

EPIC Computing & Software Task Force

Data Analysis & Preservation

The EPIC Collaboration CompSW and SimQA and the EICUG Software Working Groups are forming a joint task force to provide recommendations to the EPIC Collaboration on ensuring that Data Analysis & Preservation best practices are implemented from the start of the collaboration's activities. The discussion at the joint EPIC Collaboration and EICUG Software Working Group July 13 meeting on this topic¹ was very instructive but also indicated that a more extensive effort is needed to reach actionable recommendations.

The task force is asked to consider the following points:

- Data Analysis & Preservation in this context concerns the capturing of all aspects of scientific workflows, from data production to final publication, data archival, and potentially public release. This must however not be interpreted as a limiting definition.
- Consider which aspects of Data Analysis & Preservation can be built in from the start so they become an integral part of the collaboration's efforts.
- Consider what is needed to increase the reusability of analyses in the EPIC collaboration, that is the ability for new researchers to reuse existing components of previous analyses in order to build on them.
- Consider what the appropriate level of reproducibility is for various products the collaboration will produce, from internal reports to publications, and to what extent these apply to data, analysis, and environments.
- Consider what requirements a digital information repository for the collaboration must satisfy.
- Consider which incentives can ensure that researchers engage in Data Analysis & Preservation practices.
- Develop a **draft policy** on Data Analysis & Preservation for the EPIC collaboration, which contains the overarching philosophy and principles, but should not mention specific technology choices.
- Consider the EIC Software Statement of Principles in the development of implementation suggestions.
- Consider accessibility to the entire collaboration (and audiences beyond) in implementation suggestions.
- Provide a possible **implementation plan** that identifies possible resources (even if not presently available), a possible timeline, and a plan for management and maintenance.

¹ Indico event page: <https://indico.bnl.gov/event/15646/>, Discussion document: <https://docs.google.com/document/d/1f8ctglmbx9hzckxuudUBhN3f7jaEbzPfcUEbglDxnp/>, Live notes: https://docs.google.com/document/d/1E3J7UfVJMcSyooH-d7oVGx61qOx2mhBPYQEblW7_tI8

The task force is asked to present its findings as an EPIC internal report with an executive summary listing the specific recommendations and a presentation to a dedicated joint meeting of the EPIC Collaboration CompSW and SimQA and the EICUG Software Working Groups.