Data xx/xx/2024

LNS Data Management Plan

Written by	Checked by	Approved by
M. La Cognata (CSN3 coordinator)	S. Gammino (Director) A. Tumino (Head of the research division) V. Greco (Deputy director) E. Giorgio (Responsible for the IT Service) G. Riccobene (CSN2 coordinator) D. Gambacurta (CSN4 coordinator) G. Torrisi (CSN5 coordinator)	

Subject

Data Management Plan for research activities making use of Laboratori Nazionali del Sud (LNS) facilities.

TABLE OF CONTENTS

1 INTRODUCTION

The LNS Data Management Plan (DMP) concerns all research data collected in experiments performed by using facilities available at LNS.
The aim of the DMP is to plan the life cycle of data. Inspired to the FAIR (Findable, Accessible, Interoperable, Reusable) principles, it offers a long-term perspective by outlining how data will be collected, documented, shared and preserved and describing the process rules, the management workflow, roles and responsibilities of involved parties.
The various scenarios for the use of the facilities available at the LNS are defined in the third section. Methods for storing and ensuring the long term durability of data are described in later sections.

2 DEFINITION OF TERMS

- EXPERIMENTAL ACCOUNT: Directory created by the Information Technology (IT) service containing the data record for an experiment and therefore all the data linked to the experiment.
- DATA MANAGER: Person responsible for the data. Generally, the spokesperson of the experiment.
- DIGITAL OBJECT IDENTIFIER (DOI): Unique, long-term identifier allowing the identification of a data record. This identifier will be created by Open Science Working Group as routinely done for the sharing of products on the Open Access Repository (OAR).
- EMBARGO PERIOD: Period during which the data are available only to the experimental team. Beyond that period the data must be open to the widest audience.
- DATA STEWARD: Member of LNS IT service that is in charge of the data and metadata collection from the data manager, their transmission to the data repository, and handles the data access requests.
- RAW DATA: all kinds of data collected by experiments carried out by using LNS infrastructures.
- METADATA: all information necessary to manage and perform the analysis of the raw data, including (but not limited to) the context of the experiment, the experimental team, the experimental conditions, the data format, the logbook, software package, etc.

- SPOKESPERSON: person responsible for the experiment, identified on the scientific proposal submitted to the LNS Program Advisory Committee, and for the data management collected during the experiment.
- EXPERIMENTAL TEAM: experimental group which includes the spokesperson and any other person to whom the spokesperson designates the right to access resultant raw data and associated metadata.
- OPEN ACCESS: data accessible, upon request, by the community at large, and protected by open license.

3 DATA COLLECTION AND STORAGE

LNS are engaged in research activites making use of several facilities devoted to nuclear physics, nuclear astrophysics, development of innovative detectors, material science, applications of nuclear techniques and astroparticle physics. Hence various uses are possible and thus require different data management scenarios described below and shown in the final block diagram:

- Data obtained with LNS instrumentation and stored using LNS owned resources (Data acquisition system and servers for storage).
- Data obtained with LNS instrumentation which use a data acquisition and storage system that does not belong to LNS. However, data sharing and durability will be guaranteed by LNS by keeping a copy of the data generated which has to be provided by the spokeperson of the experiment.
- Data generated from instruments belonging to a collaboration which uses its data acquisition system and LNS servers for initial storage. This is typically the case of itinerant detectors.
- Data generated from instruments belonging to a collaboration with acquisition and servers outside LNS. The initial data storage does not use resources belonging to LNS. In this case as well, LNS will keep locally a copy of the data (provided by the spokesperson) as a guarantor of the data produced using LNS research infrastructures.
- Data generated during commercial use of LNS research infrastructures belong to the client and will not be stored on local servers unless agreed otherwise.

4 COMPATIBILITY WITH COLLABORATIONS DMP

Often, large collaboration have already a DMP, a specific computational framework, or international agreements such as MoUs. In such cases a specific agreeement between LNS and the collaboration has to established prior to PAC approval or data taking (if no PAC evaluation is foreseen) on a case-to-case basis to preserve the peculiarities of the collaborations, especially in those cases where a large number of partners are involved and specific agreements between international institutions are already present.

5 OWNERSHIP OF DATA

INFN is the owner and the custodian of the raw data (and associated metadata) produced by using its research infrastructures, unless otherwise agreed (e.g., see the last point of sect. 3).
All raw data (and associated metadata) collected in experiments approved by the LNS (excluding commercial use of LNS research infrastructures or specific cases discussed above) will be open access after an initial embargo period (see Sect. 7) during which access is restricted to the experimental team, represented by the spokesperson.

All raw data (and associated metadata) obtained as a result of proprietary research will be owned exclusively by the client who purchased the beamtime and is not covered by this DMP. Commercial users must agree with the LNS director on how they wish their raw data and metadata to be managed before the start of any experiment.

6 CURATION OF RAW DATA AND ASSOCIATED METADATA

Raw data and metadata will have read-only access for the duration of their lifetime.
Raw data formats must be well documented in the metadata.
All raw data and metadata will be organized in a well-defined structure which will be made available by LNS. The list of essential metadata is given at the end of this document.
Only data with associated metadata will be archived. The spokesperson has the responsibility to provide the metadata (in electronic or pdf format) to the data steward, complying with the FAIR policy.
The spokesperson has to inform the data steward about the requirements in terms. e.g., of disk space upon the scheduling of the experiment.
Each experiment and data set will have a unique permanent DOI. Anybody publishing results based on open access data must quote the same identifier. This DOI will be assigned by the Open Science Working Group as routinely done for the sharing of products on the OAR. A DOI will be assigned also to each derived product.
Data and metadata will be stored on a short-term basis (typically for the duration of the data taking) in dedicated LNS servers. After that, data and metadata will be transferred and stored to INFN-CNAF and preserved for at least fifteen years, unless differently agreed (for instance, in the case of existing DMP or computing models). In the latter case, it is responsibility of the spokesperson to ensure to comply with the FAIR principles. Data steward has to receive positive response from the spokesperson prior to data upload.
High level metadata such as Title, Authors, Abstract, will be made public as soon as possible using a dedicated webpage. A data steward will be in charge to take care of the curation of the data as specified in the present document.

7 ACCESS TO RAW DATA AND METADATA

Access to raw data and the associated metadata obtained from an experiment is restricted to the experimental team for an embargo period of at least five years after the end of the experiment. Any spokesperson that wishes to retain restricted access to data for a period longer than five years will have this possibility to renew this five years period by submitting a written request, specifying the reasons for the proposed prolongation, to the LNS director who will accept or reject the request. In exceptional circumstances, data can be made openly accessible before the end of the embargo period if the spokesperson informs the LNS director about this decision. Thereafter, the data will become accessible upon request to the LNS director. The data steward will contact the spokesperson to establish whether the request can be fulfilled and what data can be shared, complying with the FAIR principles as much as possible.
Data and associated metadata will be accessible through the persistent identifier.
LNS staff (e.g. scientists, computing group members and, in particular, the data steward), after director authorization, can have access to any curated data or metadata for facility related purposes (for instance, statistical purposes or machine performance evaluation). LNS will undertake that confidentiality of such data is preserved during the embargo period.
The spokesperson has the possibility to transfer parts or the totality of her/his rights during the embargo period to LNS or another registered person of the experimental team.
The spokesperson has the possibility to create and distribute copies of the raw data within the collaboration (without using LNS resources).

8 METADATA CAPTURE AND STORAGE

The spokesperson needs to ensure that experiments' metadata are complete, as this will enhance the possibilities for everybody to search for, retrieve and interpret the data in the long term.

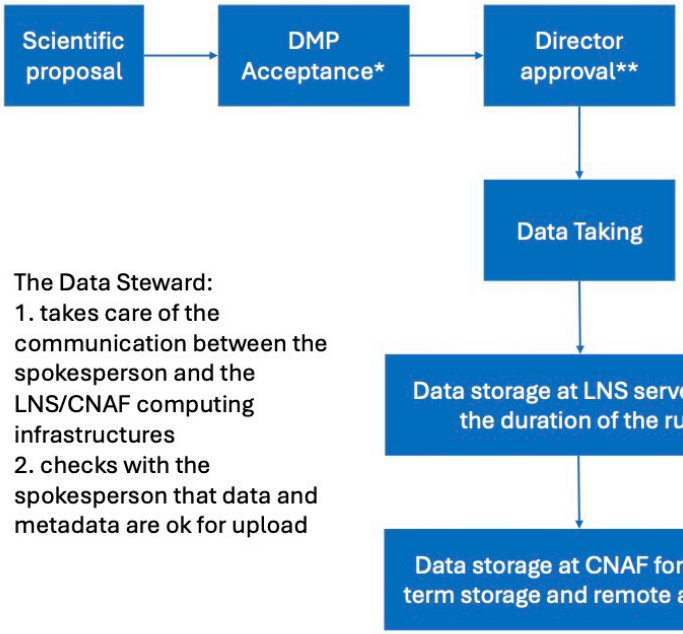
Researchers who aim to carry out analyses of raw data and metadata which are openly accessible should, where possible, contact the original spokesperson to inform her/him and suggest a collaboration if required. Researchers must acknowledge the source of the data and cite its unique persistent identifier as well as any publications linked to the same raw data.

9 PUBLICATION INFORMATION

Publications related to data collected in experiments performed at LNS must cite the persistent identifier of the data.
Once known, the DOI of the publication has to be sent to the data steward, who will inform the director and to the head of LNS research division

10 PERSONAL DATA

Personal information will be saved and archived by LNS in compliance with the principles set out in art. 6 of the EU Regulation 2016/679.



*With the following exceptions
1. Pre-existing DMP or MoU within the collaborarion or experiment
2. Commercial use of LNS infrastructure

**In the case of accelerator-based experiments, prior to PAC review

The Data Steward:
1. takes care of the communication between the spokesperson and the LNS/CNAF computing infrastructures
2. checks with the spokesperson that data and metadata are ok for upload

The Principal Investigator:
1. Gets a DOI from the Open Access Repository (OAR)
2. Makes data and metadata available to the data steward, in compliance of FAIR principle
3. Uses the unique identifier in forthcoming publications/research products

List of essential metadata

DOI
Spokesperson (Orcid)
Affiliation
Title of the experiment
Experiment approval number (or acronym, if not applicable)
Laboratory
Date of beginning of data taking
Date of end of data taking
Local contact person
Type of data (experimental data, simulation...) Type of experiment (nuclear physics, astroparticle...)
Type of beam (stable, in flight...)
Isotope
Chemical form
Beam energy
Detection setup
Type of detectors
Target
Experiment logbook
Accelerators/machines's logbook

After the embargo period, data will be openly accessible upon request if no extension of the embargo is requested/granted