EUROPEAN COMMISSION

RECETAS
TESTING ACTIONS FOR SOCIAL PRESCRIBING IN NATURAL SPACES

Horizon 2020

Grant: 945095

WP4: Evaluate NBSP through intervention studies Data Management Plan (D4.2)

[FSIE]

[Version 1] - [August 31, 2021]

www.recetasproject.eu

Document Information

Project Number	H2020 – 945095	Acronym	RECETAS
Full title	Re-Imagining Environments for Connection and Engagement: Testing Actions for Social Prescribing in Natural Spaces		
Project URL	www.recetasproject.eu		
EC Project officer	Laia Quiros		

Deliverable Number	D4.2	Title	Data Management Plan
Work package Number	WP4	Title	Evaluate NBSP through intervention studies

Date of delivery	Contractual		Actual	
Nature	Prototype ☐ Report x Dissemination ☐ Other ☐			
Dissemination Level	Public x Consortium ☐			

Authors	Name	Partner
Main Author(s)	Marta Roqué	FSIE
Contributor (s)	Laura Coll Jill Litt, Laura Hidalgo, Ashby Sachs Albert Rojas, Gabriela Garcia Lucie Cattaneo Kaisu Pitkälä Michal Karľa	FSIE ISGlobal UC AP-HM UH CU

Abstract (for dissemination)	Data Management Plan
Key words	

Version Log			
Issue Date	Version	Author	Change

INDEX

Introduction

This document is the framework for the data management of the RECETAS project.

RECETAS project

General project information

Project Title	*Re-imagining Environments for Connection and Engagement: Testing Actions for Social Prescribing in Natural Spaces*
Project description	*RECETAS explores loneliness through a transdisciplinary lens, integrating social, behavioral, health, and natural sciences, and is grounded in participatory principles. It will use randomized controlled trials (RCT) and other epidemiologic, anthropological, and health economic methods to test socially- and culturally-innovative nature-based social prescribing (NBSP) in six cities in Europe, Latin America, and Australia.*
Begin and End of the project	*01/03/2021 - 28/02/2026*
Project coordinator	*Dr. Jill Litt, ISGlobal*
Other Project Investigators (work package leaders)	*Danielle Varda (VNL, WP2), Gabriela Garcia (UC, WP3), Laura Coll Plana (FSIE, WP4), Ursula Rochau and Uwe Siebert (UMIT, WP5), Alice Cucchiaro (AViTEM, WP6), Paola Garbarino (Aketh Investments Network Services) (WP7).*
Project Partners	*FUNDACIO PRIVADA INSTITUT DE SALUT GLOBAL BARCELONA (ISGLOBAL); FUNDACIO SALUT I ENVELLIMENT (FSIE); DEPARTAMENT DE SALUT - GENERALITAT DE CATALUNYA (ASPCAT); GIP AGENCE FRANCAISE POUR DES VILLES ET TERRITOIRES MEDITERRANEENS DURABLES (AViTeM); UNIVERSITY OF THE WEST OF ENGLAND, BRISTOL (UWE); ROYAL MELBOURNE INSTITUTE OF TECHNOLOGY*RMIT UNIVERSITY (RMIT); UNIVERZITA KARLOVA (CU); UNIVERSIDAD DE CUENCA (UC); UMIT-PRIVATE UNIVERSITAT FUR GESUNDHEITSWISSENSCHAFTEN, MEDIZINISCHEINFORMATIK UND TECHNIK GMBH (UMIT);*

	HELSINGIN YLIOPISTO (UH).
Funding Source / Program	*Funded by the European Union (EU), under Grant 945095 in Program Horizon2020*

Data Management Plan for RECETAS

HORIZON 2020 projects are required to deliver a Data Management Plan (DMP) within the first six months of their project, as part of making research data findable, accessible, interoperable and re-usable (FAIR). The DMP that constitutes Deliverable D4.2 will lay out the framework and guidelines that apply specifically to the data collected, observed, generated, or created **within** Work Package (**WP) 4.** It will also establish the guidelines to an open research data policy where, as stated in the RECETAS Grant Agreement, it should be made available to third parties any information that has been collected, observed, generated, or created to validate original research findings and scientific publications within the project, as well as any associated metadata, and information about tools and instruments necessary for validating the results. It is understood that data may constitute commercial assets and thus will be reviewed within the policies of the Consortium Agreement.

A DMP is a document that describes the scale and the format(s) of the data generated, collected, or reused during the life of the project and outlines how they will be handled and shared during the project and in the longer-term. DMP are updated during the study timeline, as necessary study milestones are achieved, and key information is provided by the Consortium. Key milestones that will inform this DMP are:

- The development and report of the feasibility study of NBSP interventions in each partner city in WP3 (Deliverable 3.3),
- the writing and approval of the study protocol for the studies in WP4 (Deliverable 1.1 and 8.1),
- the writing of the informed consent forms and information sheets for participants (Deliverable 1.1 and 8.1)

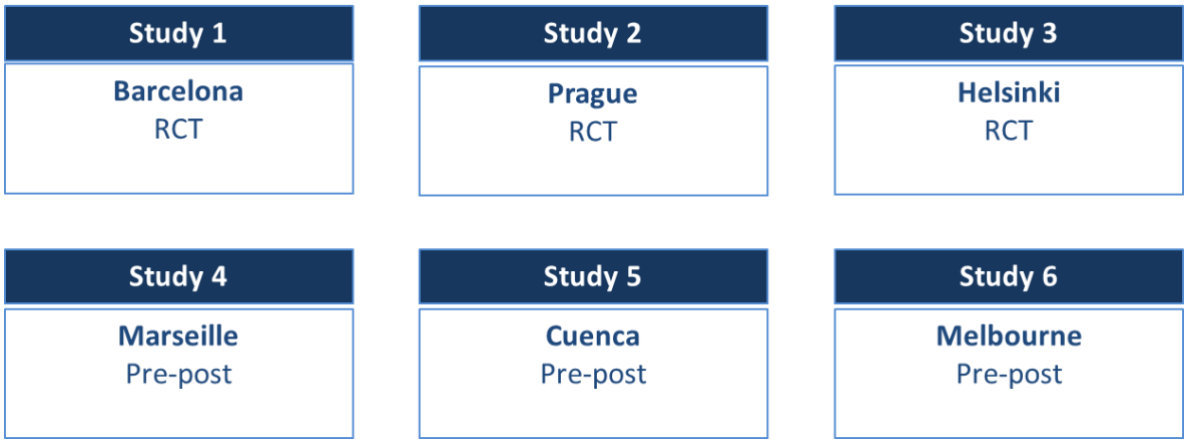
This DMP follows the Horizon2020 FAIR DMP template (See Annex), and its contents are organized as:

1. Description of data that will be collected, processed and/or generated
2. Methodology and standards applied (data types, formats, capture methods)
3. Making data FAIR (findable, openly accessible, interoperable & re-usable) during and after the end of the project
4. Allocation of resources
5. Data security aspects of how data will be curated, backed up and preserved (both during and beyond the life of the project)
6. Ethics and Intellectual Property

1. Data summary

Responsible partner	FsiE
What is the purpose of the data collection/generation and its relation to the objectives of the project?	The "Feasibility, impact and process evaluation" data collection is generated as Task 3.4 "*Test the feasibility of NBSP interventions in each partner city*" in WP3 "Co-Create NBSP Menu and Test Intervention Feasibility", led by UC, and tasks 4.3 "*Subject Assessments*" and 4.4 "*Process Evaluation and Qualitative Analysis*" in WP4 "Evaluate NBSP through intervention studies", led by FSiE. The participating partners providing data into the datasets are ISGLOBAL, UH, CU, RMIT, UC, AP-HM.
What types and formats of data will the project generate/collect?	Quantitative and qualitative data
Will you re-use any existing data and how?	No
What is the origin of the data? Methodologies for data collection	Process evaluation research applying qualitative research methodologies. Intervention studies (feasibility studies, randomised controlled trials and pre-post studies) applying quantitative methodologies.
What is the expected size of the data?	TBC
To whom might it be useful ('data utility')?	RECETAS researchers in WP1, WP5, WP6, WP7 and WP8, practitioners, third-party organizations to inform policy and practice

The "Feasibility, impact and process evaluation" data collection refers to the datasets derived from the six intervention studies (three randomised controlled trials and three pre-post intervention studies), each conducted in a different city in Europe, Latin America or Australia, and the feasibility studies that precede them. The intervention studies will test the same intervention, tailored to their community context. They will follow a common methodological framework, with individual protocols tailored to each study, and follow the same data management general guidelines.



Task 3.4 will test the feasibility of NBSP interventions in each partner city in terms of capability of implementation during the project calendar and capability of recruiting enough participants to have enough statistical power to test the impact of the intervention, based on the pre and post intervention evaluation of the indicators defined in Task 3.3.

Task 4.3 will organize and supervise assessments of impact at baseline and three follow-up time points (post-intervention, 6-month follow-up, and 12-month follow-up) of the trial participants in each study, collecting the variables set in the study protocol.

Task 4.4 involves a continuous process evaluation which will be conducted during the study, documenting the implementation of the intervention, describing, and comparing processes both in the intervention and control groups, collecting information from study participants and professionals about their perceived impacts, and the experience of taking part in the trial to explore the mechanisms of impact, and collecting data on individual and cultural contexts. Qualitative procedures including in-depth interviews, focus groups and participant observation will be used to explore the experiences and views from a purposeful sample of participants, facilitators and other relevant actors involved.

2. Methodology and standards applied

A variety of data types will be collected as part of the data assessments: interview self-reported assessments, expert assessments, qualitative data obtained through interview of focus groups. The specific data and formats will be further detailed in future versions of the Data Management Plan, once the intervention study protocols are approved.

Data will be collected in electronic or paper format in each site. Data will be directly collected or later entered into a platform verifying GPDR requirements (e.g. REDCap). This platform ensures secure remote access, restricted to authorised identified individuals. The platform is installed in a local environment at ISGLOBAL and other RECETAS consortium partners, which are member of the REDCap Consortium. Appropriate data safety and backup measures are implemented in each site, as described in Ethics Deliverable D8.2.

Data quality procedures will be established alongside the project (e.g. data completeness and consistency checks). File-naming and versioning conventions will be applied by all partners. Transfer of any participant data outside the REDCap platform should be avoided. If unavoidable, secure transfer will be conducted by encrypting the data into password-accessed files, and transferring separately the encrypted files and the password information.

3. Procedures to ensure FAIR data

3.1. Making data findable, including provisions for metadata

Are the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers)?	Yes
What naming conventions do you follow?	See text
Will search keywords be provided that optimize possibilities for re-use?	Yes
Do you provide clear version numbers?	Yes
What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.	See text

The research data produced and/or used in the project will be discoverable through metadata and Digital Object Identifiers (DOI), once the study ends. Research datasets will be stored in a repository (e.g. ZENODO), and DOI identifiers assigned to each data collection. Also, each dataset will be assigned metadata using the Dublin Core Schema (or equivalent metadata standard), including keywords following established thesaurus (e.g. MESH terms).

To ensure that the data is easy to find after the study ends, some guidelines will be established from the start regarding naming and versioning conventions of files and folders, which will simplify the data management processes during the study.

Naming and versioning conventions: Files will be named with the following structure:

- Project name (RECETAS)
- Description of the content (<18 characters)
- Version number and/or Date of creation (YYYYMMDD), as appropriate

Examples of file naming:

"RECETAS_Basal_descript_v1.0.do" or "RECETAS_BCN_Basaldata _20210525.dat"

Remember:

- Avoid special characters or spaces in a file name
- Use capitals and underscores instead of periods or spaces or slashes
- When versioning files, use consecutive numbering for major version changes, with decimals used for minor changes (v1; v1.1; v2.1; v2.2).
- 'Read me' documents with detailed info may be needed to accompany files to clarify contents, versions, etc.

Metadata: After the datasets are closed and kept in a suitable repository, specific metadata will be added using the metadata standard the Dublin Core Schema. See below the list of substantive elements that will conform the metadata for each dataset.

Example of Dublin Core Schema metadata elements to be applied to any dataset in RECETAS

Dataset name and identifier: xxx			
Substantive elements	**Label**	**Metadata**	**Notes**
Title	*DC.Title*	xxx	xxx
Keywords	*DC.Subject*	xxx	xxx
Description	*DC.Description*	xxx	Include name of the action, acronym and grant number, as well as the terms "European Union (EU)" and "Horizon 2020"
Language	*DC.Language*	xxx	xxx
Relation	*DC.Relation*	xxx	xxx
Editor	*DC.Publisher*	xxx	xxx
Date of creation	*DC.Created*	xxx	Also, length of embargo period if applicable
Type of resource	*DC.Type*	xxx	xxx
Format	*DC.Format*	xxx	xxx
Resource identifier	*DC.Identifier*	xxx	Persistent identifier, preferably DOI

3.2. Making data openly accessible

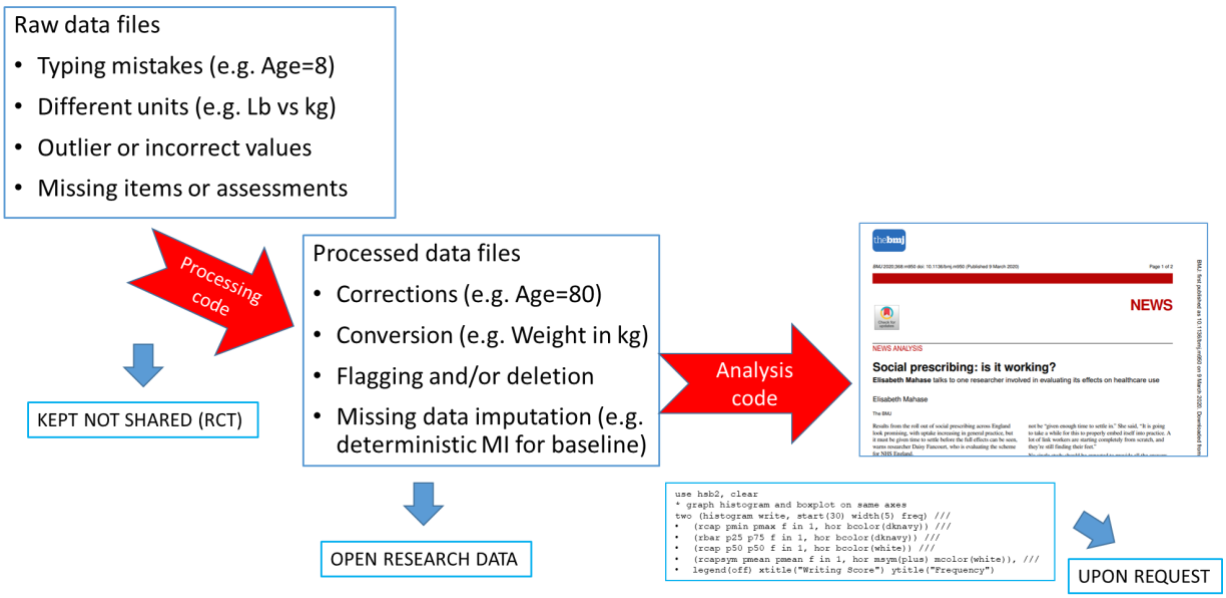
Which data produced and/or used in the project will be made openly available as the default? If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.	See text
How will the data be made accessible (e.g. by deposition in a repository)?	See text
What methods or software tools are needed to access the data?	TBD
Is documentation about the software needed to access the data included?	TBD
Is it possible to include the relevant software (e.g. in open source code)?	TBD
Where will the data and associated metadata, documentation and code be deposited? Preference should be given to certified repositories which support open access where possible.	Yes
Have you explored appropriate arrangements with the identified repository?	TBD
If there are restrictions on use, how will access be provided?	TBD
Is there a need for a data access committee?	TBD
Are there well described conditions for access (i.e. a machine readable license)?	TBD
How will the identity of the person accessing the data be ascertained?	TBD

Research information that may be **shared upon request**:

- The processed datafiles that support scientific publications and may be needed to validate results;
- related documentation (e.g., metadata, data dictionaries, ontologies) required to access the data; and
- The analysis code to reproduce and validate the results in the trials analyses and any derived manuscripts (e.g., commands to generate graphs, statistical analysis syntax).

Research data that are **not expected to be shared**:

- The original anonymized raw data collected from the randomized clinical trials (e.g., digital copies of the case report forms, or derived datasets), plus processing code (script to clear or filter data and generate the clean definitive datafiles to be analysed);
- any non-anonimizable digital research data (e.g., no-anonymized or semi-anonymized audio recordings of focus groups).



As stated in the RECETAS Grant Agreement, the project digital research data will be OPEN within the guidelines established in this Data Management Plan and other project documents. After the study end, all open research data in RECETAS will be deposited in a suitable certified long-term repository such as ZENODO, and open access will be established to identified users to access and use that data. ZENODO is a secure catch-all repository for EC funded research hosted at CERN under the framework of the OpenAIRE project. The partners are not aware of any existing standard repository for the discipline.

However, some datasets may present difficulties to be shared (e.g., qualitative datasets that are linked to semi-anonymized audio recordings of focus groups), or require an embargo period (e.g. to allow the production of manuscripts to report the project results to scientific journals). These

difficulties and restrictions will be listed for the specific datasets, alongside their causes and justification, in future updates of this Data Management Plan. All datasets, regardless of being open data or not, will be kept for a minimum period of 5 years, or a longer time following regulation in participating countries.

As the study evolves, a detailed list of datasets will be generated and added to this Data Management Plan. The dataset associated metadata will specify the software tools to access the data in each dataset. As far as possible, the datasets to be stored long term will be in non-proprietary formats and accessible with open software.

Some aspects of data accessibility that will be defined in future versions of the Data Management Plan refer to potential embargo periods, access restrictions to the datasets, need for a data access committee in RECETAS, and specific issues related to the repository arrangements and access provision.

As stated in section 6 Ethics aspects and Intellectual Property, informed consent will let research participants know about plans for preserving and sharing their personal data after the study end.

3.3. Making data interoperable

Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)?	Yes
What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?	See text
Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?	Yes
In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?	Yes

The RECETAS Consortium will strive to make the research data produced in the project interoperable and re-combinable with other datasets. The datasets to be kept long term will be presented, if possible, in formats compatible with open software applications, and have standard formats to allow recombination with other datasets in the field. To this end, the data will be accompanied with descriptive metadata and data dictionaries. A dictionary will be created for each digital dataset, describing the variables format and contents, and assigning labels to the variable values. As much as possible, standard vocabularies will be set for data types, and failing to do so, mappings to existing ontologies will be facilitated in the dataset documentation.

3.4. Increase data re-use (through clarifying licences)

How will the data be licensed to permit the widest re-use possible?	TBD
When will the data be made available for re-use? If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.	TBD
Are the data produced and/or used in the project useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why.	TBD
How long is it intended that the data remains re-usable?	TBD
Are data quality assurance processes described?	TBD

Issues related to data licensing and usability will be further detailed in future updates of the Data Management Plan.

4. Allocation of resources

What are the costs for making data FAIR in your project?	See text
How will these be covered?	See text
Who will be responsible for data management in your project?	See text
Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)?	TBD

The costs of making RECETAS data FAIR will be manageable since 1) the project will rely on existing data management software and installations in the participating entities, 2) the REDCap data management platform is free to use for partners of the REDCap consortium, and 3) the ZENODO long term repository has no costs associated. Unforeseen expenses arising will be discussed within the RECETAS consortium and covered by the partners.

Data management in WP4 will be coordinated by an appointed Data Manager in FsiE, and additional oversight will be provided locally by an appointed local Data Manager in each intervention site. The data managers in each site will take responsibility in ensuring the implementation in their site of the guidelines presented in the Data Management Plan and agreed by the RECETAS Consortium. Data Managers will meet to assess systems in place and challenges identified.

In future versions of the Data Management Plan there will be more detail regarding resources for long term preservation (costs and potential value, who decides and how what data will be kept and for how long).

5. Data security

What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)?	See text
Is the data safely stored in certified repositories for long term preservation and curation?	Yes

Electronic data security procedures and provisions are detailed at institution level in Deliverable 8.2 "POPD - Requirement No. 4", including a description of the technical and organisational measures that will be implemented to safeguard the rights and freedoms of the data subjects/research participants, and the security measures that will be implemented to prevent unauthorised access to personal data or the equipment used for processing.
Physical datasets (e.g., paper records, data extraction paper forms, audio recordings) will be kept in lockable cabinets or offices with controlled access, in each participating unit. Consent forms may be scanned and stored electronically, following each institution's policies and procedures for retaining records. If consent forms are stored electronically, the original paper forms are not required to be kept.

Electronic data will be archived in standard formats (e.g., csv or txt), with auxiliary dictionary files describing the variables, values and format. During the study conduct, electronic data will be stored in project-dedicated encrypted drives on secure servers, and subject to regular data back-up procedures. Access to electronic data and records will be controlled by passwords to restrict access to data and records to authorised individuals only. Access controls are expected to be regularly reviewed and updated as individuals join, leave, or change roles within the project.

After completion of the study and closure of the database, the electronic datasets will be stored at ZENODO or similar platform. The repository is expected to establish regular data back-up procedures, and access to electronic data and records will be controlled by passwords to restrict access to authorised individuals only. Access controls are expected to be regularly reviewed and updated as individuals join, leave, or change roles within the project. All the responsibilities concerning data recovery and secure storage will go to the repository storing the dataset. The Zenodo repository has properly addressed this issue. No associated costs will be incurred specifically for data preservation, since ZENODO is a free resource.

6. Ethics aspects and Intellectual Property

Are there any ethical or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA).	TBD
Is informed consent for data sharing and long-term preservation included in questionnaires dealing with personal data?	Yes

Unless otherwise specified, copyright and intellectual property of the data generated will remain with the RECETAS Consortium.

Data management in RECETAS will abide by the confidentiality and security obligations to protect personal data established in the Grant Agreement. Informed consent will let research participants know about plans for sharing (i.e., within RECETA's research team or more widely) their personal data, as well as plans regarding to longer-term preservation and retention of their data to support reuse. Templates of the informed consent/assent forms and information sheets for human participants will be submitted as Deliverable D8.1 " H - Requirement No. 2".

A description of the technical and organisational measures that will be implemented to safeguard the rights and freedoms of the data subjects/research participants, including a description of the anonymization/pseudonymization techniques that will be implemented in RECETAS will be submitted as Deliverable D8.2 "POPD - Requirement No. 4"

In future updates of the Data Management Plan, further detail will be provided regarding any restrictions or constraints on data sharing, such as embargo periods or restricted access.

In future updates of the Data Management Plan, additional restrictions, text modifications or specifications may be added due to requirements from the ethics committees assessing the studies in RECTAS. Also, future updates will present further detail regarding any restrictions or constraints on data sharing, such as embargo periods or restricted access.

7. Other issues

Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones?	No

References

- RECETAS Grant Agreement
- General Data Protection Regulation (EU) 2016/679 [Accessible at https://eur-lex.europa.eu/eli/reg/2016/679/oj]
- Fact sheet Open Access in Horizon 2020 [Accessible at https://ec.europa.eu/programmes/horizon2020/sites/default/files/FactSheet_Open_Access.pdf]
- Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020 [Accessible at https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf]
- Dublin Core Metadata Initiative http://dublincore.org/schemas/
- ZENODO repository http://about.zenodo.org/policies/

Annexes

Horizon2020 FAIR DMP template

The template is a **set of questions** that the DMP should answer with a level of detail appropriate to the project.

1. Data Summary

What is the purpose of the data collection/generation and its relation to the objectives of the project?
What types and formats of data will the project generate/collect?
Will you re-use any existing data and how?
What is the origin of the data?
What is the expected size of the data?
To whom might it be useful ('data utility')?

2. FAIR data

2. 1. Making data findable, including provisions for metadata

Are the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers)?
What naming conventions do you follow?
Will search keywords be provided that optimize possibilities for re-use?
Do you provide clear version numbers?
What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.

2.2. Making data openly accessible

Which data produced and/or used in the project will be made openly available as the default? If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.
Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if relevant provisions are made in the consortium agreement and are in line with the reasons for *opting out.*
How will the data be made accessible (e.g. by deposition in a repository)?
What methods or software tools are needed to access the data?
Is documentation about the software needed to access the data included?
Is it possible to include the relevant software (e.g. in open source code)?
Where will the data and associated metadata, documentation and code be deposited? Preference should be given to certified repositories which support open access where possible.
Have you explored appropriate arrangements with the identified repository?
If there are restrictions on use, how will access be provided?
Is there a need for a data access committee?
Are there well described conditions for access (i.e. a machine readable license)?
How will the identity of the person accessing the data be ascertained?

2.3. Making data interoperable

Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e., adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)?

What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?

Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?

In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?

2.4. Increase data re-use (through clarifying licences)

How will the data be licensed to permit the widest re-use possible?

When will the data be made available for re-use? If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.

Are the data produced and/or used in the project useable by third parties, after the end of the project? If the re-use of some data is restricted, explain why.

How long is it intended that the data remains re-usable?

Are data quality assurance processes described?

3. Allocation of resources

What are the costs for making data FAIR in your project?

How will these be covered?

Who will be responsible for data management in your project?

Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)?

4. Data security

What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)?

Is the data safely stored in certified repositories for long term preservation and curation?

5. Ethical aspects

Are there any ethical or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA).

Is informed consent for data sharing and long-term preservation included in questionnaires dealing with personal data?

6. Other issues

Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones?