

Grant Agreement No: 101057511

EURO-LABS

EUROpean Laboratories for Accelerator Based Sciences
HORIZON-INFRA-2021-SERV-01-07 Project EURO-LABS

DELIVERABLE REPORT

[DATA MANAGEMENT PLAN – Interim release]

DELIVERABLE: D5.8

Document identifier:	EURO-LABS_Deliverable_D5.8
Due date of deliverable:	31 August 2024
Report release date:	12/09/2024
Work package:	WP 5: Open, Diverse and Inclusive Science
Lead beneficiary:	CSIC
Document status:	Final

Abstract:

The present document reports on the Data Management Plan (DMP) providing a strategy for managing data generated and collected during the EURO-LABS project.

EURO-LABS Consortium, 2024

EURO-LABS Consortium, 2024

For more information on EURO-LABS, its partners and contributors please see <https://web.infn.it/EURO-LABS/>

The EUROpean Laboratories for Accelerator Based Sciences (EURO-LABS) project has received funding from the Horizon Europe programme dedicated to Research Infrastructure (RI) services advancing frontier knowledge under Grant Agreement no. 101057511. EURO-LABS began in September 2022 and will run for 4 years.

Delivery Slip

	Name	Partner	Date
Authored by	Antoine Lemasson María José G. Borge	GANIL/CNRS CSIC	31/08/2024
Edited by	Antoine Lemasson [Task coordinator]	GANIL	31/08/2024
Reviewed by	Antoine Lemasson [Task coordinator]	GANIL	31/08/2024
Reviewed by	María José G. Borge [WP coordinator]	CSIC	31/08/2024
Approved by	Navin Alahari [Scientific coordinator]	GANIL	12/09/2024

TABLE OF CONTENTS

1 INTRODUCTION.....	5
2 DATA SUMMARY	7
2.1 DATA PRODUCED WITHIN EURO-LABS WORKPACKAGES.....	7
2.1.1 WP 1.....	7
2.1.2 WP 2.....	8
2.1.3 WP 3.....	8
2.1.4 WP 4.....	9
2.1.5 WP 5.....	9
2.2 DATA FROM RESEARCH INFRASTRUCTURES	10
3 FAIR DATA.....	10
3.1 MAKING DATA FINDABLE, INCLUDING PROVISIONS FOR METADATA	10
3.2 MAKING DATA ACCESSIBLE.....	10
3.3 MAKING DATA INTEROPERABLE	10
3.4 INCREASE DATA RE-USE	10
4 DATA SECURITY	10
5 ETHICS.....	11
ANNEX: REFERENCES	12
ANNEX: GLOSSARY	12

Executive summary

The present Data Management Plan (DMP) is expected to provide a strategy for managing data generated and collected during the EURO-LABS project and optimize access to and re-use of research data. This data management plan was prepared following the guidelines on FAIR Data Management in Horizon Europe.

This document provides an overview of the types of datasets that are collected or generated by each work package and how each of the work packages will make these datasets findable, accessible, interoperable and reusable (FAIR). During the course of the project, work package partners are expected to collect and generate diverse datasets related to the activities of the nuclear and particles physics members of this consortium.

Data interoperability in Accelerator Based Nuclear Physics is one of the objectives of the EURO-LABS and cross fertilization of software as well as datasets will be encouraged during the project.

Within EURO-LABS, the following categories of data to be managed are foreseen:

- 1) Data generated by the project work packages*
- 2) Data collected at Research Infrastructures (RIs) with the Transnational Access support of EURO-LABS*

Data from (1) are processed according to best practices already in place and reported on in the present DMP.

Data from (2) are the responsibility of the research infrastructures. EURO-LABS encourages good research data management and adoption of FAIR principles.

This DMP also specifies the recommended licensing schemes of the data produced. EURO-LABS complies with the EU and national regulations on data handling and publishing. The project carefully addresses privacy and copyright issues prior to publicly releasing any data. All users have been made aware of the General Data Protection Regulation for Personal Data (GDPR).

The document includes the following sections: Section 1 gives a general introduction to the EURO-LABS Project. Section 2 elaborates on the data generated by the work packages during the lifetime of the project. Section 3 provides a FAIR analysis on the data generated, while Section 4 approaches the IPR uses and finally section 5 concludes this deliverable.

1 INTRODUCTION

EURO-LABS is a network of 33 research and academic institutions (25 beneficiaries and 8 associated partners) from 18 European and non-EU countries, involving 45 Research Infrastructures within the Nuclear Physics, Accelerators and Detectors for High Energy physics pillars. In this large network, EURO-LABS ensures diversity and actively support researchers from different nationalities, gender, age, and variety of professional expertise.

EURO-LABS aims at fostering the sharing of knowledge and technologies across scientific fields to enhance synergies and collaborations between the RIs of the Nuclear and High Energy communities. The main goals of the project are:

1. Provide efficient transnational access to the available resources at a major fraction of EUROPEAN Laboratories for Accelerator Based Sciences (EURO-LABS) within a network including the major European laboratories
2. Enhance collaborative targeted improvements for the existing services, that will lead to an increase of the scientific and technical opportunities at various RIs
3. Make the results from the studies conducted at the RIs of EURO-LABS during the period of the project available to the scientific community and manage the experimental data, when relevant, through a Data Management Plan (DMP) in line with the FAIR principles (Findable, Accessible, Interoperable, Reusable).
4. Organize the training of the new generation of researchers and young technical staff to best exploit the RIs, through workshops and hands-on experience at specifically chosen RIs

EURO-LABS activities are organized in the following five work packages

- **WP1: Project management and coordination**

WP1 ensures the effective management of the project, the timely flow of information and availability of project documentation to the beneficiaries, the Governing Board and the Steering Committee.

- **WP2: Access to Research Infrastructures for Nuclear Physics:**

WP2 provides Transnational Access (TA) to various Research Infrastructures (RIs) providing stable ions, radioactive ions and neutrons at various energies, combined with state-of-the-art equipment to explore nuclei under extreme conditions. Virtual access (VA) to a theoretical repository are also provided, to access model calculations, both for planning and interpretation of experiments conducted at the TA facilities. The broad access offer provided to users of EURO-LABS represents a crucial support to address emerging new challenges in the fields of nuclear and HEP.

- **WP3: Access to Research Infrastructures for Accelerator R&D:**

WP3 provides TA to a broad spectrum of installations, to test concepts for future accelerators, based on improving the present facilities, and for R&D studies for future colliders like CERN's FCC or the Muon Collider.

These facilities provide beam lines for testing advanced accelerator materials, superconducting or normal Radio-Frequency cavities, magnets and acceleration schemes. These tests use different particles and energies (low-energy protons, low-energy electrons, ultra-soft electron bunches and high-intensity high-energy electrons) and could also have connections to industrial applications.

- **WP4: Access to RI for Detectors R&D**

WP4 provides TA to various facilities having energetic beams (protons, mesons, muons, electrons) and irradiation facilities. These measurements are necessary to study the effect of detectors and associated equipment in-beam, required for the detector upgrades (for operation at the HL-LHC) and the construction of new detectors in the future.

- **WP 5: Open, Diverse and Inclusive Science**

WP5 enhances the users diversity (Task 5.1), brings the nuclear physics community into the EOSC (European Open Science Cloud) framework, develops services to enhance FAIR (Findable, Accessible, Interoperable and Reusable) data principles (Task 5.2), promotes the use of machine learning (ML) methods to improve beam quality, transport efficiency and accelerator reproducibility (Task 5.3), and trains young generations to make the European facilities better performing and more competitive (Task 5.4).

2 DATA SUMMARY

2.1 DATA PRODUCED WITHIN EURO-LABS WORK PACKAGES

Various research activities undertaken in the framework of the EURO-LABS project are expected to generate data which are collected within the following context:

- Data collected via online surveys for the definition of user needs: these data and their metadata are processed and retained via trusted providers, controlled by consortium partners. The data privacy policy is available in the project's website.
- Interviews held for user requirements, minutes and notes of meetings, validation workshops, focus group meetings and webinars: these involve documents held and shared among the project partners according to their institutional practices and on the project's collaborative platform.
- Project dissemination material: documents and publications presenting the findings, analyses, training material, tutorials and visualizations of the project, and text and images on project website and social media channels.
- Project website and the information contained therein.
- Deliverables and financial reports produced by the project: documents required by the project funder (European Commission) presenting the progress of the project.
- Software produced with the support of the project
- Databases designed with the support of the project
- Consortium meeting notes, financial information, and other internal administrative data in common document formats (spreadsheets, text documents).
- User assessment results gathered from a targeted user base of stakeholders actively engaged in the process, in the form of answers to questionnaires and notes from interviews

Below we list documents, software and specific data produced within the scope of EURO-LABS. In addition, the cases of data produced at RI supported by EURO-LABS TA and VA are described in section 2.2.

2.1.1 WP 1

The project office generates and manages the archive of all open-access deliverables of EURO-LABS. They are available through a dedicated web page and when applicable through sub-sets of metadata that referring to partners and/or work packages that have been major contributors/authors. The project office, in collaboration with the coordinators of WP2-4 and all the facility coordinators, also manages the nominative list of people benefiting from the TA offered by the project.

A harmonized user management is used throughout the project together with the exploitation of common best practices: the large number of users that are expected to apply for TA requires a proper management of their data among the facilities, in compliance with General Data Protection Regulations. A database to manage the information provided and the feedback by the users (including beam conditions, sample characteristics, dose, fluence, other environmental information, etc.) has been established to achieve standardization and to improve the level of quality assurance.

Information is gathered in a structured manner through forms to be filled at the end of TA periods, surveys and interviews to the experimental teams, taking into account users' needs and suggestions, then organized and presented in a comprehensive online portal.

The project office takes care of preparing the forms, surveys and interviews.

2.1.2 WP 2

WP2 generates and manages the archive of TA and VA access to Research Infrastructures for Nuclear Physics.

Further, under Task 2.5 a database and webpage of relevant information on the participating RI has been developed and linked with the EURO-LABS website. The collected data include:

- details of the facility (name, address, contact persons)
- list of available infrastructures and beams (with links to the technical details)
- dates of call for proposals, deadlines and details of submission and dates of PAC meetings
- total amount of access units used and still available (beam hours, visits, person-days)
- access procedures (forms to submit)
- details of the TNA support (what is offered) and details of the application
- list of publications.

At the production stage, the website is hosted by the EURO-LABS servers.

Under Task 2.4, the Theo4Exp Virtual Access provides access to theoretical tools for the EURO-LABS project as well as for the wider experimental nuclear physics community. The users access the infrastructure by using EURO-LABS Authorization and Authentication service (WP5.2). Scientific papers and publications produced by the use of Theo4Exp will be available in open access and will acknowledge EURO-LABS. The link to the different publications will be listed in the Theo4Exp webpage (<https://institucional.us.es/theo4exp>).

2.1.3 WP 3

WP3 generates and manages the archive of TA access to Research Infrastructures for Accelerator R&D.

Further, under Task 3.3 an integrated simulation and measurement framework will be developed so that users and operators can prepare, plan, perform and evaluate experiments more efficiently. This integrated simulation and measurement framework will be implemented based on B2SHARE - EUDAT services and will be tested at the KIT (Karlsruher Institut für Technologie) facilities KARA and FLUTE. Different tools are considered for the code and measurement data handling to make them publicly available. The evaluation of these tools and the implementation is part of the project. Part of this evaluation is the choice of the correct licenses to respect the local, national and EC rules on public data and code. The measurement data together with the required metadata will be made public utilizing EUDAT [1] or RADAR [2]. The source code developed within the project will be hosted in the public KIT Gitlab instance [3] or on Github [4]

2.1.4 WP 4

WP4 generates and manages the archive of TA access to RIs for Detectors R&D.

Under Task 4.3.1, a central database of irradiated material (traced, handled, stored) is deployed using section services currently offered by the installation. The data for IRRAD are presently stored and handled in the IRRAD data manager software tool. The tool is accessible to the users of the facility and for their experiment only (under SSO login). Further, the various software tools used is now upgraded to the requirements of the CERN OC11 [5] about the handling of personal data (Record of Processing Operations, disclaimers).

Under Task 4.4.1, a database handling of beam time and irradiation requests has been deployed. The outcome of the database tool is: user schedules and statistics, they are made publicly available on the CERN PS/SPS user webpage. Users will require a CERN sso (“single sign on”) account to reach the database directly.

Under Task 4.4.5, Software Development for a Cooling System and Graphical User Interface for EMC test station has been carried out. The software is stored in the ITAINNOVA database and it is available on the public repository of the EURO-LABS project. In relation to the storage of data during the test campaigns, a copy of test data will be sent to the users. A copy of these data will be kept at ITAINNOVA. The IT department of ITAINNOVA will handle this data to secure the information for long period.

2.1.5 WP 5

The data generated within the subtasks of the WP5 are listed below.

Diversity & Dissemination

- Project public website
- Intranet and collaboration workspaces
- EURO-LABS newsletter circulated to project members and to a wider community
- Modern media channels, like social media, YouTube
- Project mailing lists
- Videos of the various RI

Open NP: A portal for NP Data and tools

- Open NP Catalogue/Database of experimental datasets (Software + DataBase). The catalogue will consist in a public website based on existing data management software platforms.
- Platform to access data in RI including Authentication and authorisation methods to access data based on centralized identification authorities. The platform will be connected to EDUGAIN [6] identity federations service.

Machine Learning for Beam Diagnostics:

- virtual accessible beam diagnostic data base
- optimizer toolkit for accelerator experts and users deposited on public repositories.

2.2 DATA FROM RESEARCH INFRASTRUCTURES

Transnational Access service program within the EURO-LABS project does not produce directly data, only through the user teams accessing the participating RIs. The beneficiaries of EURO-LABS assure that the teams timely publish any scientific results in conference proceedings and scientific journals as appropriate, making them available in open-access archives. All work done and presented will acknowledge the project. For a broader impact, the EURO-LABS website contains and displays all scientific events and publications, including those reaching out to society and industry.

Datasets collected at research infrastructures with the TA and VA support of Work packages 2, 3 and 4 are managed according to the policy of the facilities. Research Infrastructure receiving the support from EURO-LABS are expected to provide a Data Management Plan for the data produced in their facility and follow the FAIR principle.

Facilities could request guidance from the WP 5.2 to establish their data management plan.

3 FAIR DATA

3.1 MAKING DATA FINDABLE, INCLUDING PROVISIONS FOR METADATA

Data and software codes related to EURO-LABS research activities are accompanied by standard general metadata and receive a DOI (Digital Object Identifier) number. The outputs are stored in the zenodo repository (<https://zenodo.org/communities/euro-labs>).

Access to the project's output is provided through the EURO-LABS project's site. Scientific publications will be also disseminated through the project website (<https://web.infn.it/EURO-LABS/>).

3.2 MAKING DATA ACCESSIBLE

Depending on each deliverable's dissemination level, deliverables are publicly available via the project's website, or be restricted.

Scientific papers and publications will be mainly available via Institutional Repositories.

EURO-LABS promotes 'Open Access' thus a machine-readable electronic copy of every publication is expected to be deposited in a suitable Open Access repository.

Software will be deposited in open repositories.

In the case of surveys containing any kind of personal data, only anonymized versions of the datasets will be openly available; raw data will be archived according to the regulations of each institute.

3.3 MAKING DATA INTEROPERABLE

Interoperability of the data collected by surveys, deliverables and publications will be achieved by the standard metadata required by OpenAIRE [6]

3.4 INCREASE DATA RE-USE

The data generated within the project will be made available under Creative Commons attribution license CC BY 4.0, unless otherwise agreed due to well justified reasons.

4 DATA SECURITY

Deliverables, reports, or any other produced document will be primarily stored in the project's collaboration online server hosted by INFN Bologna that adheres to all common security IT standards thus minimizing the risk of unintended exposure to non-authorized users.

5 ETHICS

Only anonymized data will be made openly available when datasets contain any personal data. However, if certain datasets contain data under specific restrictions that prohibit further data dissemination, they will be prevented from being publicly available either by requiring only authenticated and authorized access or by imposing Access Lists (ACLs) to the API endpoint(s) serving them.

ANNEX: REFERENCES

- [1] EUDAT: <https://www.eudat.eu>
- [2] RADAR: <https://www.radar-service.eu>
- [3] KIT Gitlab: <https://git.scc.kit.edu>
- [4] KIT Github : <https://github.com/KIT-IBPT/>
- [5] CERN OC11 The processing of personal data at CERN : <http://cds.cern.ch/record/2651311>
- [6] OpenAIRE : <https://guidelines.openaire.eu/en/latest/>
- [7] EDUGAIN federation : <https://edugain.org/>

ANNEX: GLOSSARY

Acronym	Definition
DMP	Data Management Plan
IPR	Intellectual Property Rights
GDPR	General Data Protection Regulation for Personal Data
RI	Research Infrastructure
TA	Transnational Access
VA	Virtual Access