Contract no. **[insert GA number]**

**[insert PROJECT ACRONYM]**

**[insert project title]**

|  |  |
| --- | --- |
| **Deliverable No.** |  |
| **Deliverable Full title** |  |
| **Work Package No. and Title** |  |
| **Lead beneficiary (extended name and acronym)** |  |
| **Authors**  **(Acronyms of beneficiaries contributing to the deliverable)** |  |
| **Planned delivery date** | **dd/mm/yyyy (M?)** |
| **Actual delivery date** | **dd/mm/yyyy (M?)** |
| **Dissemination level**  **(PU = Public; PP = Restricted to other program participants; RE = Restricted to a group specified by the consortium; CO = Confidential, only for members of the consortium)** |  |
| **Project website** |  |
| **Project start date and duration** | **Start date of project:**  **Duration: ? months** |

*The information in this document reflects only the author’s views. The [insert name of the agency] of the European Commission is not responsible for any use that may be made of the information it contains. The user therefore uses the information at its sole risk and liability.*

**Document History**

|  |  |  |  |
| --- | --- | --- | --- |
| **Version** | **Date (DD/MM/YYYY)** | **Created/Amended by** | **Changes** |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**Scheduled Data Management Plan (DMP) Updates**

The DMP is a document that evolves during the lifespan of the project and registers all relevant changes in the life-cycle of all the research data sets of [insert project acronym] project. Updated versions of the DMP have already been planned (see table below). Moreover, this document will be updated whenever important changes in the data or the data management policy occur.

|  |  |
| --- | --- |
| **Issue** | **Expected by project month (M)** |
| Initial DMP |  |
| Intermediate DMP |  |
| Final DMP |  |

**Partner Acronyms**

|  |  |
| --- | --- |
| **Partner extended name** (country) | **acronym** |
|  |  |
|  |  |
|  |  |
|  |  |

**Content**

[The Data Management Plan (DMP) 1](#_Toc534806354)

[1. Data Summary 1](#_Toc534806355)

[2. FAIR Data 2](#_Toc534806356)

[2.1 Making data findable, including provisions for metadata 2](#_Toc534806357)

[2.2 Making data openly accessible 3](#_Toc534806358)

[2.3 Making data interoperable 4](#_Toc534806359)

[2.4 Increase data re-use (licensing) 4](#_Toc534806360)

[2.5 Allocation of resources 5](#_Toc534806361)

[2.6 Data security 6](#_Toc534806362)

[2.7 Ethical aspects 6](#_Toc534806363)

[2.8 Other issues 6](#_Toc534806364)

[3. Data sets overview 7](#_Toc534806365)

[Annex I: Data sets tables 8](#_Toc534806366)

[WP? – [insert WP title] 8](#_Toc534806367)

[Annex II: “README” file template 9](#_Toc534806368)

[Annex III: Open Access status of project publications 11](#_Toc534806369)

# The Data Management Plan (DMP)

The DMP is a document that provides details regarding all the research data collected and generated within the [insert project acronym] project. In particular, it explains the way research data are handled, organized, licensed and made openly available to the public, and how they will be preserved after the project is completed. The DMP also provides motivations when versions or parts of the project research data cannot be openly shared on account of third-party copyright issues, confidentiality or personal data protection requirements or when open dissemination could jeopardize the project achievements.

This DMP reflects the current state of the art of the [insert project acronym] project. However, the details and the final number of the project data sets may vary during the course of research. The variations will be recorded in updated versions of this DMP.

# Data Summary

…

The project will collect and generate several type of data:

Research teams have agreed to convert research data from proprietary formats to well-known and documented open formats in order to facilitate accessibility and reusability (Tab.1).

***Table 1 - Summary of data format***

| **Type of data** | **Formats used**  **during data processing** | **Formats for**  **sharing reuse and preservation** |
| --- | --- | --- |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

Documentation files explaining all relevant details regarding data collection, processing methodologies and quality assurance are deposited along with the data sets in .odt, .rtf or .pdf format.

The existing data reused in the course of the project were from different sources: …

The data produced can be of interest to different potential users. They may include:

* …

The expected size of the data is ….

# FAIR Data

This DMP follows the EU guidelines[[1]](#footnote-1) and describes the data management procedures according to the FAIR principles[[2]](#footnote-2). The acronym FAIR identifies the main features that the project research data must have in order be findable, accessible, interoperable and re-useable, allowing thus for maximum knowledge circulation and return of investment.

## Making data findable, including provisions for metadata

…

[PROJECT ACRONYM] research data are organized in data sets, which are named collections of data units with the same focus and scope. In this DMP are suggested the following common rules for data set naming in order to improve data visibility, discoverability, citation and permanent online tracking. The recommended data set title structure consists of:

*PROJECT ACRONYM: WPnumber: WP title or description specifying WP aims: Tasknumber: Task title or description specifying Task aims: additional information specifying coverage and nature of data (if necessary): version number (optional, in case of revisions to help identifying the updates especially in repositories that do not track versioning automatically)*

Example:

*[INSERT EXAMPLE]*

The version number of the data set will be added at the end of the title in case of data revisions to help identifying the data set updates especially in repositories that do not track versioning automatically (see *Annex I* for data set names, unique identifiers and descriptions).

The DMP recommends also the following rules for file naming:

* for data set file(s)

*[PROJECT ACRONYM]\_WPnumber\_Tnumber\_coverage or other content specifications\_date (YYYYMMDD)\_vn.file extention*

Example:

*[INSERT EXAMPLE]*

* for readme file(s)[[3]](#footnote-3)

*README\_[PROJECT ACRONYM]\_WPnumber\_Tnumber\_coverage or other content specifications\_date (YYYYMMDD)\_vn.file extention*

Example:

*[INSERT EXAMPLE]*

WPnumber means “work package number” Tnumber is “task number”, and vn is the “version number” (in case of data revisions or updates).

## Making data openly accessible

As a guiding principle, [insert project acronym] seeks to make research data openly available, whenever possible, in order to allow dissemination, validation and re-use of research results.

Restrictions to access are applied only in the following cases:

* when …

As a consequence, all possible and legitimate actions and strategies are adopted to allow data sharing including:

* obtaining copyright permissions from third party data owners to be allowed to re-use, reproduce and distribute the collected data;
* converting the files to standard open formats;
* providing all relevant documentation and explanation for the data and the data sets;
* …

For data that fall under some of the restrictions described above and for which it is not possible to take any action to make them shareable, EU allows complete closure or restricted access to them. [insert project acronym] DMP indicates the versions or parts of the data sets that cannot be freely shared providing the specific motivations in *Annex I*.

At the time of publication of results, researchers deposit the project data that can be shared in a data repository in order to guarantee their discoverability, access and preservation beyond the project end.

The chosen data repositories… .

Each different data set is deposited by… .

***Table 2 – Summary of repositories.***

***The following table shows the repositories for data sets publication and preservation chosen by ...***

| **Partner** | **Repository name** | **URL** | **Type** |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

***Table 3 – Features of the chosen repositories***

|  |  |  |  |
| --- | --- | --- | --- |
| **Repository name** | **Type** | **Permanent ID** | **OpenAIRE compatibility?** |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

For each deposited data set, all relevant documentation explaining data collection procedures and analysis (such as codebooks, methodologies, etc.) will be made available along with the data, in order to guarantee intelligibility, reproducibility and the validation of the project findings. Moreover, the deposited documentation specifies the tools and software recommended to reproduce and reuse the data, when necessary. (See Tab.4 for examples of tools and software enabling reuse of the dataset).

…

## Making data interoperable

To allow data exchange and re-use among researchers, institutions, organisations, countries, etc., partners will convert all shareable data from proprietary formats and will made them available in well-known and documented open formats (see Tab.1 for details), as much as possible compliant with available (open) software applications (see Tab.4 for details).

***Table 4 – Summary of tools and software for enabling re-use of the data sets***

| **Tools/software** |
| --- |
|  |
|  |
|  |
|  |
|  |

…

## Increase data re-use (through clarifying licences)

[project acronym] distributes the shareable data by adopting licenses that allow re-use of the data and of the data sets in their entirety by other scholars and stakeholders. The data sets are made available, unless otherwise stated, under … .

In general, data are made openly available as underlying data necessary to validate the research results immediately at the time of publication of public reports and scientific papers. Data are given full citation from official project publications and web site and they are made available through institutional or public data repositories compliant with OpenAIRE requirements[[4]](#footnote-4). (See Tab.2)

It is possible that an embargo period may be applied to some data sets to allow full exploitation of research results by the partners. Embargoes applied to the datasets are specified in the descriptive tables.

The research data that are made openly available are deposited in open formats in … repositories that guarantee long term preservation to archived items, therefore they will be re-usable by third parties after the end of the project.

The research data that cannot be shared because … .

Back-up copies of the research data that cannot be shared and cannot be deposited in institutional or public data repositories will be … .

The quality of the experimental data … .

## Allocation of resources

Making data FAIR requires an investment of money and researchers’ time. In [project acronym] case, costs for … are … .

Costs related to [e.g. preparation of the data management plan and the datasets descriptive documentation, of the conversion of data files to open formats and data sets self-archiving procedures]… are … .

The cost for long term preservation of not shareable data after the end of the project … .

Responsible for the data management are …

Researchers are encouraged to identify themselves with the unique persistent identifier ORCID[[5]](#footnote-5). Registration allows for automated linkages between the researcher’s identity and his research activities and outputs.

***Table 5 – Summary and contacts of the research team leaders.***

| **Team** | **Leader** | **ORCID ID (if available)** | **mail** |
| --- | --- | --- | --- |
|  | Family name, given name |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

Moreover, partners are encouraged to identify and cite all contributors (See Tab.6) participating in data management activities, specifying their roles according to a given standard vocabulary (DataCite Metadata Schema).

***Table 6 – Summary of team members involved in the data sets collection and management.***

| **Team** | **Member** | **ORCID ID (if available)** | **Role** |
| --- | --- | --- | --- |
|  | Family name, given name |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

***Keys for “Role” column: Data Collector (such as survey conductors, interviewers…), Producer (person responsible for the form of a media product), Project Member (a researcher indicated in the GA), Researcher (an assistant to one of the authors who helped with research, data collection, processing and analysis but is not part of team indicated in the GA), Research Group (the name of a research institution or group that contributed to the data set).***

(See *Annex I* for details about data management responsibilities related to each project data set).

## Data security

… .

## Ethical aspects

… .

## Other issues

… .

# Data sets overview

The following table (Tab.7) offers an overview of the data sets expected from the project which are described more in detail in *Annex I*. It will be updated according to DMP changes and variations.

***Table 7 – Data sets list.***

***Table acronyms and abbreviations: n°= data set progressive number, LB = WP lead beneficiary, PP = project phase (starting month-ending month), CT = creator team in charge of curating the data set, C=collected, G=generated, A=available, IP=in progress, NYA=not yet available.***

| **n°** | **WP** | **LB** | **TASK or SUBTASK** | **PP** | **CT** | **DATA SET (tentative title)** | **SOURCE** | **STATUS** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |

# Annex I: Data sets tables

The analytic descriptions of the expected data sets of [project acronym] project are reported in this Annex organized by work-packages.

# WP? – [insert WP title]

WP? [insert brief description]… .

Lead: **[insert acronym]**

Participants: **[insert acronyms]**

Months: **?-?**

Potential users for the data sets of this WP include … .

| **num** | **status** | ***Dataset title*** |
| --- | --- | --- |
| **ID [ID type]** | |  |
| **Version** | |  |
| **Team in charge** | |  |
| **Creator/s** | | Family name, given name [TEAM]; … |
| **Contributor/s** | | Family name, given name [TEAM]; … |
| **Contact Person/s** | | Family name, given name [TEAM, email]; … |
| **Contents** | |  |
| **Data format** | |  |
| **Data volume** | |  |
| **Accessibility** | |  |
| **Related publication/s** | |  |

# Annex II: “README” file template

**README file**

Data Set Title: **“[insert title as defined in the DMP]”**

Data Set Author/s: **Name Surname** (Affiliation), ORCID (if available);

**[Add one or more creators, if present]**

Data Set Contributor/s: **Name Surname** (Affiliation),ORCID (if available);

**[Add one or more contributors, if present. Otherwise, cancel this line]**

Data Set Contact Person/s: **Name Surname** (Affiliation), ORCID (if available), email;

**[Add one or more contact person]**

Data Set License: this data set is distributed under a **(INSERT LICENSE)**

**[Insert the chosen license as indicated in the DMP: e.g. “this data set is distributed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license,** [**https://creativecommons.org/licenses/by/4.0/**](https://creativecommons.org/licenses/by/4.0/)**”]**

Publication Year: **(insert YEAR)**

Project Info: **[insert PROJECT ACRONYM] ([project full title],** funded by European Union, Horizon 2020 Programme. Grant Agreement num. **[insert grant agreement number]**; **[insert project website url]**

**Data set Contents**

The data set consists of:

**[Indicate the files that compose the dataset and their name and format.**

**WE STRONGLY SUGGEST YOU TO FOLLOW THE EXAMPLES PROVIDED FOR THE FILE NAMING, MATCHING THE DATASET FILENAME WITH THE README ONE**

**In the following examples the data sets were composed by only one file. In case the dataset consists of more files you can name them as described and put them in a compressed folder. In this case readme file name should match the compressed folder name]**

EXAMPLE1

* 1 textual qualitative file saved in .rtf format

“**ProjectAcronym\_WP3\_T3-2\_ItalyInterviews\_20161221\_v01.rtf**”

**[structure of the filename “ProjectAcronym\_insert WP number\_insert Task number, e.g. T3.2\_ insert Content Describing Keywords\_insert date YYYYMMDD\_insert version, if needed.format”**

**Suggested format:**

**-for textual qualitative data .rtf or .txt**

**-for tabular quantitative and qualitative data .csv**

**avoid proprietary formats such as .doc/.docx and .xls/.xlsx]**

* 1 README file

“**README\_ProjectAcronym\_WP3\_T3-2 \_ItalyInterviews\_20161221\_v01.rtf”**

**[Same naming as the dataset file. Preferred format .rtf/.txt, allowed format .pdf]**

EXAMPLE2

* 1 tabular quantitative file saved in .csv format

“**ProjectAcronym\_WP7\_T7.3\_Questionnaire\_Sweden\_20170905.csv**”

* 1 README file

“**README\_ProjectAcronym\_WP7\_T7-3\_Questionnaire\_Sweden\_20170905.rtf”**

**Data set Documentation**

*Abstract*

….

**[Insert a brief abstract describing the content of the dataset]**

Content of the files:

* file **[Insert filename]** contains ...

**[Provide a brief description of the content of the file/s. This is an example of how you could start]**

* file **[Insert filename]** contains ...
* …

*File specifics*

…

**[Provide useful info regarding file conversion etc… (Optional)**

**Please indicate instruction/technical info in order to allow potential users to correctly visualize and reuse your data (e.g. specific software, …).**

**In case of data converted in open formats it could be useful to provide some further information. For example if you deposit for long term preservation a .csv file derived from an excel you can describe the conversion. Here is an example of description of conversion using libre office calc software:**

***To create the .csv files, “LibreOffice Calc” version: 5.1.4.2 (portable) was used, with the following specifics:***

***•Character set Europa occidentale (Windows-1252/WinLatin1)***

***•Field delimiter « , » (comma)***

***•Text delimiter « “ » (quotes)*]**

*Notes*

…

**[Related to the whole dataset or to single files of a multi-file dataset (Optional)]**

*Data sources*

…

**[Optional]**

*Methodologies*

…

**[If necessary to understand how to reuse data]**

*Codebook of variables*

…

**[If necessary to understand the meaning of the variables]**

**Instructions, examples and footnotes in should be deleted from final version**

# Annex III: Open Access status of project publications

In the following table (Tab.8) it is reported the updated list describing the open access status of the project publications and the underlying data sets.

***Table 8 – Open access status of [project acronym] publications and data sets.***

|  |  |
| --- | --- |
| **Publications** | |
| **Bibliographic citation of the publication** | |
| **Archived in repository for Open Access?** | Y/N, in case of “Yes” indicate the repository and link to the deposited copy |
| **Status** | Open Access/embargo/closed access, indexed in OpenAIRE or not |
| **Related dataset/s** | Data set num: cite dataset |
| **Bibliographic citation of the publication** | |
| **Archived in repository for Open Access?** | Y/N, in case of “Yes” indicate the repository and link to the deposited copy |
| **Status** | Open Access/embargo/closed access, indexed in OpenAIRE or not |
| **Related dataset/s** | Data set num: cite dataset |
|  | |
| **Archived in repository for Open Access?** |  |
| **Status** |  |
| **Related dataset/s** |  |
|  | |
| **Archived in repository for Open Access?** |  |
| **Status** |  |
| **Related dataset/s** |  |
|  |  |

1. Guidelines on FAIR Data Management in Horizon 2020 (Version 3.0, 26 July 2016), <http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf> [↑](#footnote-ref-1)
2. The FAIR data principles (Force11 discussion forum), <https://www.force11.org/group/fairgroup/fairprinciples> [↑](#footnote-ref-2)
3. A “README” file is a document containing relevant information about data set authorship, terms of reuse and responsibilities, explaining data set content and structure, collection procedures and analysis (such as file specifics, methodologies, codebooks of variables, data sources, and further necessary notes). (See Annex II to visualize the suggested README file template). [↑](#footnote-ref-3)
4. OpenAIRE, *For Data Providers* <https://www.openaire.eu/intro-data-providers> [↑](#footnote-ref-4)
5. <http://orcid.org/> [↑](#footnote-ref-5)