LNL
Data Management Plan

E. Fioretto

DMP FOR INFN LABORATORIES

NuPECC Long Range Plan 2024

Thematic Working Group (TWG) on Open Data and Science Communication

INFN representatives
S. Bianco, E. Vigezzi

CSN3
28.03.23 – Meeting con R. Nania, S. Bianco, E. Fioretto, M. La Cognata, S. Pisano, F. Ferraro

GdL CSN3
E. Fioretto, M. La Cognata, S. Pisano, F. Ferraro

GdL Open Science
S. Bianco, M. Maggi, et al.

Informare le altre Commissioni Scientifiche Nazionali

LNL Data Management Plan

Data 20/06/2024

LNL Data Management Plan

Written by	Checked by	Approved by
Enrico Fioretto	Enrico Fioretto Tommaso Marchi (EURO-LABS TNA coordinator) Javier Valiente-Dobon (Gr3 Coordinator) Valeria Conte (Gr5 Coordinator) Antonello Ortolan (Gr2 Coordinator) Michele Gulmini (Responsible for the Information Technology Service)	

Subject

Data Management Plan for research activities making use of LNL ion beams.

Istituto Nazionale di Fisica Nucleare
DocID LNL-DIVR-M-00011 | Rev. 1.10 | Validità Bozza

TABLE OF CONTENTS

1 INTRODUCTION

The LNL Data Management Plan (DMP) concerns all research data collected in experiments performed by using ion beams delivered by the LNL accelerators.

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.

Some definitions of terms are listed in the second section. The various scenarios for the use of the beams delivered by the LNL accelerators are defined in the third section. Methods for storing and ensuring the long term durability of data, and rules to access research data and metadata are described in later sections.

2 DEFINITION OF TERMS

- SPOKESPERSON: Person responsible for the experiment, identified on the scientific proposal submitted to the LNL Program Advisory Committee, and for the data management collected during the experiment.
- DATA STEWARD: Person responsible for the management of research data throughout their life cycle, from the collection phase to the storage and sharing ones.
- EXPERIMENTAL ACCOUNT: Directory created by the Information Technology Service containing the data record for an experiment and therefore all the data linked to 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.
- EXPERIMENT CONTACT PERSON: LNL staff member who facilitates the running of the experiment.

Istituto Nazionale di Fisica Nucleare
DocID LNL-DIVR-M-00011 | Rev. 1.10 | Validità Bozza

- RAW DATA: All kinds of data collected by experiments carried out by using ion beams delivered by LNL accelerators.
- 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.
- 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.
- BEAMTIME COORDINATOR: Person in charge of the coordination of the accepted experiments and the preparation of the beamtime schedule maximizing the number of experiments to be performed in the available accelerator beamtime.
- OPEN ACCESS: Data accessible, upon request, by the community at large, and protected by open license.

3 DATA COLLECTION AND STORAGE

Legnaro National Laboratories (LNL) are engaged in research activites making use of ion beams and mainly addressed to nuclear physics, nuclear astrophysics, development of innovative detectors, material science and applications of nuclear techniques. Hence various uses of the beamtime are possible and thus require different data management scenarios described below and shown in Fig. 1:

- Data achieved with LNL instrumentation and stored using local resources (Data acquisition system and servers for storage).
- Data achieved with LNL instrumentation which use a data acquisition and storage system that does not belong to LNL. However, data sharing and durability will be guaranteed by LNL 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 LNL 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 LNL. The initial data storage does not use resources belonging to LNL. In this case as well, LNL will keep locally a copy of the data (provided by the spokesperson) as a guarantor of the data produced using LNL research infrastructures.
- Data generated during commercial use of LNL research infrastructures belong to the client and will not be stored on local servers unless agreed otherwise.

Moreover, metadata from these experiments will probably be used to categorize experiments and data records accurately.

4 OWNERSHIP OF DATA

INFN is the owner and the custodian of the raw data (and associated metadata) produced by using research infrastructures installed in its National Laboratories. Often, large collaborations have already a DMP or are ruled by international agreements such as MoUs. In such cases specific agreeements between LNL and the Management Boards of the collaborations have to be established.

The LNL data management policy is shown in Fig. 1 and 2.

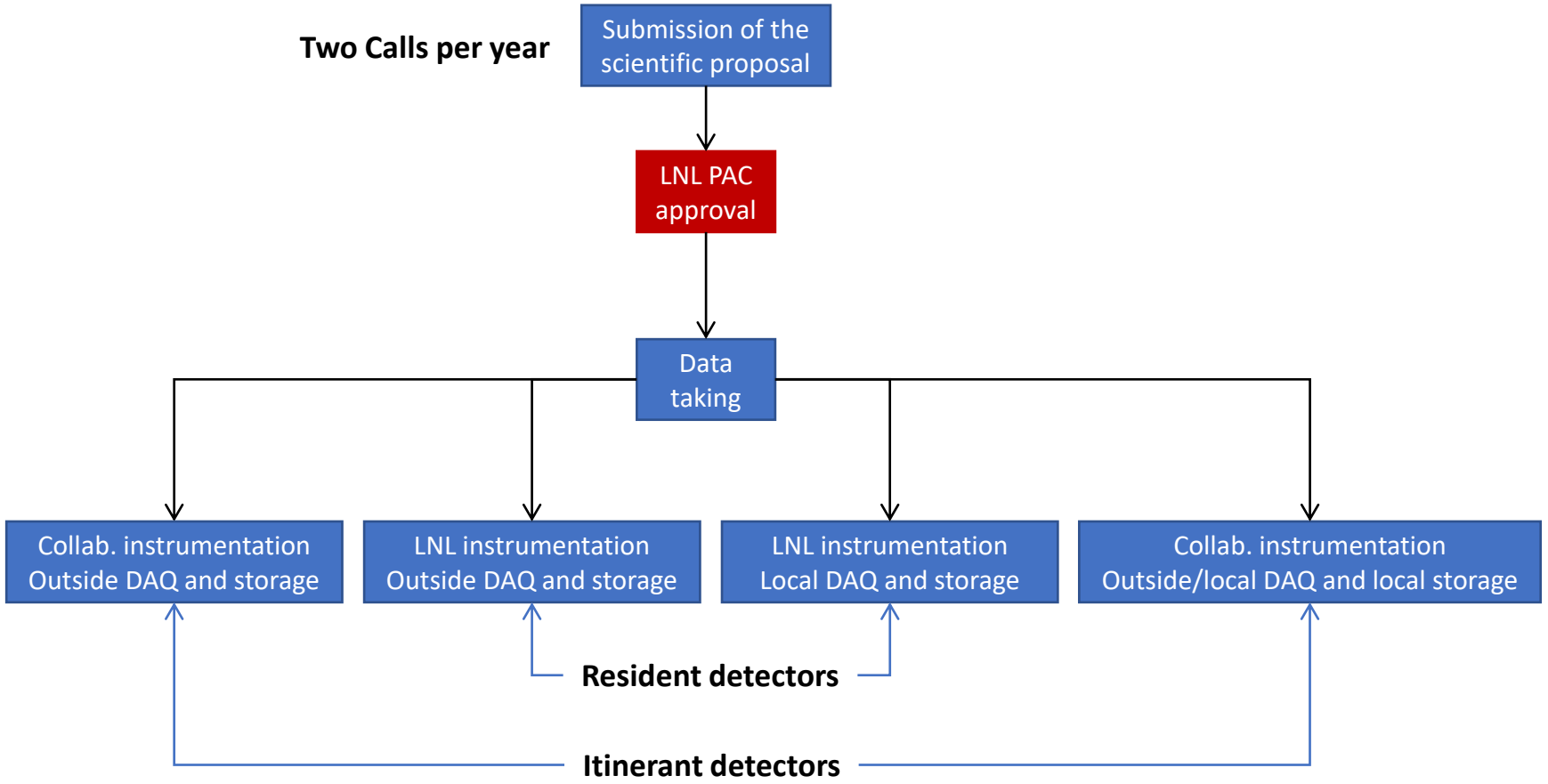
All raw data (and associated metadata) collected in experiments approved by the LNL Program Advisory Committee (excluding commercial use of LNL research infrastructures) will be open access

2. DEFINITION OF TERMS

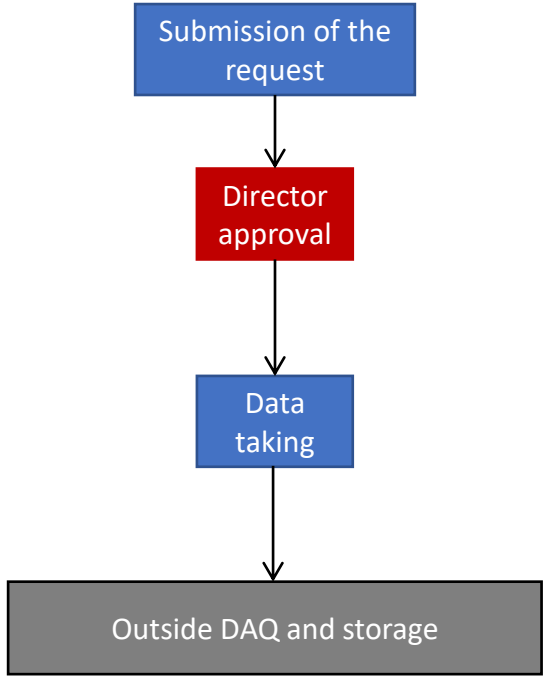
- SPOKESPERSON: Person responsible for the experiment, identified on the scientific proposal submitted to the LNL Program Advisory Committee, and for the data management collected during the experiment.
- **DATA STEWARD: Person responsible for the management of research data throughout their life cycle, from the collection phase to the storage and sharing ones**.
- EXPERIMENTAL ACCOUNT: Directory created by the Information Technology Service containing the data record for an experiment and therefore all the data linked to 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**.
- EXPERIMENT CONTACT PERSON: LNL staff member who facilitates the running of the experiment.
- RAW DATA: All kinds of data collected by experiments carried out by using ion beams delivered by LNL accelerators.
- 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.
- BEAMTIME COORDINATOR: Person in charge of the coordination of the accepted experiments and the preparation of the beamtime schedule maximizing the number of experiments to be performed in the available accelerator beamtime.

3. DATA COLLECTION AND STORAGE

Scientific use of LNL research infrastructures



Commercial use of LNL research infrastructures



3. DATA COLLECTION AND STORAGE

Instrumentation @ Tandem-ALPI-PIAVE accelerator complex

Experimental set-up	Beamline	DAQ	Data storage
SIRAD	(+70°)	Outside	Outside
Radiobiology	(+60°)	Local	Local
STARTRACK	(+50° B)	Local	Local
Sliding seal scattering chamber	(+50° A)	Outside/Local	Outside/Local
GAMIPE	(+40°)	Local	Local
MESH	(+30°)	Outside	Outside
PRISMA	(+20°)	Local	Local
AGATA + complementary detectors	(+20°)	Outside	Local
EXOTIC	(+10°)	Local	Outside
PISOLO	(-20°)	Local	Local
GALILEO	(-30°)	Local	Local
ex 8PLP	(-40° B)	Outside	Outside
LIRAS	(-40° A)	Outside	Outside
GARFIELD		Local	Local
ATS		NA	NA

3. DATA COLLECTION AND STORAGE

Instrumentation @ Tandem-ALPI-PIAVE accelerator complex

Experimental set-up	Beamline	DAQ	Data storage	Data size/exp
SIRAD	(+70°)	Irradiation		
Radiobiology	(+60°)	Irradiation	Processing in lab	
STARTRACK	(+50° B)	Local	Local	30-50 GB
Sliding seal scattering chamber	(+50° A)	Outside/Local	Outside/Local	
GAMIPE / ELCOM	(+40°)	Outside	Outside	200 GB / 50 GB
MESH	(+30°)	Irradiation		
PRISMA	(+20°)	Local	Local	500 GB
AGATA + complementary detectors	(+20°)	Outside (Agata) Local (ancillary)	Local (data taking) GRID long term	9 TB
EXOTIC	(+10°)	Local	Local	150 GB
PISOLO	(-20°)	Local	Local	100 GB
GALILEO	(-30°)	Local	Local	2 TB
ex 8PLP	(-40° B)	Outside	Outside	
LIRAS	(-40° A)	Outside	Outside	
GARFIELD		Local	Local	500 GB
ATS		NA	NA	5 TB

4. OWNERSHIP OF DATA

INFN is the owner and the custodian of the raw data (and associated metadata) produced by using research infrastructures installed in its National Laboratories.

Often, large collaborations have already a DMP or are ruled by international agreements such as MoUs. In such cases specific agreeements between LNL and the Management Boards of the collaborations have to be established.

All raw data (and associated metadata) collected in experiments approved by the LNL Program Advisory Committee (excluding commercial use of LNL research infrastructures) **will be open access after an initial embargo period 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 facility management on how they wish their raw data and metadata to be managed before the start of any experiment.

5. 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 LNL.
Only data with associated metadata will be archived.

The spokesperson has to inform data steward about the requirements in terms. e.g., of disk space upon the scheduling of the experiment. **The spokesperson has the responsibility to provide the data and the metadata (in electronic or pdf format) to the data steward, in compliance with the FAIR principles**. The spokesperson has 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.

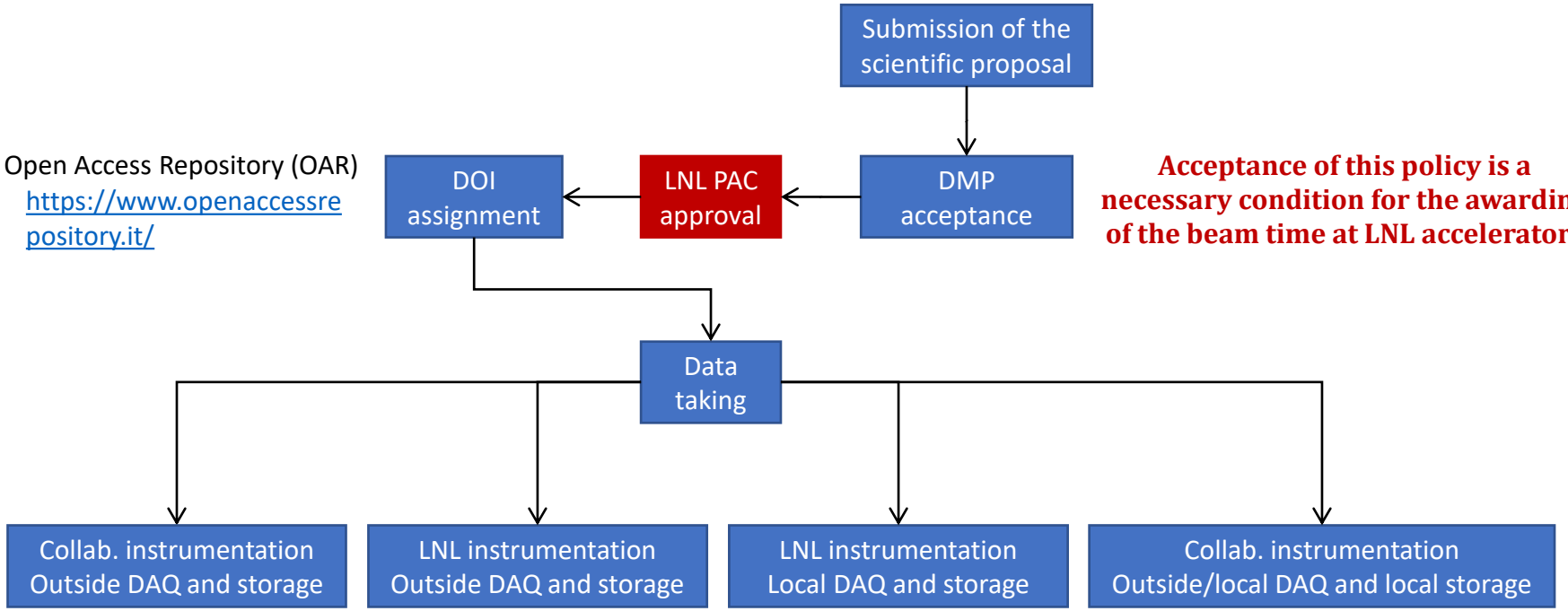
Each experiment and data set will have a unique permanent Digital Object Identifier (DOI). Anybody publishing results based on open access data must quote the same identifier. This DOI will be assigned by the Open Science Working Group as routinely done for the sharing of INFN products on the Open Access Repository (OAR).

High level metadata such as Title, Authors, Abstract, will be made public as soon as possible using the dedicated webpage (https://opendata.lnl.infn/). This information will be available through the persistent identifier. A data steward will be in charge to take care of the curation of the data as specified in the present document.

5. Curation of Raw Data and associated Metadata

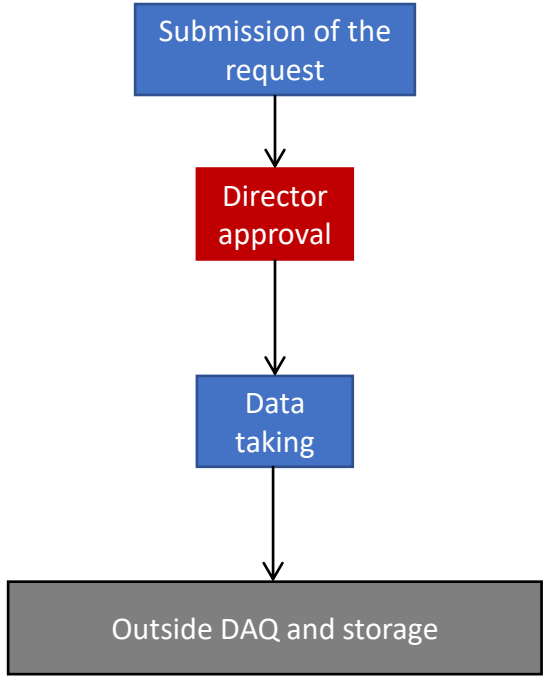
Scientific use of LNL research infrastructures

Open Access Repository (OAR) https://www.openaccessrepository.it/



Acceptance of this policy is a necessary condition for the awarding of the beam time at LNL accelerators

Commercial use of LNL research infrastructures



5. EXAMPLE OF METADATA

Items	Description	Field Mandatory	Comment	Format	Example
Identifier	Digital Object Identifier of the experiment data taking	Y	Assigned by the INFN Open Science WG	DOI format	
Alternate identifier	Local experiment number	Y	Assigned by the beam coordinators		LNL 23.034
Spokesperson	Full name	Y		First name and family name	
	ORCID	N	It allows cross referencing (data, publications, etc.)	ORCID Format	
	Affiliation	Y		Free text	
	E-mail address	Y		E-mail format	
Experiment title	Experiment title (as in the scientific proposal submitted to the PAC)			Free text	
Publisher	Laboratory where the experiment took data	Y		Free text	INFN – Laboratori Nazionali di Legnaro INFN – Laboratori Nazionali del Sud INFN – Laboratori Nazionali di Frascati INFN – Laboratori Nazionali del Gran Sasso
Publication Year	Year when the data will be made open upon request to the researcher community	Y	Included the embargo period	YYYY	2023
Contributors	Scientist or Staff involved in the experiment	N		Free text	
Experiment contact person	Local contact person	Y if the contact person exists			
	Full name	Y		First name and family name	
	ORCID	N		ORCID Format	
Date	Experiment start date	Y		Day Month Year	21 March 2022

+ Data format

Resource type	Resource type	Y			Experimental dataset
Logbook	Pdf file or electronic logbook	Y			
Topic	Research domain	Y		Free text	Nuclear Physics – Ion-Matter Interaction – Industrial Applications – Interdisciplinary activities – etc.
Beam	Isotopes	Y		Mass numbers followed by the chemical symbols	68Ni
	Production mode	Y			ISOL – In-Flight – Source
	Energy in MeV	Y		Beam energy followed by the symbol "MeV"	6.72 MeV
Detectors	Detector type	N			Gamma; Neutron; Charged Particles; Light; None
	Detector(s) used	Y			AGATA – PRISMA – GARFIELD – ATS – GALILEO – SIRAD irradiation facility – etc.
Target	Target nature	Y			Isotope, Organic Compound, Inorganic Compound, None
	Target description	N		Free text	
Description	All additional information that does not fit in any of the other categories	N		Free text	

6. 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 LNL 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 LNL director about this decision. **Thereafter, the data will become accessible upon request to the LNL director**.

Data and associated metadata will be accessible through the persistent identifier.

Authorized LNL staff (e.g. scientists, computing group members and, in particular, data stewards) have access to any curated data or metadata for facility-related purposes. LNL 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 LNL 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 LNL resources).

Researchers who aim to carry out analyses of raw data and metadata which are openly accessible should, where possible, should also 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.

7. METADATA CAPTURE AND STORAGE

Data and metadata will be stored on a short-term basis (one year maximum) in dedicated LNL servers and preserved for immediate access and data analysis.

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.

High level metadata such as Title, Authors, Abstract, will be made public as soon as possible using the dedicated webpage (https://opendata.lnl.infn/).
This information will be available through the persistent identifier.
A data steward will be in charge to take care of the curation of the data as specified in the present document.

8. PUBLICATION INFORMATION

Publications related to data collected in experiments performed at LNL must cite the persistent identifier of the data.
Once known, the DOI of the publication has to be sent to the Information Technology Service (calcolo@lnl.infn.it), in the person of the data steward, and to the Research Division (sdr@lnl.infn.it).

LNL Data Management Plan