ICM - Institut du Cerveau: ICM - DMP template (English)

I. ROLES

1. DMP author

Exemple de réponse:

John Smith

Sorbonne Université, Institut du Cerveau-Paris Brain Institute-ICM, Inserm, CNRS, APHP, Hôpital de la Pitié Salpêtrière, Paris, France.

Email: john.smith@ICM-institute.org

Recommandations:

Please indicate the author(s) of this Data Management Plan (DMP), including their full name, email, institution, and team, core facility or department.

2.Project funder

Exemple de réponse:

The project funder is INSERM.

Recommandations:

Please indicate the funder of the research project.

3.Sponsor

Exemple de réponse:

The project sponsor is APHP.

Recommandations:

Please indicate the organization which is responsible for research involving the human person (RIPH), ensures its management and verifies that its funding is provided.

If the study does not involve human persons, or only uses secondary human data (RNIPH), no sponsor needs to be identified.

4.Project Controller (“responsable de traitement”)

Exemple de réponse:

ICM is the controller (“responsable de traitement”)

Recommandations:

If the study involves personal data (RIPH or NRIPH), please enter the organization which defines the purposes and the means of personal data processing, i.e., the objectives and the means to achieve them.

5.Partner(s)

Exemple de réponse:

Organization	Unit or lab	City	Country	Description	Role
Nice CHU	sequencing platform	Nice	France	Genomic data acquisition	Processor
ICM	Data Analysis Core	Paris	France	Genomic data analysis and storage	Processor
ICM	Data Analysis Core	Paris	France	REDCap database	Processor (Data Manager)

Recommandations:

Indicate only those partners involved in data processing (collection, analysis, storage, etc.). For each partner, the name and unit/department must be indicated and whether they collect, analyze, store, have access, etc. to the data. This information also helps define the scope of the project.

Possible roles of the partners include either joint controller, when they define the purposes and means of data processing, or processor when they either collect, analyze, store, have access, etc. to the data on behalf of, on instructions from and under the authority of the controller.

6.Scope

Exemple de réponse:

The project will be performed at ICM, Paris, France and consists of a monocentric study only.

Recommandations:

Please specify whether the project is monocentric or multicentric.

The project is monocentric when the research project is conducted in a single recruitment center.

The project is multicentral when the research project is conducted in several recruitment centers.

When involving other centers, indicate whether the project involves local, national, or European partners. If centers are located outside of Europe, please specify the countries.

7.Data management policy

Exemple de réponse:

This Data Management Plan (DMP) is based on the "DMP ICM" model provided by ICM - Institut du Cerveau – Paris Brain Institute (2023, version 2), and is in line with the Data Policy of the Paris Brain Institute (2023, version 1)

Recommandations:

Please indicate which data policy or policies are adhered to in the study. These can include national, funder, sectorial, or departmental procedures or policies. In all cases, the ICM data policy must be adhered to. In addition, this template DMP can be indicated.

8.Data Manager

Exemple de réponse:

The Data Analysis Core will provide a 20% Data Manager for the duration of the project.

Recommandations:

Please define who will have the role of data manager within the project. If no data manager(s) is/are yet defined, consider contacting the Data Analysis Core at ICM, and budgeting data management activity within the project.

9.Access

Recommandations:

Define clearly who (name or function, with affiliation) will have access to the data (raw data, processed data, health data, or results), and how they will be given access to the data. For large data transfers the ICM recommends the use of Globus, and sFTP. Please contact DSI for questions and further information concerning secure data transfer.

Exemple de réponse:

Name	**Organization**	**Data type**	**How (Software/Hardware)**	**Agreement type**
Partners	Dublin brain institute	All data	Account on each software	Partner agreement
Data Managers Zujovic Research Team	Research team, ICM	Clinical	REDCap with individual read-only access	Data Sharing Agreement
Bioinformaticians	DAC, ICM	MRI images	Lustre 2 ICM	
Data engineers	ICM	Microscopy images	Xnat accounts through VPN	
Data engineers	ICM	Cell imaging	Omero accounts through ICM vpn and Lustre 2 ICM	

10. ICM plateforms

- iGENSEQ
- HISTOMICS
- Data Analysis Core (DAC)
- QUANT
- ePHYS
- Centre de Ressources Biologiques (CRB)
- ICV
- CENIR

Recommandations:

Tick here the ICM plateform you are planning to use during your project.

Make sure to contact them before any production is started to plan for it.

II. DATA GOVERNANCE (DURING THE PROJECT)

1.File naming

Recommandations:

Reliable access requires unique and precise naming of data files according to an explicit nomenclature common to all project partners.

This naming convention can (and generally should) include information about the version, e.g., using a date or version number. For example, good naming conventions are the following:

- Maximum 30 characters
- Names of partners (consistent between files)
- Version number (“v1”, “v2”)
- Dates in ISO format: YYYY-MM-DD
- Only standard alphabet letters, numbers, and the following characters are allowed: “ **.**,_ , **-**“
- The following characters will not be used: Special or accented characters, except the underscore (“_”), e.g.: “. : ! ? % & ( ) # / * é è ê à ä ù ç”

Exemple de réponse:

All data stored in REDCap, Xnat and Omero will use the project naming convention implemented in the database.

All other files will be named as follows: “MyProject_Filename_Date_Version_Status”, e.g.: “Profa_ExportAllData_20230320_V1”

2.Metadata

Recommandations:

Specify here the procedures (software and/or methods) as well as the metadata standards used to create and manage metadata for each type of data you are processing.

Metadata are data used to define or describe data, regardless of the type of medium used (printed or electronic).

The use of standard metadata is necessary to make your data findable and reusable, publishable and sharable between partners. Metadata that belong to non-digital data should be digitized as much as possible and done so according to FAIR principles.

Metadata standards are templates that specify all the metadata required to describe a resource.

It is recommended to use a metadata standard specific to your discipline and scientific community.

If none exists, it is possible to create a metadata schema adapted to your needs. In that case, please indicate what type of metadata will be created and how. Examples of available metadata standards are shown here.

Exemple de réponse:

Metadata for each data type will be handled, if possible, with the scientific software handling this data type.

If not possible, metadata will be handled using Excel.

- **REDCap** is used for managing clinical data. REDCap is a browser-based, metadata-driven, community electronic data capture application that allows controlled access to data during the project. REDCap is hosted by the ICM IT department and is validated with all security settings to store sensitive data. REDCap stores many types of metadata in a database that can be exported to a CSV file.
- **XNAT** is used to store MRI data. The DICOM file contains a header that stores patient's annotations (patient's name, age, sex, type of scan, image dimensions, etc) as well as all of the image data (that can contain information in three dimensions). The NIFTI format used for fMRI time series acquisition contains metadata (frequency, phase, and slice encoding directions, per-slice acquisition time, slice acquisition order, etc.) in the NIFTI-1 header.
- **OMERO** is used to store microscopy images.
- **OMICS** metadata will be saved on a spreadsheet summarizing sequencing techniques, kit used, and sequencing specifications, according to MINSEQE guidelines.

3.Quality Control

Recommandations:

Please explain how the consistency and quality of data collection will be controlled and documented.

Access and actions on data should be tracked and documented, and appropriate data quality and consistency checks must be in place, in accordance with those that have been defined at the onset of the study.

This may include processes such as calibration, repeated samples or measurements, standardized data capture, data entry validation, peer review of data, or representation with controlled vocabularies.

Exemple de réponse:

All data will be collected, and their completeness validated by a dedicated data manager before the statistical analyses. The data manager will perform quality checks and checks on data accuracy every 3 months.

REDCap offers a quality assurance workflow to help clinical data studies ensuring completeness and correctness.

A standard operating procedure exists and will be followed by the engineer responsible for data management.

4.Backup during project

Recommandations:

Research data should be regularly backed-up in managed, controlled storage areas. Describe here how data is backed up during the project.

The Paris Brain Institute strongly recommends using the infrastructure provided by the Information Technology department (DSI) to process and store research data and scientific software, considering the sensitivity of the data. Backup conditions are decided together by the research team and DSI, and include access permissions, backup frequency, type and retention period. Per default, the institute has the following backup procedures:

- Lustre file server (unstructured data): Monthly backup
- REDCap (structured text database): Daily backup
- XNAT (structured image database): Daily backup

All backup services are stored on a cloud (AWS S3) located at the EU, and meet the following security measures:

- Full database encryption (AES 256 encryption) to ensure data security
- Protection of exchanges on the internet via an Https certificate
- Management of access to databases and administration interfaces by the DSI
- Rights & access are determined according to user profiles with compartmentalized teams
- VPN protection or encryption of transmissions

Exemple de réponse:

The project benefits from the backup policy of the institute:

- Lustre file server : Monthly backup
- XNAT : Daily backup

All backup services are stored on a cloud (AWS S3) located at the EU, and meet the following security measures:

- Full database encryption (AES 256 encryption) to ensure data security
- Protection of exchanges on the internet via an Https certificate

- Management of access to databases and administration interfaces by the DSI
- Rights & access are determined according to user profiles with compartmentalized teams
- VPN protection or encryption of transmissions

III. DATA COLLECTION (GENERAL)

1.Period

Exemple de réponse:

Start date: 28/10/2020.

Inclusion period: 21 months

Participation period: 7 months ± 30 Days

Total duration: 28 months ± 30 Days

End date: 28/09/2022

Recommandations:

Please indicate the start and finish dates of the data collection period of the project.

2.Description

Recommandations:

Only data that is relevant within the context of the scientific questions defined in the research project (described in 1st tab of Opidor DMP) should be collected. In case of personal and/or animal data, please indicate those in next sections “Data Collection (personal)” and/or “Data Collection (animal)”.

Please specify the data that you will collect, including:

A) Type of data you plan to collect

B) A short description of the data and the purpose for its collection

C) The data format, often clear from the filename extension (e.g., pdf, xls, docx, txt, rdf,).

D) The software used to extract/collect store the data, if any

E) The origin of the data, i.e., where the data is collected

F) Expected total volume during collection of data.

Give preference to open and standard formats and justify the use of non-open formats, e.g., DOC versus RTF. You can find examples and explanations of good practices regarding file formats here.

If the expected data volume is higher than 1To, the DSI needs to be informed.

For questions please contact the Data Management Plan team of the Data Analysis Core.

Exemple de réponse:

This project gathers different types of data:

A) Type	**B) Description and purpose**	**C) Formats**	**D) Software**	**E) Origin**	**F) Exp total Volume**
Clinical data	Clinical assessment to identify patients with genetic mutation X	CVS & XML, PDF, Jpeg, Docx	REDCap	CIC ICM	1Go
Biological data	Human blood samples and urinary pregnancy test to exclude patients for neuroimaging	CVS & XML (REDCap)	Tumorotek	Biobank APHP	1Go
Microscopy data	Microscopy images to detect changes in neuronal arborization, including immunohistochemistry. electron microscopy	OME-TIFF	Omero	Quant ICM	20To
Neuroimaging data	MRI and PET to evaluate cortical inflammation.	DICOM standard (raw data) / NIFTI format (processed ones)	Xnat	Cenir ICM	1To
Electroencephalography	EEG to determine modulated cortical responses			Cenir ICM	80Go
OMICS	RNA Seq for Lymphocyte profiling	CSV		Igenseq ICM	200Go
Video files	Patient’s body movement recording to detect anormal movement link to Parkinson	MKV		Aphp	50Go
Statistical scripts	Scripts for statistical analysis in R	ASCII Text files (R)			<1Go
Source code	Scripts for data analysis	.MAT (Matlab)			<1Go

IV. DATA COLLECTION (ANIMAL)

1.Period

Exemple de réponse:

Start date: 28/10/2020.

Total duration: 28 months ± 30 Days

End date: 28/09/2022

Recommandations:

Please indicate the start and finish dates of the animal data collection period of the project.

2.Description

Recommandations:

Only data that is relevant within the context of the scientific questions defined in the research project (described in 1st tab of Opidor DMP) should be collected.

Please specify the data that you will collect, including:

A. Type of data you plan to collect

B. A short description of the data and the purpose for its collection

C. The data format, often clear from the filename extension (e.g., pdf, xls, docx, txt, rdf, etc.). Give preference to open and standard formats and justify the use of non-open formats, e.g., DOC versus RTF. You can find examples and explanations of good practices regarding file formats here.

D. The software used to collect the data

E. The origin of the data, i.e., where the data is collected

F. Expected total volume during collection of data. If the expected data volume is higher than 1To, the DSI needs to be informed.

For questions please contact the Data Management Plan team of the Data Analysis Core.

Exemple de réponse:

This project gathers different types of data:

A) Type	B) Description and purpose	C) Formats	D) Software	E) Origin	F) Exp total Volume
Animal sleep behaviour	24 hour video recordings	.MOV	In-house software	ICM animal facility	1Go

3.Authorization

Exemple de réponse:

The principles of the 3Rs (Replacement, Reduction and Refinement) will be used to guarantee ethical regulations for humane animal research:

Replacement: To conduct our project, there is currently no alternative to the use of animals.

Reduction: the experiments are designed to use a minimal number of animals consistent with reliable statistical significance.

Refinement: During and after surgeries, animals will receive the best analgesic coverage. We will closely observe the recovery and well-being of the animals. If any complication is observed, animals will be removed from the study with fast and humane euthanasia.

APAFIS number: 23542323

date of APAFIS: 12/01/2025

All animal protocols are performed in accordance with the guidelines published in the National Institute of Health Guide for the Care and Use of Laboratory Animals. All the data collected from animal testing is validated by an ethical committee research with partners outside Europe, respecting the H2020 manual.

Recommandations:

Animal research must be authorized by APAFIS. Please report the APAFIS number and date here.

The principles of the 3Rs (Replacement, Reduction and Refinement) need to be used to guarantee ethical and humane animal research. Please specify how the research respects these three aspects.

V. DATA COLLECTION (PERSONAL)

1.Period

Exemple de réponse:

Start date: 28/10/2020.

Inclusion period: 21 months

Participation period: 7 months ± 30 Days

Total duration: 28 months ± 30 Days

End date: 28/09/2022

Recommandations:

Please indicate the start and finish dates of the personal data collection period of the project.

2.Description

Recommandations:

Only data that is relevant within the context of the scientific questions defined in the research project (described in 1st tab of Opidor DMP) should be collected.

Please indicate here what type of personal data you will be collecting, and for what purpose. For each data type, indicate the tool used to collect and securely store the data. We recommend following best practices in personal data collection described here.

Please also indicate the (estimated) number of patients or participants of whom data will be collected. This is required for estimating risk in the Privacy Impact Assessment (see below).

You can use the following categories to help identify the type of data:

A. Civil status, identity, identification data (name, surname, sex, order number, photograph, date and place of birth)

B. Professional life (CV, training, diplomas, distinctions, professional situation)

C. Personal life (life habits, family situation)

D. Localization or geolocation data (address, phone number, movements, GPS data, GSM)

E. Economic and financial situation (income, bank data)

F. Connection data (IP addresses, connection identifiers, computer traces, time stamps)

G. Internet (cookies, tracers, navigation data, audience measurements)

H. Sensitive data (racial or ethnic origins, political, philosophical, religious opinions, trade union membership, SS number, health, genetic or biometric data, life or sexual orientation, convictions, or offences

I. Other (please specify)

Exemple de réponse:

Category (*)	**Exhaustive list of processed data and purpose**	**Origin**	**Software**	**Conservation during collection and analysis period**	**Archiving duration**
A)	Patient Sex for data clustering Order number for pseudonymization	ICM	REDCap	2 Years	15 Years
B)	Last degree obtained. Professional status (unemployment, retirement,)	ICM	REDCap	2 Years	15 Years
C)	French language skills	ICM	REDCap	2 Years	15 Years
D)	Postal code	APHP	ORBIS	2 Years	15 Years
E)	bank details (RIB/IBAN) for payment	APHP	ORBIS	2 Years	15 Years

(*) Category explanation:

A. Civil status, identity, identification data (name, surname, sex, order number, photograph, date and place of birth ...)

B. Professional life (CV, training, diplomas, distinctions, professional situation ...)

C. Personal life (life habits, family situation)

D. Localization or geolocation data (address, phone number, movements, GPS data, GSM, ...)

E. Economic and financial situation (income, bank data, ...)

F. Connection data (IP addresses, connection identifiers, computer traces, time stamps, ...)

G. Internet (cookies, tracers, navigation data, audience measurements, ...)

H. Sensitive data (racial or ethnic origins, political, philosophical, or religious opinions, trade union membership, SS number, health, genetic or biometric data, life or sexual orientation, convictions or offences

I. Other

3.Informed consent

Recommandations:

The implementation of the research involving personal data requires the following:

- An information sheet about research
- An information sheet about data processing (you will find by clicking on the link below a checklist of mandatory information and a standard information sheet: ...)
- A consent certificate to research (when required)

Please indicate in your response the following:

- How the information sheet is delivered to the study participants
- How the consent certificate to research is obtained (when required)
- How consent forms are stored and who has access to them (when required)

Exemple de réponse:

A written information document will be given to all potential participants to provide them with all the information to allow them to decide whether to participate in this study, including any potential risks, and including the mandatory mentions of articles 13 and 14 of the GDPR.

A research consent form will be signed by those wishing to participate in the research.

The consent forms will be stored physically in a safe that is only accessible by the principal investigator.

4.Pseudonymization

Exemple de réponse:

The pseudonymization of all data will be done using ICM pseudonymization software called Identity Manager. This solution provides a simple and secure interface to store identifying data in a physical database separate from scientific data. Identifying data will be collected in Identity Manager, which will then generate a custom project identifier, which will be used to collect scientific data from REDCap.

Recommandations:

Only indirectly identifying data can be used in a research protocol, except data collected in the correspondence table.

Indicate here whether the data that is collected allows for direct identification (e.g., name), or whether data is pseudonymized, i.e., only allowing indirect identification. Indicate the method of pseudonymization by which this is done to secure the confidentiality of persons of whom data is collected. For an explanation of pseudonymization, see here

5.Privacy Impact Assessment (PIA)

Recommandations:

A Privacy Impact Assessment (PIA) must be created when the processing of data might pose a risk to the rights and freedoms of natural persons. The sensitivity of the data processed, their quantity, the category of data subjects all affects the impact assessment and are therefore required information.

A **PIA** is required when:

1. The data processing is on the list for which the CNIL considers a PIA mandatory
2. Or the data processing does not respects a reference methodology
3. Or the data processing is necessary for the constitution of a health data warehouse
4. Or two or more of the following questions are answered affirmatively:

- Assessment/scoring (including profiling)
- Automatic decision with legal or similar effect
- Systematic monitoring
- Collection of sensitive or highly personal data
- Collection of large-scale personal data
- Cross-referencing of data
- Vulnerable persons (e.g., patients, elderly or minors)
- Innovative use (use of new technology)
- Exclusion from the benefit of a right or contract

You can also follow the decision tree for PIA.

After you receive the PIA, you can upload it here to combine it easily as part of the DMP.

Please contact the DPO to evaluate the need, and if so, perform the PIA, after providing (to the best of your ability) the information outlined above.

Exemple de réponse:

The research project will involve the processing of health data of vulnerable persons (e.g., patients, elderly or minors).

A PIA will be made by the DPO of the ICM according to the ICM template.

6.Reference methodology

Exemple de réponse:

The project conforms to a reference methodology of the CNIL (MR003: 2211517 v0).

Recommandations:

Please indicate whether the research project conforms to a reference methodology (RM):

- The MR-001 concerns *interventional research*, i.e., intervention on human persons that is not justified by his/her usual care, research involving only minimal risks and constraints, and research requiring an examination of genetic characteristics.
- The MR-003 concerns non-interventional research involving humans, mentioned in the third paragraph of Article L. 1121-1 of the Public Health Code.
- The MR-004 is dedicated to research, studies or evaluations in the health field that do not involve human persons and are of public interest.

If you comply to a reference methodology at ICM, and ICM is responsible for the data treatment, you can refer directly to the conformity declaration and use the following identifiers:

- MR-001: 2211516 v0
- MR-003: 2211517 v0
- MR-004: 2217035 v0

7.CNIL authorization

Exemple de réponse:

CNIL authorization number is 123456

Recommandations:

If no reference methodology (previous section) is applicable, an authorization from the CNIL is required. You can upload your application form for to the CNIL here to simplify communication with the DPO. Once you have received the CNIL authorization number, please indicate it here.

VI. REUSING DATA (GENERAL)

1.Description

Recommandations:

Please specify the data that you will reuse, including:

A. Type of data you plan to reuse

B. A short description of the data and the purpose for its reuse

C. The data format, often clear from the filename extension (e.g., pdf, xls, docx, txt, rdf, etc.).

D. The software used to extract/collect store the data, if any

E. The origin of the data, i.e., where the data has been shared, stored or collected

F. Expected total volume. If more than 1Tb, please inform the IT department.

Please indicate whether a Data Sharing Agreement (DTA) exists, and if so, please add the reference.

Exemple de réponse:

A) Type	B) Description and purpose	C) Formats	D) Software	E) Origin	F) Exp total Volume
EEG	EEG record of interval timing in Parkinsons patients (anonymized)	.eeg, .vhdr, .vmrk	EEGlab	http://predict.cs.unm.edu/	50Go

2.Origin and agreement

Exemple de réponse:

Status: Published Dataset

Origin: Harvard University, USA

Publications DOI: 1234987546

Terms of re-use: Comparison of data for reproducibility.

The research will comply with these terms by not re-sharing the data, and publishing results solely in terms of reproducibility.

Recommandations:

Indicate the origin and the publications related to the data that is being reused, e.g., by referring to their Data Object Identifier (DOI). If material is protected by specific rights specify the access or re-use restrictions found in their user agreements. If there are restrictions on access or re-use, explain the means used to ensure compliance with those terms. In case you plan to add new data to these existing datasets, state if the existing license allows this.

VII. REUSING DATA (PERSONAL)

1. Description

Exemple de réponse:

A) Type of data	B) Description and purpose	C) Format	D) Software	E) Origin of data
Name, Date of Birth, clinical history, genetic markers	Genetic disease re-analysis of new markers	CSV	IDManager	ANR project 12345 (2012), PI: John Walker

Recommandations:

Describe the data that you plan on re-using:

A. Type of data

B. A short description of the data and the purpose for its re-use

C. The data format, often clear from the filename extension (e.g., pdf, xls, docx, txt, rdf, etc.)

D. The software used to extract/collect store the data (if any)

E. The origin of the data, i.e., where the data has been shared, stored or collected

Please indicate whether a Data Sharing Agreement (DTA) exists, and if so, please add the reference.

If data comes from an INSERM or AP-HP clinical trial, please check with INSERM or AP-HP before reusing such data.

2.Informing participants

Recommandations:

The reuse of personal data is considered processing of data.

This means that research participants must be informed of this.

Only the re-use of anonymous (not pseudonymous) data is not subject to the information obligation.

Explain here how the data subjects have been informed of the reuse of their personal data.

It is highly recommended anticipate the re-use of personal data by the following practices:

- Informing participants of the possible re-use of their personal data and acquiring their consent
- Directing participants to the Transparency Portal where they can find research projects reusing their personal data
- Explaining to participants how to object to the re-use of their personal data

Contact the DPO if you need more information.

Exemple de réponse:

At the end of this research, and if participants agreed to the principle, their coded personal data, including their biological samples, may be used in the framework of further research projects in collaboration with private or public partners, in France or abroad.

Participants will find all the necessary information specific to these projects at the following address: https://institutducerveau-icm.org/fr/recherche/etudes-clinique/.

In the context of personal data re-use, personal data may be transferred outside the European Union.

The transfer of personal data will be governed by appropriate and adapted guarantees in a contract concluded between Paris Brain Institute and the recipient of the personal data. Participants have the right to obtain a copy of these guarantees.

Participants can object to the re-use of their personal data by contacting the investigator.

VIII. PROCESSING DATA

1.Period

Exemple de réponse:

Start date: 28/10/2020.

End date: 28/09/2022

Recommandations:

Please indicate the start and finish dates of the data processing period.

2.Volume

- Other (detail)
- >1To
- 500Go - 1To
- 50Go - 500Go
- < 50Go

Recommandations:

Describe how much space you will need for processing the data you collected.

IX. DATA GOUVERNANCE (AFTER THE PROJECT)

5.Archiving

Recommandations:

Long-term preservation of data is a task for the research institution. At ICM, archiving (long-term storage) is provided by the IT services on demand, using Amazon Web Services (“Glacier”) geographically based in EU. Please contact the DSI for more information.

Alternatively, data collected in the framework of a limited (in scope and in time) project can be stored via the scientific journal storage facilities when the results are published. Most journals and funding bodies have specific requirements. Some funding bodies, as well, provide lists of recommended repositories.

You should only archive data (long-term storage) that might be reused in future research, e.g., data that are:

- Unique, non-reproducible or difficult to reproduce
- Necessary for a global understanding of the research
- Necessary to validate the results presented in a scientific publication

Please indicate here:

- What data will be archived
- When data will be archived
- How the data will be archived, e.g., on a local or distant server, external drive, or repository. When using a data repository, specify the repository or cloud service and the arrangements in place. Preference should be given to certified repositories supporting open access, when possible.
- The lifetime of the data, considering the existing legal or prescribed requirements.
- The total (anticipated) volume

Exemple de réponse:

Long term storage of raw image files is performed every month using AWS cold storage.

The expected size over the course of the project (2 years) is 35 Tb.

6.Destroying

Recommandations:

Specify whether some types of data will be destroyed (such as personal data according to CNIL guidelines), when and how.

X. PUBLISHING DATA AND SOFTWARE

1.Description

Recommandations:

Describe the type of data, where or how they will be publicly available, and under what license. If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.

In the user agreement or license specify the conditions for access (whether open or restricted with controlled access procedures). Also specify whether there are methods or software tools needed to access the data (e.g., web access or open-source code).

Exemple de réponse:

Type	**Description**	**Where**	**License**
Analysis scripts	Analysis underlying publication DOI:0.1/12345	gitlab.com/myproject/myscripts (web access)	No commercial Use - No Modification CC BY-NC-ND
Data	Simulated data for predicting recovery	https://eosc-portal.eu/ DOI:0.2/67890 (web access)	Creative Commons CC2.0

At the end of the project, raw data from RNA sequencing will be hosted at the European nucleotide archive (https://www.ebi.ac.uk/ena), under the CC BY 4.0 license.

The repository uses a predefined and rich list of metadata.

2.Intellectual Property Rights (IPR)

Recommandations:

Please indicate here:

- Whether there are IPR or copyrights issues to consider when publishing the products of the research
- Whether there are agreements with other stakeholders concerning publishing the products of the research
- Whether permission is needed to collect/reuse the products of research
- Whether rights will be transferred to another organization for data distribution and archiving
- Whether other restrictions need to be considered

Please contact CART for any questions or uncertainties regarding these topics.