

MASA Consortium: MASA DMP template V.2 - Context

Description of the project and Keywords

Project type

- Publication: publishing of sources from the literature
- Publication: catalog
- Valorization
- Mediation
- Publication: publishing of epigraphic sources
- Other
- Digitization
- Study
- Survey
- Archaeological inventory or prospecting
- Archaeological excavation or prospecting

Recommandations:

If necessary use the "additional information" field to describe the project more precisely by indicating the discipline, theme or type of work carried out.

For example: archaeozoology, experimental archaeology, preventive archaeology, underwater work, heritage inventory, etc.

Exemple de réponse:

Digital epigraphic publication using TEI/EpiDoc encoding.

Type of funding

- Doctoral research project
- Preventive archaeology
- Survey (France)
- Excavation programme (outside France)
- Project funded by the European Commission
- Collective research project
- Other (specify)
- ANR-funded project

Inclusion of the project in a scientific programme with funding

Recommandations:

If the project is part of a scientific programme with funding from a private or a public institution, specify the scientific theme areas associated with that programme.

For example:

- H2020 Programme
- Scientific theme area or programme of a research structure linked to the project leader or team.

Chronology

Recommandations:

Indicate here the keyword(s) that describe the chronological period associated with the project. If the repository is online and has perennial identifiers (Uniform Resource Identifier or URI) or "permalinks", link the keyword to this identifier.

If possible, use a controlled vocabulary, such as PeriodO (<https://perio.do/en/>) or PACTOLS (<https://pactols.frantiq.fr/>)

Exemple de réponse:

Upper Palaeolithic (<https://ark.frantiq.fr/ark:/26678/pcrt9U8BH9pVRu>)

Sites

Exemple de réponse:

Cyprus (<https://ark.frantiq.fr/ark:/26678/pcrtS4MSAHu2F3>)

Salamis (Cyprus) (<https://ark.frantiq.fr/ark:/26678/pcrtMSZVqxly6Q>)

Recommandations:

If possible, use a controlled vocabulary.

If the repository is online and has perennial identifiers (Uniform Resource Identifier or URI) or "permalinks", link the keyword to this identifier.

Examples of geographical reference domains and indexes which are useful for projects related to archaeology or the study of ancient worlds:

- [Pleiades](#) geo-historical index of ancient places;
- [GeoNames](#) online index;
- "Places" domain of the [Pactols](#) thesaurus;

If it is an archaeological projection, fill in the coordinates according to the **Lambert93 projection** (the official projection for maps of mainland France which is linked to the RGF93 geodetic system). Indicate the level of **accuracy used** and justify this if necessary (e.g. use of the municipality's barycentre to protect the site).

Subject or theme

Recommandations:

If possible, use a controlled vocabulary. If the repository is online and has persistent identifiers (Uniform Resource Identifier or URI) or "permalinks", link the keyword to this identifier.

Using the [Pactols](https://pactols.frantiq.fr) thesaurus (<https://pactols.frantiq.fr>) or the Art and Architecture Thesaurus (www.getty.edu/research/tools/vocabularies/aat/) is recommended.

Exemple de réponse:

hut
post-hole
fire place
daub
palisade
social groups
modern pottery
animal bones
metal object
glass
architectural terra-cotta

Identification of associated archaeological operations (optional)

Identifier of the operation or operation(s)

Recommandations:

This field only needs to be filled in if this number exists. For example for excavations in France, indicate the identifier of the operation in the national reference system here.

Name of the operation or operation(s)

Recommandations:

Title(s) of the archaeological operation(s) as it/they appear(s) in the reference system for archaeological operations or as it/they appear(s) in the official authorization.

Administrative status of the operation

Exemple de réponse:

Preventive excavations, scheduled excavations, Ministry of Foreign Affairs excavations, etc.

Nature of the intervention(s)

Exemple de réponse:

Diagnostic, excavations, prospecting, works supervision, underwater excavations, etc.

Director of the excavations

Exemple de réponse:

Perrot, Jean / CNRS.

Recommandations:

Several personal identifier systems can be used. If the person identifier does not exist in ORCID, other registries may be used (e.g. ISNI, IDREF, VIAF, wikidata, etc.).

Geographical coordinates of the operation

Recommandations:

Indicate the Lambert 93 coordinates here.

Indicate the level of precision used, and justify it if necessary (e.g. use of the town's barycentre to protect the site)

MASA Consortium: MASA DMP template V.2 - Data management

Responsibilities

Head(s) of data management for the project

Recommendations:

The data manager(s) must be involved in the project from its creation to its completion. He/she is responsible for the collection, organization and storage of the data.

He/she is often a different person from the scientific manager.

The responsibilities shared between several sites, partners and/or institutions must be described here. Provide the name of the person in charge of the data management coordination.

Exemple de réponse:

Hugo, Saturnin (saturnin.hugo@cns.fr) / CNRS USR 1000. <http://orcid.org/0000-0003-4317-1800>.

Other people involved in the data management

Recommendations:

Set out the roles, duties and responsibilities of all those involved in the data management, e.g. data entry/capture, metadata modelling, metadata production, data quality management, file management, data storage and backup, archiving, data sharing...

You may also describe the data workflow process in order to clarify the different responsibilities.

For collaborative projects, explain how data management responsibility shared between partners is coordinated.

Resources

Resources (budget and time) assigned to data management.

Recommendations:

Explain how the resources required (time for example) to prepare data for sharing or the preservation (curation) of data have been evaluated.

Carefully review and justify all resources required to ensure that the data will be FAIR (Findable, Accessible, Interoperable, and Reusable)

Definition of the FAIR principles

[FAIR Principles - GO FAIR \(go-fair.org\)](https://www.go-fair.org/)

Quality control

Data and file quality control

Recommendations:

Indicate here whether a data and file quality control system was set up at the project level or possibly at the dataset(s) level.

You can refer to the Guidelines on FAIR Data Management Horizon 2020

https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

The Facil file format validation tool can be used to check that the file formats are eligible for archiving by the National Computing Center for Higher Education (CINES):

<https://facile.cines.fr/>

MASA Consortium: MASA DMP template V.2 - Data overview

Constructing a corpus

Origins of the data

Exemple de réponse:

Collection of new data, conversion, transformation of existing data, sharing or exchange of data, buying data

Organisation

File classification tree

Exemple de réponse:

Example of a file classification tree for a project linked directly to one or more archaeological operations (INRAP, France):

- Raw data
- Scoping documents
- Spatial data (all the georeferenced data)
- Study (data and documentation from the processing of raw data stored in the "data" folder; specialized studies)
- Illustration (intermediate files which were necessary to create the figures produced in the context of the project for promotion and dissemination such as posters, publications etc.).

Recommandations:

Indicate here the classification tree of the files used and produced during the project. This must be set up at the start of the project and used throughout the project.

For advice on developing a classification plan, you can refer to the guide "Traçabilité des activités de recherche et gestion des connaissances: guide pratique de mise en place" (Traceability of research activities and knowledge management: a practical implementation guide) developed by the CNRS Mission for Transversal and Interdisciplinary Initiatives (2018) which is available on Doranum,

Nature of the datasets described

Recommandations:

List the datasets here and briefly describe their contents.

If you develop a software, you can use the PRESOFT management plan template, which is available on DMP OPIDoR and is documented at <http://www.france-grilles.fr/presoft/>

Volume

Expected file size

Recommandations:

Indicate the unit used (Go, Mo, To, etc.)

You will find examples in the guide "[Traçabilité des activités de recherche et gestion des connaissances: guide pratique de mise en place](#)" (Traceability of research activities and knowledge management: a practical implementation guide) developed by the CNRS Mission for Transversal and Interdisciplinary Initiatives (2018).

Final file size

Recommandations:

Indicate the unit used (Go, Mo, To, etc.)

Final number of files

Data description

Data Producer

Recommandations:

If they are only produced by one person, this person should be cited as the producer of the dataset.

If the data in the dataset were produced collectively, use a generic signature such as "XXX project team".

Specify: Last name, First name/ Affiliation, function within the project. Personal identifier.

Exemple de réponse:

Rodier, Xavier/ CNRS, archaeologist. 0000-0002-1243-3167

Nature of the Data

Recommandations:

File formats are described below (question 5).

Exemple de réponse:

- experimental or observational measurements

- maps and plans
- quantitative data
- 3D model
- processed data
- photographs
- inventories
- spatial data
- samples
- physical collections
- software
- educational materials
- bibliography
- GIS
- administrative data
- logical data/database model

Did the production of these data require the reuse of other data?

Recommendations:

If the production of the dataset described here required the reuse of data produced for other purposes:

- indicate which data are concerned
- describe briefly the reused data :
 - the producer of the data
 - the owner of the data
 - the data storage location
 - the legal framework for reuse

If the reused dataset is used as such with no modification, create a dedicated dataset tab.

N.B. : If reusing existing data has been considered but has finally been rejected, explain why you chose not to reuse them for this project.

Exemple de réponse:

The production of these data requires the reuse of public data (CC by licence) produced by Helen Walpole (INRAP). Data storage : INRAP servers / Zenodo repository.

How were the data produced?

Recommendations:

Depending on the nature of the data, specify the method(s) of acquisition; e.g. type of camera, survey, scan, vectorization, etc.

If appropriate, specify :

- the settings,
- the units of measurement,
- the degree of accuracy,
- or any other indicator or information on how the condition of production affected the data (quality, weather conditions, settings...).

For datasets from databases, specify :

- the data model
- the version of the software used.

Exemple de réponse:

Topographic surveys using scale stick with a GPS- **** Leica Viva GS08 GNSS receiver ****

Format of the Data

Exemple de réponse:

- 250 .jpeg files
- 12 .mp4 files
- 1 .csv file
- 5 .PLY (ASCII 1.0) files
- 2 PDF (1.4) files

In this example, the csv file and one of the PDF files contain the same information, i.e. metadata concerning the images and videos in the dataset. For instance, csv files facilitate the reuse of the data while PDF 1.4 (PDF/A) files are intended to facilitate the long-term archiving and accessibility of the data.

Recommendations:

Indicate the native format of the files which make up the dataset described (jpeg, ai, cord, xls, ai; dwg, etc.).

We recommend :

- standard interoperable formats
- open and non proprietary formats

You can also refer to the [Internet Assigned Numbers Authority \(IANA\)](https://www.iana.org/) registry or the Archaeology Data Service (ADS) best practices guide for file formats.

Justify the technical choices, i.e. widespread use in the community, the result of expertise or recommended format for storage in a data repository, etc.

Data Naming

Recommendations:

Please indicate here the naming rules adopted for the data produced during the project.

Take into account the constraints which are specific to the institution which produced it (institutional repositories, regulations, etc.).

You can refer to the ADS service instructions for the permanent archiving of archaeological

data: <https://archaeologydataservice.ac.uk/advice/PreparingDatasets.xhtml#FileNaming>.

They must respect the principles of long-term file accessibility :

- the file name should:
 - be brief: do not exceed 31 characters, do not count the extension (!\ A maximum of 255 characters is allowed for the full name: file path and file name.
 - be accurate: the subject of the document, its type, its date of creation or modification, its version, its identification number in a database that references an image...
 - to avoid any loss of information and to make sure the digital data and documents created during the project are coherent, use unique and persistent identifiers, i.e. DOI, ARK, Handle.... If the format of the file must be changed for archiving, only change its extension (don't change its name). For instance, cyprus_inventory_20210209_v1.xls => cyprus_inventory_20210209_v1.csv
- The file name must comply with the following regulations on characters and practices and use:
 - use the 26 letters of the Latin alphabet,
 - use the 10 Arabic numbers,
 - do not use spaces. You can either use the underscore "_" (Alt + 8), which is preferable to the hyphen (Alt + 6) or you can include capitals at the beginning of certain terms, for instance cyprus_map_20210302.jpg or CyprusMap20210302.jpg
- Avoid the characters and practices below:
 - diacritical marks: à, â, ä, é, ê, ë, ì, í, ò, ô, ù, û, ü, ÿ, ç
 - special characters like ;, : !? % & () # / *
 - words with no actual meaning: the, a, some, and
 - vague designations such as "miscellaneous", "other", "to be classified"

Processing files*

Recommendations:

Indicate if you have renamed the files for use in the framework of the project. If so, specify the naming rules adopted.

They must respect the principles of long-term file accessibility and to do so the file name should:

- be brief: it must not exceed 31 characters, not counting the extension (!\ A maximum of 255 characters is allowed for the full name: file path and file name.
- be accurate: according to the cases concerned, it can or must include the subject of the document, its type, its date of creation or the version,
- be unique to avoid any loss of information and to make sure the digital data and documents created during the project are coherent. If the format of the file must be changed for archiving, the name should not be changed.

In addition, the file name must comply with the following regulations on characters and practices and use:

- the 26 letters of the Latin alphabet,
- The 10 Arabic numbers,
- the underscore "_" (Alt and key 8) is preferable to the hyphen (Alt and key 8) and should replace and indicate spaces,
- include capitals at the beginning of certain terms which means you can avoid " " by staying close to the preceding term.

And avoid the characters and practices below:

- diacritical marks: à, â, ä, é, ê, ë, ì, í, ò, ô, ù, û, ü, ÿ, ç
- special characters like ;, : !? % & () # / *
- words with no actual meaning: the, a, some, and
- vague designations such as "miscellaneous", "other", "to be classified",
- the name of the agent who created or manages the file

● Indicate the format of the files constituting the dataset described here (jpeg, ai, cord, xls, ai; dwg, etc.). Indicate if you have converted source files.

● Indicate here the indexing system(s) (controlled vocabularies, ontology, etc)

Use the repositories of the disciplinary field in question and provide for their enrichment during the project if necessary.

Indicate whether you have generated a new indexing system for using data in the project.

Examples of repositories used in the archaeological community:

- People: [ISNI](#) and [ORCID](#)
- Places: [Geonames](#)
- Subjects of the thesaurus [PACTOLS](#)
- For institutions: [Research Organization Registry](#)

Conditions for reuse*

Recommendations:

Indicate here the reuse conditions : open license, agreement, etc.

N.B. : Data ownership should be described in section 1 of the Data Overview tab.

Associated Metadata

Metadata standard

Recommendations:

Indicate here the metadata standard used to describe the file(s) which make up the dataset.

Ex: [Dublin Core](#), [Inspire](#), [IPTC](#), etc.

Exemple de réponse:

- Dublin Core
- Inspire (Geospatial metadata)

What is the workflow for the production of metadata?

Recommendations:

Were the metadata produced at the time of data creation by the actual data producer or were they produced subsequently? If so, by whom and when?

Which tool(s) were/was used to produce the metadata?

Recommendations:

Indicate which software was used to enter the metadata.

Which indexing systems were used to describe the data?

Recommendations:

- Indicate the indexing system(s) here (controlled vocabularies, ontology, etc.)
- Use the area of knowledge terminology and provide for these to be enriched during the project if necessary.

Ex :

- People: [ORCID](#)
- Places: [Geonames](#)
- Subjects of the thesaurus [PACTOLS](#), [AAT](#)
- Ontology: CIDOC-CRM...

Exemple de réponse:

- People: [ISNI](#) and [ORCID](#)
- Places: [Geonames](#)
- Subjects of the thesaurus [PACTOLS](#)
- Ontology: CIDOC-CRM

Is there documentation associated with the data?

Recommendations:

Specify here actions taken to facilitate long term access to the data (ex : existence of a reference database or inventory, etc.)

Original metadata

Recommendations:

Describe the original metadata (format...).

Added metadata

Recommendations:

If new metadata have been added during the project, specify their format, how they have been created and who is responsible for them (ex : creation, management...).

Data storage

Print documents produced during the project

Recommendations:

Describe the type (s) of document produced exclusively on paper in addition to digital data.

Indicate also the packaging and the place (s) of storage of the data other than digital.

Projected volumetry.

Recommendations:

Indicate here the expected volume of data produced during the project, expressed in terms of storage (GB, MB, TB, etc.). This estimate is useful for planning the storage space required.

Digital data storage infrastructure during the project

Recommendations:

- Describe the infrastructure (equipment, hardware) and location of digital data storage.

- Make sure the storage infrastructure is compatible with your institution's regulations.
- Explain how data will be recovered in the event of an incident.
- You can specify the backup devices (to be distinguished from archival policy and devices)

Exemple de réponse:

- Data stored on the Nanterre University Huma-Num Box made available by the TGIR Huma-Num, servers managed at the IN2P3 lab (CNRS).
- 3D models stored on the Archéogrid Archéovision platform (https://www.archeogrid.fr/home_conservatoire3d), made available by TGIR Huma-Num, servers managed at IN2P3 (CNRS).

Data safety

What are the risks and threats for data?

Exemple de réponse:

- No secure access to the server.
- No automatic backup.

Guarantee of data confidentiality

Recommandations:

Specify the measures taken to ensure:

- the protection of personal data;
- server security.

Guarantee of data integrity and traceability

Recommandations:

Describe here the processes put in place to ensure data integrity and traceability. Certain aspects like quality control and access management are covered by the data management plan and do not need to be repeated here. Describe the backup, duplication and version management procedures.

Access to Data

How should data be read?

Recommandations:

Say if the use of proprietary softwares is required to read the data.

How is access to data guaranteed?

Recommandations:

Indicate whether the data is accessible on a local server, an intranet, on the Internet, in free access or with authenticated access. Specify the different levels of authorization and their rules.

This field must be updated as the project progresses.

Exchanges and sharing

Recommandations:

Explain how the data will be shared and can be found during the course of the project (for example: by deposit in a trusted data warehouse, indexation in a catalog, using a secure data service, direct processing of data requests, etc...)

Protection of sensitive data during the project

Do any of the data require special legal protection and why?

Recommandations:

List the sensitive data and specify here the nature of the risks or constraints, likely to justify access restrictions (eg protection of people, sites, copyright).

Check the compliance of your policy with the general data protection regulations (GDPR).

Are safeguards in place to protect sensitive data?

Recommandations:

List here the measures taken to ensure data protection (for all data, not just sensitive data), security of servers, workstations, the internal network, incident management, etc.

Legal framework

Legal framework for the use, reuse and preservation of data

Recommandations:

Provide here all the relevant legal information explaining the conditions of use, reuse and preservation of data.

- Who owns the data produced and/or reused?

- What rights apply to the databases and their contents?
- How are the intellectual property rights on the data allocated? etc.

For collaborative projects, explain how the ownership of the data is allocated. Be aware that the sharing and dissemination of data must be defined before the project starts. In the case of software development, please remember to refer to specific software legislation.

Archiving policy (created data only)

Retention period of the data produced

Recommendations:

Please refer to your institution's dedicated services.

The archiving policy must comply with the legal framework.

In the case of CNRS research and service units, please note the instruction n°DAF/DPACI/RES/2007/002 of January 15th 2007: [Traitement et conservation des archives des délégations du Centre national de la recherche scientifique \(CNRS\) et des archives des unités de recherche et de service](#) (*Management and conservation of the archives of the delegations of the CNRS and the archives of research and service units*).

Final volume of project data

Recommendations:

Specify the overall volume and also the subset volume (e.g. classified by dataset, producer, site, etc.).

Arrangements for data conservation at the end of the project

Recommendations:

Specify the technical means, processes, conservation sites and human and institutional resources dedicated to data conservation.

Terms and conditions for access to data at the end of the project

Recommendations:

Specify how it will be possible to access the data.

Dissemination of data

Post-project data dissemination policy

Recommendations:

Describe here the dissemination policy which apply to the different data or datasets produced during the project particularly in relation to the target communities. Indicate where the data are deposited and referenced at the end of the project.

The French National Research Agency (ANR) recommends "explaining how data should be retrieved and shared (e.g., by deposit in a trusted data repository, indexing in a catalogue, use of a secure data service, direct processing of data requests, or use of any other mechanism).

Exemple de réponse:

Dissemination to the community of French archaeologists:

- Integration of new concepts in the PACTOLS thesauri, in connection with the Frantiq CNRS GDS Unit.
- Integration of XX and YY datasets in the Open Archo platform: <https://masa.hypotheses.org/openarchaeo> (triplestore dedicated to archeology developed within the framework of the MASA consortium: its intuitive query is inspired by the British Museum's Research Space search engine)

Dissemination to the international research community:

Publication of articles with a link to the data concerned, data papers. Deposits in the open archive HAL (Hyper Articles en Ligne)

- Deposit of datasets in Zenodo

Dissemination to heritage curators and cultural mediators:

- Participation in the Annual International Conference on Cultural Mediation and Rock Art

Have there been any publications resulting from the use of these datasets?

Recommendations:

- Monograph, scientific article, conference papers, data paper, etc.
- Specify at least: author, year of publication, title, publisher and place of publication. Also specify how to identify and access the publication and its associated data.
- If a public Zotero collection is associated with the project, indicate this.

The French National Research Agency (ANR) recommends "explaining when the data will be made available", "indicating the expected publication deadlines", "explaining whether there is a claim of exclusive use of the data and, if so, for what reason and for how long".

Exemple de réponse:

- Claridge, A. and Rendell, H. 2013 The Evolution of Rome's Maritime Façade: archaeology and geomorphology at Castelporziano (Data Paper), *Internet Archaeology* 35. <https://doi.org/10.11141/ia.35.11>
- Jean-Michel Chazine, Jean-Georges Ferrie. Recent archaeological discoveries in East Kalimantan, Indonesia

Is certain data covered by an embargo?

Recommendations:

Embargo = Temporary protection allowing the postponement of free distribution and/or reuse.

Specify which data, for how long, for what reason.

The French National Research Agency (ANR) recommends "indicating whether data sharing will be deferred or restricted, for example for publication, intellectual property protection, or patenting".

Potential for reuse of the data (target user groups, resources required, etc.).

Recommendations:

Specify:

- the type of user groups or potential areas of application
- which data elements can be correctly reused: associated documentation, contextualization, data qualification (degrees of accuracy, certainty, etc.), DOIs, etc.

The ANR recommends "indicating who will be able to use the data. If it is necessary to restrict access for certain communities or to impose a data sharing agreement, explain how and why. Explain the steps that will be taken to overcome or minimize these restrictions."

If necessary, you may refer back to other sections of the DMP that have already been filled in.

Exemple de réponse:

Dataset relating to the iconography of rock paintings (see the section on license types).

File documentation_paintings.txt associated with the dataset.

Types of potential user groups:

- heritage curators (museums)
- cultural mediators
- prehistory doctoral students
- (pre)history teachers

Conditions for reuse

Recommendations:

1. Specify the licenses that apply to the project data set as a whole or to each data set and explain how the datasets can be reused.

- **Open licenses are often recommended :**
 - [Etalab Licence](#)
 - [Creative Commons Licences](#)
 - [GNU Licences](#)
- French law recommends using the open license, Etalab: <https://www.legifrance.gouv.fr/eli/decret/2017/4/27/2017-638/jo/texte>

2. Check the existence of elements which are required to cite the data and give an example of how to cite the data you have produced. Datacite recommends checking the existence of at least these 5 elements: Author (Year of publication): Title. Publisher. Identifier / Creator (Year of Publication): Title. Publisher. Identify

3. If specific software and technical and/or scientific skills are required to reuse the data or data sets, provide the relevant information here.

Exemple de réponse:

Open licence (Etalab). Cite the data in this form - Author (Year of Publication): Dataset title. Publisher. Dataset identifier (DOI)