ICM - Institut du Cerveau: DMP template V2.sw (English) [IN CONSTRUCTION]

ROLES

Project funder

Recommandations:

Please indicate the funder of the research project.

Exemple de réponse:

Ex. 1: The project funder is Paris Brain Institute (ICM), Paris, France

Ex. 2: The project funder is *L'Institut national de la santé et de la recherche médicale* (Inserm), France

Partner(s)

Exemple de réponse:

Organization	**Unit or lab**	**City**	**Country**	**Description**	**Role**
Institute Pasteur	Team Bakker	Paris	France	Sample preparation & analysis	Processor
Nice CHU	sequencing platform	Nice	France	Genomic data acquisition	Processor
Paris Brain Institute (ICM)	Histomics platform	Paris	France	Microscopy Data acquisition	Processor
Paris Brain Institute (ICM)	iGenSeq sequencing platform	Paris	France	Genomic data acquisition	Processor
Paris Brain Institute (ICM)	Data Analysis Core	Paris	France	Genomic Data analysis and Storage	Processor
Paris Brain Institute (ICM)	Data Analysis Core	Paris	France	REDCap database	Processor (Data Manager)

Recommandations:

Indicate only those partners involved in data processing (collection, analysis, storage, etc.).

- **Description** specifies whether they collect, analyze, store, have access, etc. to the data. This information also helps define the scope of the project.
- **Role** can be *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 *data controller*.

Data Manager

Exemple de réponse:

Ex. 1: A Data Manager will be recruited to ensure quality control and FAIR data collection, archiving and data sharing.

Ex. 2: The Data Analysis Core will provide a 20% Data Manager for the duration of the project and will provide a data quality report every 2 month.

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 (dac@icm-institute.org), and budgeting data management activity within the project.

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.

All ICM policies are available here : https://institutducerveau-icm.org/en/our-commitments/

DATA GOVERNANCE DURING THE PROJECT

Where is your data stored during the project?

Recommandations:

Indicate whether your data are stored on:

- Your computer
- Shared storage space provided by IT
- External file server

Be careful, you absolutely must **not** store your data on free online storage (e.g., Dropbox, Google Drive, OneDrive), nor on external hard drives, because these are generally neither safe nor secure.

Exemple de réponse:

Ex. 1: As the project only includes collaborators and personnel at the Paris Brain Institute, a common project folder has been created on the storage servers with the name: *ProjectID*.
Ex. 2: Each partner stores the data it produces on its own servers. At the Paris Brain Institute, a common project folder has been created on the storage servers with the name: *ProjectID*.

Data organization

Recommandations:

The filing classification scheme refers to the folder tree structure set up to classify data files. The purpose of this organization is to allow any collaborator to easily and quickly locate and retrieve the data he or she needs, as well as facilitate the versioning, backup and archiving of data.

We strongly recommend to follow the example response.

Exemple de réponse:

A file classification scheme has been created for the common storage space of all project partners.

Data is first organized by collection method (microscopy, phenotyping, sequencing...) and then chronologically.

Raw data and processed data are stored in different folders.

Below is an overview of the classification scheme:

- I. Collection method 1 (e.g. microscopy)
- I.1 Date of first collection (e.g. 2021-01-12)
- I.1.1. Processing step 1 (ex: quality check)
- I.1.2. Processing step 2 (e.g. raw data)
- I.1.3. Processing step 3 (e.g.: analyzed data)
- I.2 Date of second collection (e.g. 2021-01-19)
- II. Collection method 2 (e.g. sequencing)

File naming convention

Exemple de réponse:

All documents, regardless of data type, will be named according to the following pattern :

- “[ProjectID]_[Date]_[description]_[version].[extention]”, e.g.: “Profa_20230320_A1B2_1.tiff”

For the sake of human readibility and compatibility across operating systems, the following additional conventions are required:

- Limit file names to no more than 25-35 characters.
- Use leading zeros to facilitate numerical sorting, e.g.: 001, 002, 010, 011, 100, 101, ... instead of 1, 2, 10, 11, 100, 101, ...
- Use a period followed by a file extension, e.g.: .tiff, .jpg, .gif.
- Avoid uppercase letters.
- Avoid symbols or spaces that could cause complications across operating platforms, e.g.: ! ? % & ( ) # / * é è ê à ä ù ç
- Avoid blank spaces within the character string; use hyphens or underscores instead.
- Avoid special or accented characters, e.g.: ! ? % & ( ) # / * é è ê à ä ù ç.
- Avoid overly complex or lengthy naming schemes prone to human error, such as filenameconventionjoesfinalversioneditedfinal.doc.

Recommandations:

Access that is reliable over time and between collaborators requires the agreement of unique and precise naming of data files. For example, good naming conventions are:

- Unique and consistently structured, including the project identifyer, date, description and version.
- Persistent and not tied to anything that changes over time or location

For the sake of human readibility and compatibility across operating systems, the following additional conventions are required:

- Limit file names to no more than 25-35 characters.
- Use leading zeros to facilitate numerical sorting, e.g.: 001, 002, 010, 011, 100, 101, ... instead of 1, 2, 10, 11, 100, 101, ...
- Use a period followed by a file extension, e.g.: .tiff, .jpg, .gif.
- Avoid uppercase letters.
- Avoid symbols or spaces that could cause complications across operating platforms, e.g.: ! ? % & ( ) # / * é è ê à ä ù ç
- Avoid blank spaces within the character string; use hyphens or underscores instead.
- Avoid special or accented characters, e.g.: ! ? % & ( ) # / * é è ê à ä ù ç.
- Avoid overly complex or lengthy naming schemes prone to human error, such as filenameconventionjoesfinalversioneditedfinal.doc.

Quality Control

Exemple de réponse:

Ex. 1: In order to guarantee the quality of the data, various measures have been implemented:

- Independent repetition of the experiments (minimum of three repetitions on three different days)
- Standardization of data collection (all animals raised under the same conditions, temperature control, same stimulation conditions)
- Regular review of data with PI

Ex. 2: 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.

Ex. 3: 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 data manager.

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.

Describe who has access to which data, and under which agreement (if external to ICM).

Exemple de réponse:

Name or function	**Organization**	**Data type**	**How (Software/Hardware)**	**Agreement type**
Partners	Dublin brain institute	Health data (including samples)		Data Transfer Agreement
Data Managers Pr Smith Research Team	DAC, ICM	Clinical	REDCap with individual read-only access	
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	
Claire Smith Julien Smith Mathieu Smith Pierre Smith	ICM	Living animal	Phenoparc computers	

Recommandations:

- **Name or function** of who will have access to the data during the project.
- **Data type** needs to be specified.
- **How (Software/Hardware)** access will be provided during the project. Contact Front-Office-DSI@icm-institute.org for support in secure data transfer.
- **Agreement type** is required when external collaborators or partners are given access. Contact legal@icm-institute.org for support in agreement creation.

Backup & versioning during project

Recommandations:

Research data should be regularly backed-up in managed, controlled storage areas.

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
- Identity Manger (Personnal data) : 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

For further assistance, add share this DMP with Front-Office-DSI@icm-institute.org, and request support.

Exemple de réponse:

The project benefits from the backup policy of the institute:

- Lustre file server: Monthly backup
- REDCap (structured text database): Daily backup
- XNAT: Daily backup
- Identity Manger (Personnal data): 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

REUSING DATA

Data re-use

Recommandations:

Please state the reasons if re-use of any existing data has been considered but discarded.

Description of data that will be re-used.

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 by referring to their Data Object Identifier (DOI)

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

G. Indicate whether a re-use **agreement or conditions** exists, and under which conditions. 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. If a Data Transfer Agreement (DTA) exists, and if so, please add the reference. Please share the DMP with legal@icm-institute.org and ask for support if needed.

Exemple de réponse:

A) Type	**B) Description and purpose**	**C) Formats**	**D) Software**	**E) Origin**	**F) Exp total Volume**	**G) Agreement or conditions**
EEG	EEG record of interval timing in Parkinsons participant (anonymized)	.eeg, .vhdr, .vmrk	EEGlab	http://predict.cs.unm.edu/	50Go	Creative Commons (CC0)

DATA COLLECTION

Data collection period

Recommandations:

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

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

Description

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 participants with genetic mutation X	CVS & XML, PDF, Jpeg, Docx	REDCap	CIC ICM	1Go
Biological data	Human blood samples and urinary pregnancy test to exclude participants 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, JPEG, TIFF, ND2 (Nikon confocal), CZI (Apotome, Axioscan), NDPI (Nanozoomer), LIF (confocal)	Omero, Zen, Nanozoomer	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	FASTQ, BCF	Seurat	Igenseq ICM	200Go
Animal sleep behaviour	24 hour video recordings	.MOV, .szk .szd .szp .szv .mbd .dsn .mpg .rst .tcg .tcr	Clever system, TopScan, Anymaze	Phenoparc	1To
Animal behaviour	Mutation/lesion impact on behavior	.xlsx	Excel	Phenoparc	<1Go
Animal electrophysiological data	Speed of nerve conductance	.csv	Mobius Med64	Ephys ICM	<1Go
Cell sorting (cytometry)	FACS-sorting	FCS, PCP	FlowJo	ICAN	
Video files	participant'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

Recommandations:

Data should only be collected when relevant within the context of the scientific questions defined in this research project. 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,). Check the list here

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** of data during collection. If the expected data volume is higher than 1To, the DSI needs to be informed at Front-Office-DSI@icm-institute.org.

If open and standard formats will not be used, justify this, e.g. due to the use of certain software, or skills of technicians. You can find examples and explanations of good practices regarding file formats here.

LIVE ANIMAL DATA COLLECTION AND EXPERIMENTATION

Data collection 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.

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 necessitates authorization from APAFIS. Each project involving animal experimentation requires two aspects: a project validated by APAFIS and competent personnel. Please report the APAFIS number assigned to this project. (One number for the project or for each animal experimentation group).

PERSONNAL DATA COLLECTION

Period

Recommandations:

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

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

Description

Exemple de réponse:

We will be collecting personal data from 245 participants, that will include the following:

Category (*)	**Exhaustive list of processed data and purpose**	**Origin**	**Software**	**Conservation during collection and analysis period**	**Archiving duration**
A)	Participant 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

Recommandations:

Please indicate here what **type** of personal data you will be collecting, and for what **purpose**. Only data that is relevant within the context of the scientific questions defined in the research project should be collected. For each data type, indicate the

software used to collect and securely store the data. We recommend following best practices in personal data collection described here.

You are also obliged to specify the **data retention duration** after project completion. In France the retention period is 15 years, but this could change depending on European or French law. E.g., it is 25 years when pertaining to clinical trials on medicinal products for human use (EU regulation 536/2014). Please contact legal@icm-institute to validate.

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)

Informed consent

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 locked and secure place that is only accessible by the principal investigator.

Recommandations:

Please read the template information sheet here

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

- An information sheet about research.
- An information sheet about data processing.
- A consent certificate for research (when required).

Please indicate in your response the following:

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

You can share your DMP with the DPO to ask for assitance: dpo@icm-institute.org

4.Pseudonymization

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

For any help about this question, please to add this contact in your DMP and send an email to : dpo@icm-institute.org (DPO).

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.

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:

- Collection of sensitive data (such as health data) or highly personal data
- Data processing of vulnerable persons (e.g., patients, elderly or minors)
- Collection of large-scale personal data
- Cross-referencing data from separate databases
- Innovative use (use of new technology)
- Exclusion from the benefit of a right or contract
- Transfer of personal data outside the European Union
- Assessment/scoring (including profiling)
- Automatic decision with legal or similar effect for the data subjects
- Systematic monitoring of data subjects

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 : dpo@icm-institute.org

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.

DPO : dpo@icm-institute.org

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.CNIL Reference methodology

Recommandations:

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

- The RM-001 concerns personal data processing in the context of health research with the consent of data subjects
- The RM-003 concerns personal data processing in the context of health research not requiring the consent of data subjects.
- The RM-004 is dedicated to research in the context of research that do not involve human person.

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

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

For any help about this question, please to add this contact in your DMP and send an email to : dpo@icm-institute.org (DPO).

Exemple de réponse:

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

7.CNIL authorization

Recommandations:

If no reference methodology (previous section) is applicable, an authorization from the CNIL is required.

Please, contact the DPO : dpo@icm-institute.org

Once you have received the CNIL authorization number, please indicate it here.

Exemple de réponse:

CNIL authorization number is 123456

Please 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).

VII. REUSING PERSONAL DATA

1. Description

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.

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

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 : dpo@icm-institute.org

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

Recommandations:

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

Exemple de réponse:

Start date: 28/10/2020.

End date: 28/09/2022

2.Volume

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

Recommandations:

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

For any help about this question, please to add this contact in your DMP and send an email to : Front-Office-DSI@icm-institute.org (DSI).

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) if is required by law or 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

For any help about this question, please to add this contact in your DMP and send an email to : Front-Office-DSI@icm-institute.org (DSI).

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, when and how.

For any help about this question, please to add this contact in your DMP and send an email to : Front-Office-DSI@icm-institute.org (DSI).

X. PUBLISHING DATA

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).

For any help or any information about copyright, please contact the direction of innovation : di@icm-institute.org

Exemple de réponse:

Type	**Description**		**License/Agreement**	**Raw data (yes/no)**
Data	Simulated data for predicting recovery	https://eosc-portal.eu/ DOI:0.2/67890 (web access)		No

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 legal department for any questions or uncertainties regarding these topics : legal@icm-institute.org

XI. PUBLISHING SOFTWARE

1.Description

Exemple de réponse:

Type	Description		License/Agreement	Raw data (yes/no)
Analysis scripts	Analysis underlying publication DOI:0.1/12345	gitlab.com/myproject/myscripts (Web access)	No commercial Use - No Modification CC BY-NC-ND	No

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.

Recommandations:

Describe the type of software, where or how they will be publicly available, and under what license.

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).

For any help or any information about copyright, please contact the direction of innovation : di@icm-institute.org

2.Intelectual Property Right (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