

# Charter of the Physics Working Group

Version 1.0 dated February 12, 2024; prepared by the lead and the collective of the working group.

## 1 Introduction

The Physics Working Group brings together certain projects of large research infrastructures mentioned in the document "Roadmap for Large Research Infrastructures" in the chapter on Physical Sciences and Engineering, as well as many smaller groups from various institutions engaged in research in the field of physical sciences. In both large and smaller projects, the level of data care varies currently. For instance, the research infrastructure CERN-CZ is closely linked to projects at CERN, which serve as an example for other fields in long-term data preservation, documentation, and accessibility. For some other groups, this issue is addressed less systematically and depends on the efforts of individual workers to manage data care. The EOSC-CZ projects provide an opportunity to extend knowledge and standards to other physical disciplines and groups.

## 2 Goals

Even large international projects face a shortage of experts for long-term data management. CERN has developed documents committing to making measured data accessible. All data are released with permanent identifiers. Data and related data services adhere to open and FAIR principles. Researchers and experiments are expected to develop data management plans for their research activities. Data products of varying complexity are available through the CERN Open Data portal. The development and adoption of formats, tools, and standards for working with data is a continuous effort by the community, conducted in the Data Preservation and Long-Term Analysis in High Energy Physics (DPHEP) group, the open science working group, and various initiatives at national, European, and international levels. Similar plans and commitments to open data access have been published by other institutions in the high-energy physics area, linked to national large infrastructure projects (Pierre Auger Observatory – AUGER-CZ, Brookhaven National Laboratory – BNL-CZ, Cherenkov Telescope Array CTA-CZ, Research Infrastructure for Experiments at Fermilab - FERMILAB-CZ). Even in these large institutions, the level of access to open data varies. Often, active PhD students dedicate additional efforts to create tools and document procedures for data analysis for other specialists not involved in the project, or even for the wider professional public. After their departure, it is difficult to find further enthusiasts to continue their work. Systematic positions dedicated to long-term data management are few; even fewer are focused on open data and supporting access to it. One of the goals of EOSC projects is to ensure long-term support for these tasks. Within EOSC-CZ, we aim to contribute in a balanced way to these large international projects.

A significant goal of the Physics group is to create a community where experiences and already developed procedures from more advanced groups can be transferred to groups that have not yet been able to sufficiently address the problems of long-term data storage. Some solutions suitable for large projects will be less applicable in small groups, will need to be appropriately modified, or other suitable procedures created. Individual groups will face the problem of which repository to use for long-term data storage. In some areas, established international repositories already exist, and newly produced data can be stored or at least indexed in them. For some areas, various websites allow the use of experimental data, but they are scattered in many places, not easily searchable, lacking a unified approach to data.



The goal of the physics working group will be to address this issue for several areas. We want to create conditions for making data from existing databases of large infrastructures available within the EOSC-CZ project and ensure compatibility of tools for working with data.

### 3 Outputs and Their Applications

The outputs of the Physics Working Group will be data stored in repositories and indexed in metadata databases. Other results will include websites with more detailed instructions for accessing data and specific examples of processing. The group will utilize hardware resources available in individual projects or operated within the National Repository Platform project. The outputs will serve either internally within the project for long-term data preservation for their potential use in verifying new theories or combining with newly acquired data. In line with current trends and commitments of many institutions, we expect that many data will be freely available and thus usable by professional groups and individuals who did not participate in their creation and can use them in testing new hypotheses or previously unperformed analysis. An important aspect is also the use of data for educating students and the public.

### 4 Membership and Expected Members

At the establishment of the working group, all representatives of large research infrastructures in the field of Physical Sciences and Engineering, and representatives of significant institutions engaged in research in the field of physics were approached. A wider audience was also addressed at many events organized as part of the promotion of EOSC-CZ. All interested parties could and still can join the group. Currently, the group covers representatives of large research infrastructures, several universities, and several institutes of the Academy of Sciences of the Czech Republic.

