

# COHERENT COLLABORATION DATA MANAGEMENT PLAN

---

The goal of our data management plan is to ensure the reliable archival of the raw experimental and simulated data, to assure that these data are readily accessible within the collaboration, and to enable broad participation in scientific analyses and the publishing of scientific results in peer-reviewed journals. This DMP is intended to conform to FAIR (Findable, Accessible, Interoperable, and Reusable) principles.

Where reasonably possible, data necessary for validation of results are to be preserved and shared by the Collaboration. Research data generated as part of our research generally fall into categories of experimental data, simulation data, or computational analysis results. Experimental data include (but are not limited to) instrument-specific output that we store in raw form or in compressed formats using appropriate data compression algorithms, and instrument metadata which define experimental conditions (but can be stored automatically), as well as a description of algorithms and processing used to transform the raw data.

ORNL will take primary responsibility for the archival of the raw experimental data using ORNL institutional-supported production systems such as the HPSS archive system. The resources required for this project are within the scope of what is typically approved within an ORNL Leadership Computing Facility Director's Discretionary proposal<sup>1</sup>. OLCF project HEP106 has been renewed annually since 2016 and now hosts 325 Tb of raw experimental data. Simulation/analysis data include output from computational software, including (but not limited to) Monte Carlo simulations as well as tools that process and analyze simulation results. Research data do not include: preliminary analyses, drafts of scientific papers, plans for future research, peer reviews, or communications with colleagues.

Data analyses will be documented as internal research reports or published research articles, which include information on the background, description of the experimental or computational procedures, data generated from experiments and simulations, and summaries of the results. Digital data include data used to generate figures presented in publications, digital photographs, data collected through digital means, and non-digital data represented by digital information. Data and metadata (descriptions of the data, including format, processing from raw data, software used to generate data, and other relevant information) presented in publications will be made publicly accessible in digital, machine-readable format, along with necessary documentation. This will be provided in digital supplemental information for the journal, or through other public repositories where appropriate. Where no such repository exists, we will use institutional repositories where available that provide long-term preservation and sharing. In the absence of institutional options we will preserve the information locally and make it available via the project web site, as well as upon request to the project team member responsible for the data. Displayed data will be made as accessible as soon as possible to the public in accordance with the principles stated in the Office of Science Statement policies for Digital Data Management<sup>2</sup>.

COHERENT has established a public-facing web site<sup>3</sup>. A part of the content of this web site is a list of project publications and theses<sup>4</sup>, with links to their archival sources (usually the journals in which they were published). In general, we will rely on external and institutional archival repositories to provide long-term access to the project data beyond the lifetime of the project. Keepers of

---

<sup>1</sup><https://www.olcf.ornl.gov/for-users/olcf-policy-guide>

<sup>2</sup><https://www.energy.gov/datamanagement/doe-requirements-and-guidance-digital-research-data-management>

<sup>3</sup><https://sites.duke.edu/coherent>

<sup>4</sup><https://sites.duke.edu/coherent/publications/>

## COHERENT COLLABORATION DATA MANAGEMENT PLAN

---

locally preserved project data shall make reasonable efforts to maintain their archives and respond to external requests beyond the lifetime of the project.

**For each major result, COHERENT provides a data release, which is internally peer-reviewed according to collaboration Publication Board procedures.** Information on use of the data is posted on the arXiv. The data themselves are made available through ORNL OLCF, Zenodo, or as arXiv ancillary material, depending on the size of the data set.

See <https://sites.duke.edu/coherent/data/> for a list of COHERENT data releases. We note that we have received positive feedback from the theoretical community about our consistent practice in this regard, and collaborators have been actively involved in discussions of future data curation (e.g., at the Magnificent CEvNS workshop venue).

Updated August 2024