HORIZON 2020
Research Infrastructures

H2020-INFRADEV-2016-1
INDIVIDUAL SUPPORT TO ESFRI AND OTHER WORLD-CLASS
RESEARCH INFRASTRUCTURES

IDEAAL
International Development of gAnil-spirAL2
Grant Agreement Number: 730989

D3.4 – Data Management Plan

Version: V2
Author: Nicolas. Ménard
Date: 30/09/2020
PROJECT AND DELIVERABLE INFORMATION SHEET

IDEAAL Project Ref. №	730989
Project Title	International DEvelopment of gAnil-spirAL2
Project Web Site	https://ideaal.ganil-spiral2.eu/
Deliverable ID	D3.4
Deliverable Nature	Data Management Plan
Deliverable Level*	PU
Contractual Date of Delivery	30/09/2020
Actual Date of Delivery	30/09/2020
EC Project Officer	Blagovesta CHOLOVA

* The dissemination level are indicated as follows: PU – Public, PP – Restricted to other participants (including the Commission Services), RE – Restricted to a group specified by the consortium (including the Commission Services). CO – Confidential, only for members of the consortium (including the Commission Services).

DOCUMENT CONTROL SHEET

Document	Title: Data Management Plan	
	ID: D3.4	
	Version V1	
	Available at: https://ideaal.ganil-spiral2.eu/	
	Software Tool: Microsoft Office Word 2007	
	File: D3.4_Data Management Plan_IDEAAL.pdf	
Authorship	Written by:	Nicolas MENARD (GANIL)
	Reviewed by:	Sabrina LECERF (GANIL)
	Approved by:	Marek LEWITOWICZ (GANIL)

DOCUMENT STATUS SHEET

Version	Date	Status	Comments
V0	23/09/2020	For internal review	
V1	29/09/2020	For internal review	
V2	30/09/2020	Submitted on EC Participant Portal	
		Final version	

DOCUMENT KEYWORDS

Keywords	DMP, Policy, DATA, Management, Plan

Disclaimer

This deliverable has been prepared by Work Package 3 (Excellence of Access to Infrastructure) of the IDEAAL Project in accordance with the Consortium Agreement and the Grant Agreement n°730989. It solely reflects the opinion of the parties to such agreements on a collective basis in the context of the Project and to the extent foreseen in such agreements.

Copyright notices

Deliverable D3.4 Data Management Plan

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE 1 : DATA MANAGEMENT STRATEGY *page 6*

REFERENCES AND APPLICABLE DOCUMENTS
[1]

LIST OF ACRONYMS AND ABBREVIATIONS

DMP	Data Management Plan
GANIL	Grand Accélérateur National d'Ions Lourds
DOI	Data Object Identifier
FAIR	Findable, Accessible, Interoperable, and Reusable

EXECUTIVE SUMMARY

The purpose of the task 3 of the work package 3 of the IDEAAL project is to provide a data management strategy for the data produce by the GANIL-SPIRAL2 infrastructure.

Study, discussions, thinking and work on this objective lead to produce a Data Manage Plan based on two documents :
- The first one, namely the Data Management Plan describes the whole process and responsibilities of the parties involved
- The second one, namely the Data Policy focuses on the role and responsibility of the spokesperson of a GANIL experiment; it has to be accepted by the spokesperson before the proposal foe experiment is submitted.

INTRODUCTION

The experiments, which are taking place in GANIL, produce data; those data, called raw data in this document, are produced, explored and analysed by physicists and collaboration to produce scientific results. The Data Management Plan of GANIL concerns all scientific data produced in experiments using GANIL-SPIRAL2 facilities.

The Data Management Plan at GANIL is governed by two documents: the "Data Policy" and the "GANIL Data Management Plan". The Data Policy briefly describes how data are managed and how responsibilities are distributed. The GANIL Data Management Plan describes the whole lifecycle of data.

SECTION 1 – THE DATA POLICY

The data management policy defines the ownership of, the curation of and access to experimental data and metadata collected and/or stored at GANIL. Approval of this policy is a condition for awarding a beamtime.

This two page document is addressed to the spokesperson of experiment and it is organized in six chapters: general principles, definitions, raw data and associated metadata, good practice for metadata capture and storage, publication information and personal data.

SECTION 2 – DATA MANAGEMENT PLAN

The document "Data Management Plan" describes the whole process and the lifecycle of data generated by the infrastructure. It defines the roles and responsibilities of each parties involved: GANIL, the spokesperson and the experiment contact person.

Additionally, the document defines raw data, data record, data set and metadata and describes the general workflow. Some of intangible rules are also set down:

- All data produced during GANIL experiments are covered by the DMP
- All data covered by the GANIL DMP will be identified by a specific and unique DOI
- The spokesperson has to gather all the metadata required in the DMP and she/he is encouraged to add every document and information that could help to reuse data in the future
- Each data record will be organized in a specific way described in the document; this will lead to have a homogenous organization of the data record
- An embargo period of three years is defined; the embargo period can be extended once for maximum three years
- GANIL is in charge of transferring the data record on the long-term storage infrastructure
- GANIL is in charge of registering the metadata, the dataset on the site stipulated in the document; it is also in charge of attributing a DOI
- The FAIR principles should be followed as far as possible.

The figure below shows the Data Management Strategy adopted by GANIL-SPIRAL2. The full text of the Data Policy and of DMP can be found in Annex 1 and 2 respectively.

GANIL DATA MANAGEMENT STRATEGY

ENSURE DATA PRESERVATION | OFFER OPEN ACCESS | INCREASE GANIL AND RESEARCHERS VISIBILITY
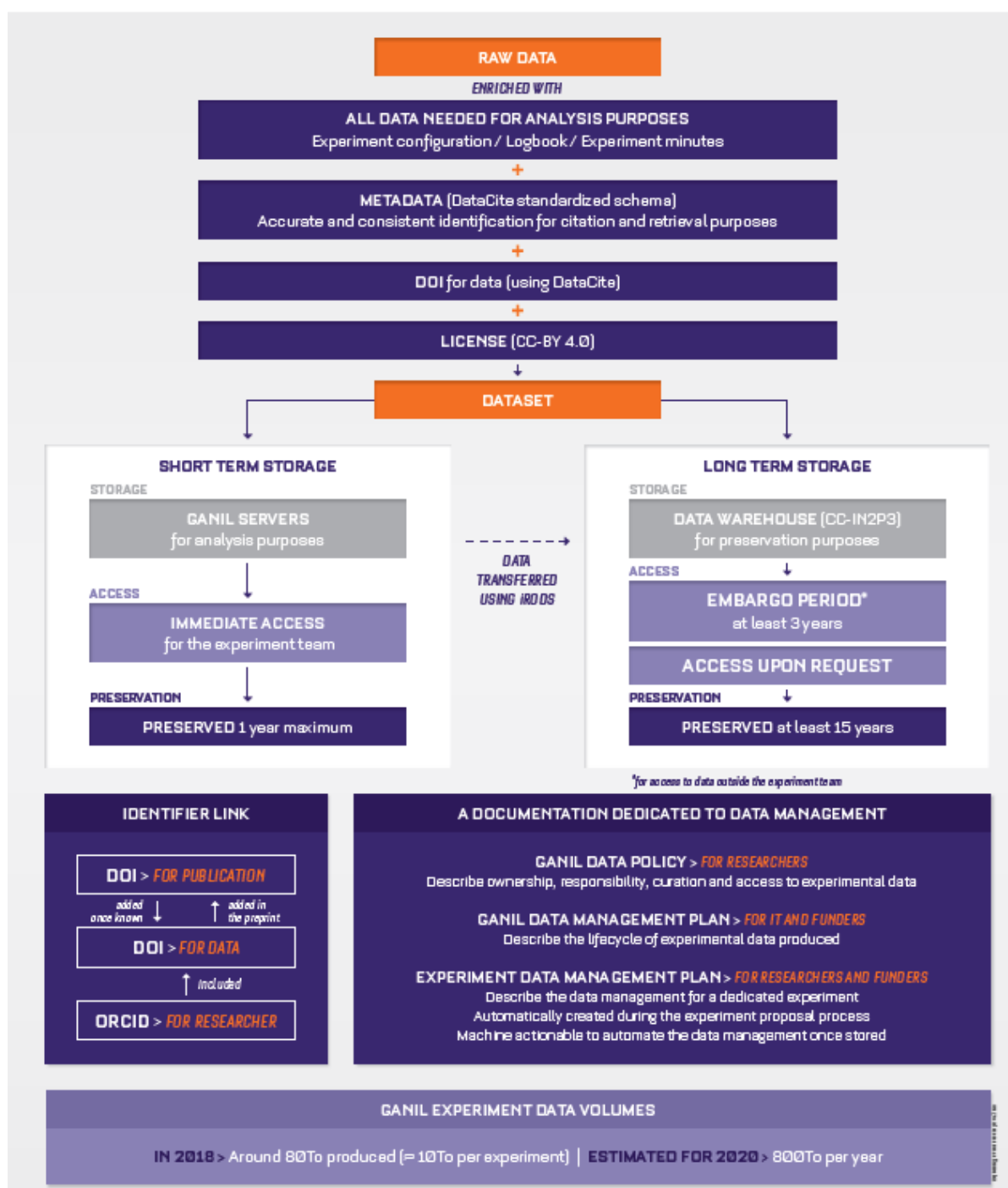
RAW DATA

ENRICHED WITH

ALL DATA NEEDED FOR ANALYSIS PURPOSES
Experiment configuration / Logbook / Experiment minutes

+

METADATA (DataCite standardized schema)
Accurate and consistent identification for citation and retrieval purposes

+

DOI for data (using DataCite)

+

LICENSE (CC-BY 4.0)

DATASET

SHORT TERM STORAGE

STORAGE
GANIL SERVERS
for analysis purposes

ACCESS
IMMEDIATE ACCESS
for the experiment team

PRESERVATION
PRESERVED 1 year maximum

DATA TRANSFERRED USING iRODS

LONG TERM STORAGE

STORAGE
DATA WAREHOUSE (CC-IN2P3)
for preservation purposes

ACCESS
EMBARGO PERIOD*
at least 3 years

ACCESS UPON REQUEST

PRESERVATION
PRESERVED at least 15 years

*for access to data outside the experiment team

IDENTIFIER LINK

DOI > FOR PUBLICATION

added once known ↓   ↑ added in the preprint

DOI > FOR DATA

↑ included

ORCID > FOR RESEARCHER

A DOCUMENTATION DEDICATED TO DATA MANAGEMENT

GANIL DATA POLICY > FOR RESEARCHERS
Describe ownership, responsibility, curation and access to experimental data

GANIL DATA MANAGEMENT PLAN > FOR IT AND FUNDERS
Describe the lifecycle of experimental data produced

EXPERIMENT DATA MANAGEMENT PLAN > FOR RESEARCHERS AND FUNDERS
Describe the data management for a dedicated experiment
Automatically created during the experiment proposal process
Machine actionable to automate the data management once stored

GANIL EXPERIMENT DATA VOLUMES

IN 2018 > Around 80To produced (= 10To per experiment) | **ESTIMATED FOR 2020** > 800To per year

Figure 1 : Data Management Strategy at GANIL-SPIRAL2

ANNEX 1 DATA POLICY

DATA POLICY

English version

Ind.	Date	Historique des modifications
A	06/01/2020	Initial version
B	29/09/2020	Revised version

Rédacteur	Vérificateur	Vérificateur	Approbateur
DSTA N Ménard	PHYSIQUE M. Lewitowicz	PHYSIQUE G. Defrance	DIRECTION N. Alahari

1 - GENERAL PRINCIPLES

a. The present data management policy pertains to the ownership of, the curation of and access to experimental data and metadata collected and/or stored at GANIL.

b. Acceptance of this policy is a condition for the award of beamtime.

c. Users must not attempt to access, exploit or distribute raw data or metadata unless they are entitled to do so under the terms of this policy.

d. Deliberate infringements of the policy may lead to denial of access to raw data or metadata and/or denial of future beam time requests at GANIL.

e. All data and metadata will be subject to the data protection legislation of France.

2 - DEFINITIONS

1. FOR THE PURPOSES OF THIS POLICY:

a. The term Raw Data pertains to data collected from experiments performed on GANIL instruments and includes data collected from experiments performed on CRG (Collaboration Research Group) instruments. This definition includes data that are created automatically or manually by facility specific software and/or facility staff expertise in order to facilitate subsequent analysis of the experimental data.

b. The term Metadata describes information pertaining to data collected from experiments instruments, including (but not limited to) the context of the experiment, the experimental team, experimental conditions and other logistical information.

c. The term Principle Investigator (PI) pertains to the PI identified on the experiment proposal. For experiments outside of the facility proposal system, the PI is the person initiating or performing the experiment. In the terminology used at GANIL the PI is the spokesperson.

d. The term Experimental Team includes the PI and any other person to whom the PI designates the right to access resultant raw data and associated metadata.

e. The term Open Access means accessible, upon request, by the community at large, and protected by open license.

f. Data enjoying open access are called Open Data, which for GANIL refers to data that have been released under the terms of the Creative Commons license CC-BY-NC 4.0, that obliges the users of such data to cite GANIL, the corresponding experiments and scientific teams as the source of the data. This license prevents use of data for commercial purposes.

3 - RAW DATA AND ASSOCIATED METADATA

1. OWNERSHIP OF DATA

a. GANIL is the owner and the custodian of the raw data and associated metadata produced by the facility.

b. All raw data and the associated metadata obtained as a result of peer reviewed access to GANIL and in-house research (excluding commercial use of GANIL) will be open access after an initial embargo period during which access is restricted to the experimental team, represented by the PI.

c. All raw data and the associated metadata obtained as a result of proprietary research will be owned exclusively by the client who purchased the access and is not covered by the GANIL data policy. Commercial users must agree with the facility management how they wish their raw data and metadata to be managed before the start of any experiment.

2. CURATION OF RAW DATA AND ASSOCIATED METADATA

a. All raw data and metadata will be curated in well-defined formats, for which the means of reading the data will be made available by GANIL.

b. Only data with metadata will be archived. The PI has the responsibility to communicate the metadata to the G2i (Groupe Infrastructure Informatique).

c. Raw data and metadata will be read-only for the duration of their life time.

d. Raw data and metadata will be migrated or copied to archival facilities for long-term curation.

e. Each experiment and data set will have a unique persistent identifier. Anybody publishing results based on open access data must quote the same identifier. This Digital Object Identifier (DOI) will be created by GANIL, and registered with the DataCite organization (https://www.datacite.org/).

f. High level metadata such as Title, Authors, Abstract, will be made public as soon as possible using DataCite. This information will be available via the persistent identifier landing page on the web. Experiments details will be retrievable using https://search.datacite.org/.

3. ACCESS TO RAW DATA AND METADATA

   a. 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 three years after the end of the experiment. Thereafter, the data will become accessible upon request. Access details will be accessible on the persistent identifier landing page. Any PI that wishes data to retain restricted access for a period longer than three years will have this possibility to renew this three year period by submitting a written request, specifying the reasons for the proposed prolongation, to the facility management who decide on the request. In exceptional circumstances, data can be made openly accessible before the end of the embargo period if the PI informs G2i to do so.

   b. Authorized GANIL staff (e.g. instrument scientists, computing group members) have access to any curated data or metadata for facility related purposes. GANIL will undertake that confidentiality of such data is preserved during the embargo period.

   c. The PI has the possibility to transfer parts or the totality of her/his rights during the embargo period to GANIL or another registered person of the experimental team.

   d. The PI has the possibility to create and distribute copies of the raw data within the collaboration (without using GANIL resources).

4 - GOOD PRACTICE FOR METADATA CAPTURE AND STORAGE

   a. The experimental team 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.

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

5 - PUBLICATION INFORMATION

   a. Publications related to data from experiments carried out at GANIL must cite the persistent identifier of the data in their publication.

   b. Once known, the DOI of the publication has to be sent to the G2i (g2i-dmp@ganil.fr)

6 - PERSONAL DATA

   a. Personal information will be saved and archived by GANIL. The names, first names and affiliations of the participants in the experiments are part of the metadata.

ANNEX 2. DATA MANAGEMENT PLAN

<div style="background:#1a2a5a;color:#fff;text-align:center;">

GANIL DATA MANAGEMENT PLAN

</div>

English version

Ind.	Date	Historique des modifications
A	06/01/2020	Initial version
B	29/09/2020	Revised version

Rédacteur	Vérificateur	Vérificateur	Approbateur
DSTA N. Ménard	PHYSIQUE M. Lewitowicz	PHYSIQUE G. De France	DIRECTION N. Alahari

Deliverable D3.4 Data Management Plan

Table of Contents

1 - DATA COLLECTION

The Data Management Plan of GANIL concerns all scientific data produced in experiments using GANIL beams.

The Data Management Plan at GANIL is governed by two documents: the "Data Policy" (ref. GANIL-0560) and the "GANIL Data Management Plan", the present document.

The present document, aims to describe the process, management workflow, roles and responsibilities of involved parties.

The various scenarios for use of the beams from the GANIL accelerators are defined in this first section. Methods for storing and ensuring the long term durability of data are described in later sections.

The present document contains also a description of the data generated by the facility and its organization.

1.1 - OBJECTIVE OF DATA COLLECTION

GANIL is engaged in research with ion beams with the main focus of the laboratory being fundamental nuclear physics. This is supplemented by strong programs in accelerator based atomic physics, condensed matter, radiobiology and industrial applications. Hence a variety of uses of the beam are possible and thus require different data management scenarios described in this document.

1.2 - USES OF THE GANIL BEAMS.

Each of the items described below will be the subject of dedicated type of data management.

- Scenario 1: Acquisition and GANIL servers (scenario 1, section 5 -): Standard use of the facility, Data produced during the experiments are stored using resources provided at GANIL.
- Scenario 2: Acquisition and servers outside GANIL (scenario 2, section 5): Certain research teams may request, while submitting the experimental proposal, to use an acquisition and storage system that does not belong to GANIL. Data durability will, however, be guaranteed by GANIL by keeping a copy of the data generated.
- Scenario 3: Instruments belonging to a collaboration with acquisition and GANIL servers (scenario 3, section 5): Data generated from instruments belonging to a collaboration which has chosen to use GANIL resources for storage.
- Scenario 4: Instruments belonging to a collaboration with acquisition and servers outside GANIL (scenario 4, section 5): The initial data storage does not use resources belonging to GANIL. In this case as well, GANIL will keep locally a copy of the data as a guarantor of the data produced using GANIL facility.
- Scenario 5: Data generated during commercial use of GANIL resources belong to the client and will not be stored on GANIL servers unless agreed otherwise (scenario 5, section 5).
- Scenario 6: Interdisciplinary Research (scenario 6): Experiments conducted by CIRIL (Centre for Interdisciplinary Research with Heavy Ions) using the GANIL facility. GANIL is responsible for ensuring durable, open data from CIRIL experiments (scenario 6, section 5 -).

Furthermore, Metadata from these experiments will probably be used to categorize experiments and data records accurately.

1.3 - DESCRIPTION OF THE DATA

1.3.1 - DATA RECORD

A Data record (see definition page 23) of an experiment contains necessarily the Raw data (see definition page 23) generated during the experiment at GANIL.

- MultiFrame Metaformat (MFM)

The data acquisition system of GANIL enables data to be generated in the MFM format. This format represents around 95% of the raw data generated by the facility.

The specifications for this format can be consulted through the following link: [MultiFrame Specifications](#).
- Data format IN2P3

Certain experiments (from CIRIL in particular) generate data in an older format: data format IN2P3. A Data Record must contain at least raw data and metadata; any other data, information or files can be added to help future exploitation of data record.

1.3.2 - ACQUISITION PARAMETERS

The data acquisition parameters can be found in the "Ganacq Manip" file.
This file contains the entire parameter set for the acquisition system, the parameters for the electronic modules (GECO) and the different analysis codes to visualize the experiment in progress (GRU, VIGRU and specific codes used by the experiment).

1.4 - SYSTEM FOR ORGANIZING, NAMING AND MANAGING DIRECTORIES AND FILES.

1.4.1 - GENERAL

For each experiment conducted at GANIL, a dedicated directory for the experiment is created when the experiment is scheduled to run. This is the account attributed to the experiment.
This account is created by the IT Group (G2i) of GANIL.
In the specific case where an experiment is to be repeated, using an already existing experiment number, a new directory will be created (i.e. Experiment e700 gave rise to the creation of directory "e700" on the storage servers, when the experiment is repeated, a directory "e700v1" is created).

1.4.2 - DIRECTORIES AND FILES

When an experiment account has been created the following files will be created : /home/eNNN and /data/eNNNX where NNN is the number of the experiment.
All the files described below and therefore the experimental data are available in /data/eNNNX. The complete flow chart is shown in Annex 1. All data are read-write.
- Data

This file contains the raw data produced during an experiment.
The data are automatically stored in the above directory.
- Config

This file contains the electronic configuration used for the experiment and in particular the "Ganacq Manip" file.
- Logbook

The logbook (or manip book) will be saved in this file.
Paper logbooks must be scanned and saved under the supervision of the spokesperson for the experiment
The report on the experiment is also to be stored in the Logbook file.
When the Logbook has been validated it must be duplicated in this file by the spokesperson if there is one for the experiment in question.
- MISC

This file is intended to contain all the information that may be of use in the analysis of the data:
  - Preparatory documents;
  - Reduced data;
  - Software package for reading the data;
  - Any other document considered important.
- Metadata

This file contains the metadata for characterising the experiment accurately (see section 2 for more information).
Part of the metadata can be recovered automatically, the other part must be entered manually.

Eventually, it will be possible to recover all metadata automatically.

- Temp

This file is a backup space for the experimental team. Unlike the other files, this space will not be backed up on a long-term basis in the context of the procedure described in Section 3.

1.4.3 - NAMING RULE

The main objective of long-term storage, is data durability and reuse, so the system for naming files and saved documents must follow the standards used by the scientific community at GANIL.

1.4.4 - DATA FLOW AND RESPONSIBILITIES

The figure below shows the different files in the experiment account and the person responsible for execution of backups.

The spokesperson for an experiment will be responsible for creation and communication of the metadata characterizing the experiment and for adding the documents and files as described below. The spokesperson can be guided by the Experiment Contact Person (ECP) who will be a staff member of GANIL or CIRIL.

The experimental team is encouraged to store all documents, files, source codes and software to facilitate reuse of the data in the future.

At the end of the experiment, the spokesperson ensures data are stored in the dedicated directory on the GANIL infrastructure.
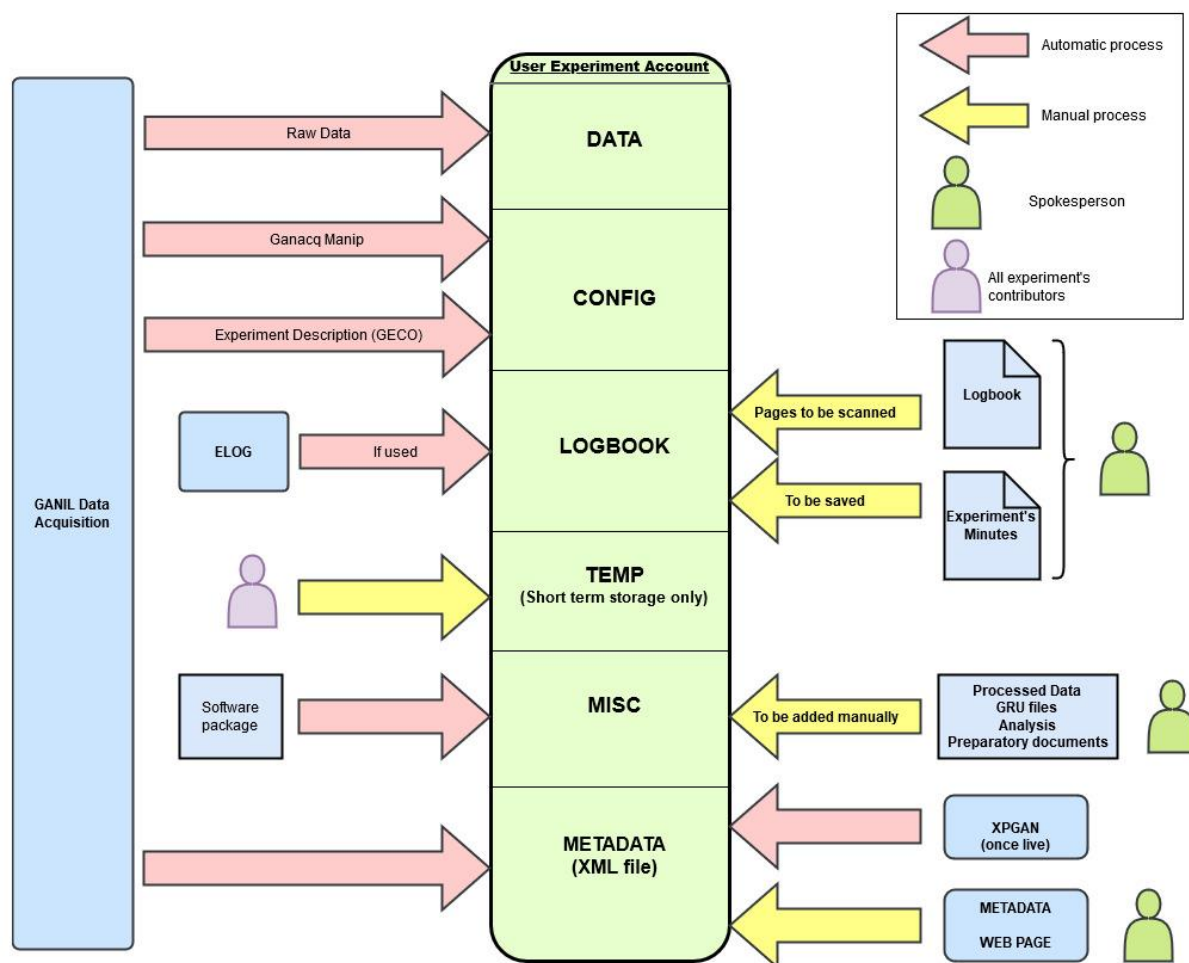

Figure 2: Feeding data into the experiment account (GANIL Data Acquisition)

As described in the laboratory data policy, GANIL is the owner of the data generated by the facility and guarantees their durability.

2 - CATEGORIZING DATA USING METADATA

A metadata file will be created for each data record. This file will include "administrative" information such as the spokesperson or the date of the experiment and more technical information characterizing the experiment.

2.1 - USE OF METADATA

Metadata are essential for:

- Accurate characterization of a data record and therefore an experiment;
- Cataloguing experiments and data;
- Data search and reuse;
- Understanding the data, by ensuring the link between the data producer and the persons reusing them.
- In the event of reuse of the data by a third party, metadata would facilitate citation of the data.

2.2 - STANDARD USED

The standard used at GANIL is the "DataCite Metadata Schema" (https://schema.datacite.org/). The scheme has been updated to include the specific features of experiments conducted at GANIL.

2.3 - DIGITAL OBJECT IDENTIFIER - DOI

To facilitate third party awareness of the data, GANIL has chosen to create a unique identifier for each data record. This identifier, known as the Digital Object Identifier (DOI) will be registered on the DataCite platform enabling third parties to find experiments conducted at GANIL (see details in chapter 6.6 -Implementation of DOI page 22)

2.4 - LIST OF METADATA

The spokesperson, with the support of G2i, ensures the metadata gathering.
The complete list of metadata can be found in the Annex 2.

2.5 - METADATA STORAGE

Metadata created for each data record will be stored on the GANIL servers.
In addition, a XML file containing all the metadata of the data record will be stored in the "Metadata" file for the experimental account in question.

2.6 - CREATION OF A DATA PORTAL

The objective is to create an interface for searching the Metadata database.
The interface will be accessible from the GANIL website and will provide the community an extensive view of the experiments that have taken place at GANIL and the conditions to access the experimental data generated by the facility.

3 - DATA STORAGE, BACKUP AND SECURITY

3.1 - STORAGE

3.1.1 - GANIL RESPONSIBILITY

As specified in the above sections, GANIL is the owner and guarantor of the data generated by the facility. GANIL must therefore ensure conservation and access while simultaneously guaranteeing data security.

3.1.2 - STORAGE AT GANIL

In 2020 the volume of scientific data generated by the facility and stored locally was around 190 TB. Taking into account the operation of SPIRAL2, it is estimated that a total data volume of 600TB to 800TB per year will be produced from 2021.
In 2020 local storage capacity at GANIL is more than 1 PB with about 900 TB available.

3.1.3 - DATA STORAGE

- Local storage (short term)

From the start of the experiment, data are stored in /Data directory to provide the experimental team with an immediate access to the data.
  - Storage site

Data are stored on Isilon servers in room 85, which is a controlled area in the main building of GANIL. Access to this room is only possible to authorized personal and the room is under constant video surveillance.
  - Duration of storage

To enable the experimental team to access the data with ease for a reasonable period, for data analysis, the data is locally available for a maximum period of one year.
To optimize storage space, G2i may, in agreement with the spokesperson and/or ECP, shorten this period if necessary.
  - Access to data

Data are accessible from dedicated work stations in the experimental areas.
Data can also be accessed from a GANIL work station.
There is an SSH gateway for accessing data from a remote terminal.
  - Management of access rights

Access to data is protected by a login and a password. This access is managed by the G2i team. Information for accessing the data will be communicated to the spokesperson of the experiment who will notify the experimental team.
  - Rights to data

Data in the experiment account are read-write
  - Cost of local storage

The servers installed in 2019 and 2020 at GANIL for a useful capacity of around 1 PB will be warranted for 5 years.
The cost per 1TB is therefore around 62€/year (excluding manpower costs).
Annual investment was 170k€ in 2019 and 100k€ in 2020.
  - Local storage manager

As noted earlier, G2i is responsible for operations linked to local storage of data.
- Off-site storage (long-term conservation)

To guarantee the durability of the scientific data generated by the facility, the data in the experiment account and therefore the entire data record, will be duplicated to an off-site storage facility.
  - Types

After the end of the experiment, all files in the experiment account (except for the TEMP directory described in section 1) will be duplicated to a dedicated space in a data warehouse.
The transfer will be carried out manually by G2i using the iRODS data management software.
As this transfer requires a complete Metadata file, the G2i team may contact the spokesperson for the experiment to obtain any missing metadata.
  - Remote storage site

The data will be stored at CC-IN2P3 in Lyon.
  - Duration of storage

By default, data will be stored for a minimum period of 15 years. However, when the proposal for the experiment is submitted, the spokesperson for the experiment may define a longer period for conserving the data.

o Access to data

Data will be accessible to the experimental team throughout the data life cycle.
At the request of the Spokesperson for the experiment, data can be made open to people outside the experimental team/collaboration before the end of the embargo period (as defined in section 5).

o Management of access to data

Access requests must be sent to G2i who will define the type of access to data with IN2P3 computing centre (CC-IN2P3).

o Access level to data off site

Data stored off site will be in read only mode.

o Cost of off-site storage

The cost of archiving of the data from GANIL is paid by IN2P3 through its annual budgetary allocation to CC-IN2P3, at no cost for GANIL.
In case the storage to CC-IN2P3 no longer possible, the cost of storage is estimated to be around 100€ to 150€ per TB/year.

o Off-site storage manager

The Group responsible for data storage at CC-IN2P3 storage group can be contacted for any questions regarding the off-site storage.

o Request for off-site storage

Each year, GANIL needs to request storage space according to the estimated data production for the following year. This request is made by the G2i team

3.2 - BACKUP

Daily backup of locally-stored data is organized and managed by G2i. Data are backed up to CC-IN2P3 every day.
If data are lost locally G2i will ensure that relevant recovery procedures are implemented.
Disclaimer: depending on the amount of data to save, the frequency of backup can be more than once a day.

3.3 - SECURITY

As specified in the previous sections, access to physical servers is severely restricted   and only possible for certain members of G2i staff.
Data from an experiment with a spokesperson from GANIL that is stored locally is accessed via a login and password provided by G2i.
Transfers (copy, backup) will be carried out using secure protocols

3.4 - RISK MANAGEMENT

3.4.1 - DESTRUCTION OF LOCAL SERVERS

If local storage servers are no longer usable, there are 3 possible procedures to ensure rapid data recovery:

- The ARP (Activity Recovery Plan), automatic daily backup;
- TIVOLI, automatic daily backup
- iRODS, manual backup after each experiment.

3.4.2 - DATA PROTECTION

Data transfers are encrypted and monitored by the G2i team.

3.4.3 - DATA INTEGRITY

Data integrity is guaranteed by the protocols used.

4 - CONSERVATION AND RESPONSIBILITIES

The type of conservation depends on the scenario for the experiment (the scenarios are described in section 1.2.)

As far as possible and irrespective of the scenario for use of the GANIL facility, both the raw data as well as all the information that facilitate subsequent reuse of the data must be stored.

The spokesperson must ensure that storage of data corresponds to the selected scenario for the experiment (Section 5).

5 - SCENARIOS FOR DATA COLLECTION, STORAGE AND ACCESS

[1] Raw data and all data permitting subsequent reuse of the data (see section 1, fig 1)

Scenario	STORAGE and IDENTIFICATION			ACTIONS and RESPONSIBILITIES	
	Short-term storage (local)	Long-term storage (off-site)	DOI creation	Actions	Party responsible for the action
1: Experiment with GANIL acquisition and GANIL storage	Data will be stored locally for a maximum period of one year	At the end of the experiment data is copied to CC-IN2P3	Yes	Ensure that data[1] is correctly stored in the experiment account and transferred to CC-IN2P3	Spokesperson or Data Manager
2: Experiment with non-GANIL acquisition and non-GANIL storage	Data must be transferred to GANIL servers at the end of the experiment[2]	Data transfer to CC-IN2P3 as soon as possible	Yes	Ensure transfer of data[1] to GANIL servers	
3: Instrument(s) belonging to a collaboration with GANIL acquisition and GANIL storage	Data will be stored locally for a maximum period of one year	At the end of the experiment data copied to CC-IN2P3	Yes	Ensure as soon as possible at the end of the experiment that all the data[1] are correctly backed up in the experiment account.	
4: Instrument(s) belonging to a collaboration with non-GANIL acquisition and non-GANIL storage	Data must be transferred to GANIL servers at the end of the experiment[2]	Data transfer to CC-IN2P3	Yes (to be discussed with the collaboration if created by the collaboration)	Ensure transfer of data[1] to GANIL servers	
5: Industrial applications	To be discussed with GANIL management before any experiment		No		
6: CIRIL	To be defined according to the type of experiment (action to be taken by CIRIL)	Data transfer to CC-IN2P3	Yes		Spokesperson or Data Manager

[2] If the volumes generated are too large a direct copy to CC-IN2P3 could be considered

[3] Experiment Contact Person: GANIL or CIRIL staff member in interface between the experiment team and GANIL

6 - DATA SHARING

One of the objectives of the French Ministry of Higher Education, Research and Innovation is to make open access to research data from programs with public funding mandatory (national plan for open science).

GANIL has therefore chosen to follow these recommendations and, thus, implement a Data Management Plan that complies with the FAIR principles described below.

6.1 - FAIR DATA PRINCIPLES

6.1.1 - FIND

Data must be easily findable
The use of metadata characterizing the experiments conducted at GANIL and the creation of the unique, Digital Object Identifier (DOI) will make it possible to find experimental data using search engines and in particular https://search.datacite.org/.
The DOI for GANIL experiments are registered on the DataCite platform https://www.datacite.org/ (see section 5.6).

6.1.2 - ACCESS

Data must be accessible and open as far as possible.
GANIL will offer access upon request.
The various types of access are specified on the DOI main page for the data in question.
With the spokesperson's and GANIL's agreement, a link with the connection identifiers will be provided to the applicant to access only the data in question. The applicant will be able to download the data.

6.1.3 - INTEROPERATE

Data must be interoperable to facilitate subsequent analysis
The data generated by GANIL will have standard formats and standards used by the nuclear and particle physics community.
Metadata use the DataCite standard (DataCite Metadata Schema 4.1), details are available at https://schema.datacite.org/.

6.1.4 - REUSE

The ultimate goal is to optimise the reuse of data.
A software package will be part of the data record to enable the raw data to be read and reused.
These open source software packages will also be available on request on http://wiki.ganil.fr/gap/wiki/GANIL-SPIRAL2-DAQ-Packages-Distribution.
In addition, GANIL may also provide elements (codes and related information) to facilitate the use of data.
Finally, a "Creative Common" license[1] will be used for all the data shared by GANIL.

6.2 - WHICH DATA ARE SHARED?

All data in long-term storage may be open to access. That is, all data in the data record for an experiment.

6.3 - PERSON RESPONSIBLE

The Data Manager is responsible for the data as defined when the experiment is submitted.
By default, the Data Manager is the Spokesperson for the experiment.

6.4 - EMBARGO PERIOD

The Embargo Period is a minimum period of 3 years, starting from the end of the experiment, during which the data will only be open to the experimental team.
The duration chosen when submitting the experiment can be longer and last up to 6 years (3 years renewable once).

---

[1] See https://creativecommons.org

6.5 - Type of open access

On request, the Data Manager may agree or not to data access after the embargo period.
Any refusal after the embargo period must be justified.
If requested, the Data Manager may share the data before the end of the embargo period with agreement from GANIL management.

6.6 - Implementation of DOI

A DOI attribution agreement has been signed with INIST under reference VDR-2019-000009.

6.6.1 - Creation of the long-term identifier

This identifier can be defined when the experiment is planned.
Its creation is linked to its registration on the DataCite platform (see below).

6.6.2 - Naming nomenclature for this identifier

When the agreement was concluded with INIST, a prefix was attributed to GANIL for naming its DOI.
The prefix is as follows: **10.26143**.
The suffix will be chosen by GANIL and will take the following form: "GANIL-YYYY-experiment number" where YYYY is the year of the experiment.
For example, for experiment E699, the DOI will take the following form: **10.26143/GANIL-2018-E699**.

6.6.3 - Creation of the metadata XML file

When the metadata linked to the experiment in question are known (upon G2i's request to the Spokesperson), G2i can create an XML file according to the DataCite scheme (see section 2).

6.6.4 - Creation of the Landing Page

A website displaying the metadata and data access conditions must be created and hosted by GANIL.
The URL for this site must be communicated to DataCite when registering the DOI.

6.6.5 - Registering the DOI on the DataCite platform.

G2i will be responsible for registering DOIs with the DataCite Metadata Store platform.
The login and password supplied by INIST can be used to connect to the platform and upload the following 3 elements:

- DOI, that is, the identifier itself (prefix and suffix);
- metadata XML file, in compliance with the DataCite schema;
- URL for the Landing Page for the experiment in question.

When the DOI has been registered, the experiment can be searched for on the web.

6.6.6 - Cost

An annual rate of **180 Euros excl. tax** will be paid by GANIL to INIST for the duration of the attribution of DOI names.

7 - DEFINITIONS

Experimental account: Directory created by the IT Infrastructure Group (G2i) 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 for the experiment concerned.

Digital Objet Identifier (DOI) : Unique, long-term identifier permitting identification of a data record. This identifier will be created by G2i and recorded on the DataCite platform.

Data record: Data contained in the experiment account. Raw data produced by the facility and all the data required for analysis of the raw data: configuration of the experiment, various parameters, logbook, report on the experiment, software package, etc.

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 (ECP) : GANIL or CIRIL staff member who facilitates the running of the experiment.

Landing Page: Web page visible when the DOI for a data record is consulted. This page can be used to display details of the experiment (metadata) and the types of access to the data. Contact person and embargo period in particular.

Metadata: Metadata are structured information used to describe data.
Data that can be used to accurately characterize an experiment and therefore a data record. GANIL metadata follow the schema proposed by DataCite.

Raw data: data produced by an experiment, written by acquisition systems and which are not filtered, computed or processed.

Spokesperson: This is the person responsible for the experiment, generally the person who proposed the experiment and is responsible for the data management of the experiment.

8 - ANNEXES :

8.1 - ANNEX 1: PROPOSED TREE DIAGRAM FROM THE EXPERIMENT ACCOUNT

```
📁 1 DATA
   📁 Acquisition
   📁 Data
   📁 Run
📁 2 CONFIG
   📁 Ganacq Manip
📁 3 LOGBOOK
   📁 ELOG
   📁 Exp Minutes
   📁 Logbook
📁 4 TEMP
📁 5 MISC
   📁 Analysis
   📁 Other
   📁 Processed Data
   📁 Software
📁 6 METADATA
```
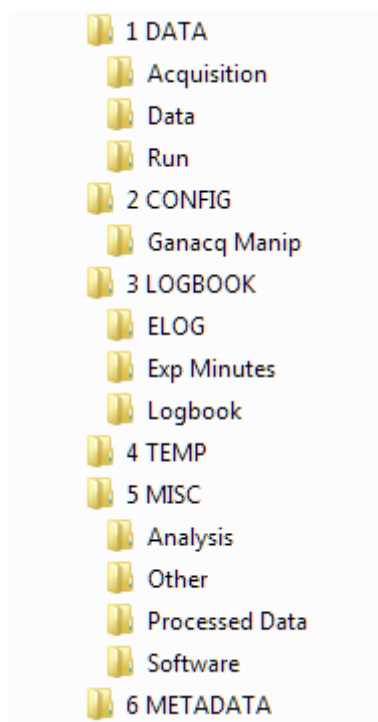
8.2 - ANNEXE 2: COMPLETE LIST OF METADATA

Attribute Name	ID	Description	nb	Datacite Mandatory	GANIL Mandatory	Comment	Format	Example
Identifier	1	DOI	1	Y	Y		DOI format	10.1234/ganil_e638
identifierType	1.1	Predefined value	1	Y	Y			DOI
Creator	2	Spokesperson's description	1-n	Y	Y			
creatorName	2.1	The full name of the Experiment's Spokesperson	1	N	Y		family name, first name	Durand, Jean-Pierre
nameIdentifier	2.4	ORCID	0-n	N	N	Allow cross referencing (data, publication, researcher)	ORCID Format	https://orcid.org/0000-0001-7069-0000
nameIdentifierScheme	2.4.1	Predefined value	1	Y if nameIdentifier exists	N			ORCID
affiliation	2.5	Institutional affiliation of the creator	0-n	N	N		Free text	GSI
creatorEmailAddress	2.6	Email address of the spokes person	1-n					
Title	3	Experiment title	1-n	Y	Y		Free text. No particular format or language	Study of nuclei at GANIL
Publisher	4	Laboratory where the experiment takes place	1	Y	Y		Legal name of the laboratory	GANIL
PublicationYear	5	If known, the year when the data will be made open upon request to the community of researchers	1	Y	Y	Included an embargo period	YYYY	2021
Subject	6	Will include all GANIL/other labs attributes (see line 32)	0-n	N	Y		Free text. No particular format or language	
Contributor	7	Scientist or Staff involved in the experiment	0-n	N	N			
contributorType	7.1	Type of contributor	1	Y if Contributor exists	Y if Contributor exists		Controlled List Values	Experiment ContactPerson; Researcher
contributorName	7.2	The full name of the contributor	1	Y if Contributor exists	Y if Contributor exists		family name, first name	Durand, Jean-Pierre
nameIdentifier	7.5	ORCID	0-n	N	N	Allow cross referencing (data,	ORCID Format	https://orcid.org/0000-0001-7069-0000

						publication, researcher)		
nameIdentifierScheme	7.5.1	Predefined value	1	Y if nameIdentifier exists	Y if nameIdentifier exists			ORCID
Date	8	Experiment start date	0-n	N	Y		Month Year	March 2018
dateType	8.1	Predefined value	1	Y if Date exists	Y		Controlled List Values	Created
ResourceType	10	Predefined value	1	Y	Y			Experimental data
resourceTypeGeneral	10.1	Predefined value	1	Y	Y			Dataset
AlternateIdentifier	11	GANIL Experience number	0-n	N	Y		e (lowercase) followed by 3 digits	e658
alternateIdentifierType	11.1	Predefined value	1	Y if AlternateIdentifier exists	Y			GANIL Experiment Number
RelatedIdentifier	12	Publication's DOI, ISBN	0-n	N	N	Allow cross referencing (data, publication, researcher)	Free text	10.1234/elsevier_154236
relatedIdentifierType	12.1	Predefined value	1	Y if RelatedIdentifier exists	N	Allow cross referencing (data, publication, researcher)	Controlled List Values	DOI
relationType	12.2	Predefined value	1	Y if RelatedIdentifier exists	N	Allow cross referencing (data, publication, researcher)	Controlled List Values	IsSupplementTo
Rights	16	Any rights information for this resource (License)	0-n	N	N		Title of the license	CC BY 4.0 (Creative Commons)
rightsURI	16.1	URI of the license	0-1	N	N			https://creativecommons.org/licenses/by/4.0/
Description	17	All additional information that does not fit in any of the other categories	0-n	N	N		Free text	
descriptionType	17.1	Predefined value	1	Y if Description exists	N			Abstract
GANIL/Other labs Attributes								
topic	6.1	Research domain	1	N/A	Y		Free text	Nuclear Physics; Ion-Matter Interaction; Industrial Applications; interdisciplinary
productionMode	6.2	Emission's type	1	N/A	N		Controlled List Values	In Flight; ISOL; Source
beam	6.2.1	Isotopes	1-n	N/A	Y	From CC Database	Mass number followed by the	68Ni

								Chemical symbol	
beamType	6.2.2	Type of Beam		N/A	Y			Controlled List Values	Isotope; Molecule; Nanoparticles; Other
totalBeamEnergy	6.2.3	Beam energy in MeV		N/A	Y		From CC database	Beam energy followed by the symbol "MeV"	6.72 MeV
detectorType	6.3	Detector's Type	1-n	N/A	N			Controlled List Values	Gamma; Neutron; Charged Particles; Light; None
detectors	6.3.1	Detector(s) used		N/A	Y			Controlled List Values and free text	VAMOS; AGATA; ACTAR; FAZIA
targetType	6.4	Target's nature	1-n	N/A	Y			Controlled List Values	Isotope, Organic Compound, Inorganic Compound, None
targetMatter	6.4.1	Target's Description	0-n	N/A				Free text	