysis – Jerome Lauret

- No meeting next two weeks