



Let's care

BUILDING SAFE AND CARING SCHOOLS
TO FOSTER EDUCATIONAL INCLUSION
AND SCHOOL ACHIEVEMENT

D7.9 Data Management Plan



Funded by
the European Union

Horizon Europe research and innovation programme
project number 101059425

Funded by the European Union under grant agreement of the project 101059425. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.



DELIVERABLE ID

Deliverable No.	D7.9	Work Package No.	WP7	Task/s No.	T7.5
Work Package Title	Project management and legal and ethical challenges				
Linked Task/s Title	Data management				
Dissemination level	PU	<i>(PU-Public, SEN-Sensitive)</i>			
Due date deliverable	31/03/2023	Submission date	31/03/2023		
Deliverable name	Data Management Plan				

DOCUMENT CONTRIBUTORS

Deliverable responsible	1. COM	
Contributors		Organisation
EVA BAJO MARCOS		COM
PIETER GRYFFROY		TLX
FLORIAN MAURER		FHV
Reviewers		Organisation
ALBA COUSO		COM
GIULIA DI LISIO		COM
ANTONIO MILÁ		COM
ABEL MUÑIZ		ZIC

DOCUMENT HISTORY

Version	Date	Comments
0	2023-01-24	First draft version
1	2023-03-24	Second draft with contributions from DMB and other partners
2	2023-03-30	Final version including all content
3	2023-03-31	Final review to correct forma issues in accordance to Let's Care visual identity



TABLE OF CONTENTS

Executive Summary	5
1. Conformation and tasks of the Data Management Board (DMB)	6
2. Research Data Management (RDM) and the FAIR principles	6
3. Data management strategy	9
3.1. Summary of the data in the Let's Care project.	10
3.2. Involvement and roles in the data management process	15
3.3. Compliance of the FAIR principles	16
3.3.1. Making the data Findable	16
3.3.2. Making the data Accessible	18
3.3.3. Making the data Interoperable	19
3.3.4. Making the data Reusable	20
4. Ethics and data security	21
References	21



LIST OF ABBREVIATIONS

Abbreviation	Description
GA	Grant Agreement
ERC	European Research Council
DMB	Data Management Board
DMP	Data Management Plan
RD	Research Data
RDM	Research Data Management
FAIR	Findability, Accessibility, Interoperability, and Reusability
IEB	Independent Ethics Board
ESL	Early School Leaving
CESSDA	Consortium of European Social Science Data Archives
DMEG	Data Management Expert Guide
CC	Creative Commons
ERC	European Research Council
DPIA	Data Protection Impact Assessment
DPO	Data Protection Officer
GPDR	General Data Protection Regulation
DCMI	Dublin Core Metadata Initiative
DDI	Data Documentation Initiative
PDI	Persistent Data Identifiers
CMS	Content Management System



Executive Summary

The Data Management Plan (DMP) summarises how the data collection, storage, protection, retention and destruction will be handled during the project execution. A Data Management Board (DMB) led by COM and co-led by TLX and FHV will oversee that the best practices towards data management and the obligations of the grant agreement (GA) related to open access are accomplished.

The present report constitutes the first version of the Let's Care Data Management Plan (DMP) corresponding to M6 (March 2023). New updates to this document will be done during the project implementation if the need to do so is detected.



1. Conformation and tasks of the Data Management Board (DMB)

The Let's Care Grant Agreement (GA) No 101059425 foresees the constitution of a **Data Management Board (DMB)** to:

1. Define the types of data produced and collected during the project.
2. Closely monitor the collection of data during the project's implementation.
3. Assess the risks and develop specific mitigation measures within a data management strategy and a specific data management plan.
4. Supervise and advice the accomplishment of the GA contractual obligations towards data management and open access of peer-reviewed scientific publications accomplishing the Horizon Europe (2021-2027) obligations and implementing the technical, ethical and legal measures applicable in the Consortium countries.

Based on the technical expertise profiles of the teams the GA appoints **COM** as the leader of the DMB, and **TLX** and **FHV** as co-leaders.

2. Research Data Management (RDM) and the FAIR principles

Research Data (RD) refers to any piece of information collected, observed, or created to produce *research results*, and so accordingly covers different *types* of data (e.g. qualitative/quantitative; small/big data...). RD can present several *formats* (e.g. digital – such as audio-visual, geospatial, etc. – or physical – such as text, tabular, etc. –) and be obtained from different *sources* (e.g. administrative and historical data, survey records, etc.) but generally its purpose is to produce valuable project results. The *size* and *complexity* of RD will vary along the research phases of a project, and the data management measures applied in each stage must adapt to the new requirements of the project and the participants involved. In the increasing complexity of research parties and processes involved, RD has a longer lifespan than the projects in which they are generated and therefore they must be responsibly curated (Figure 1).

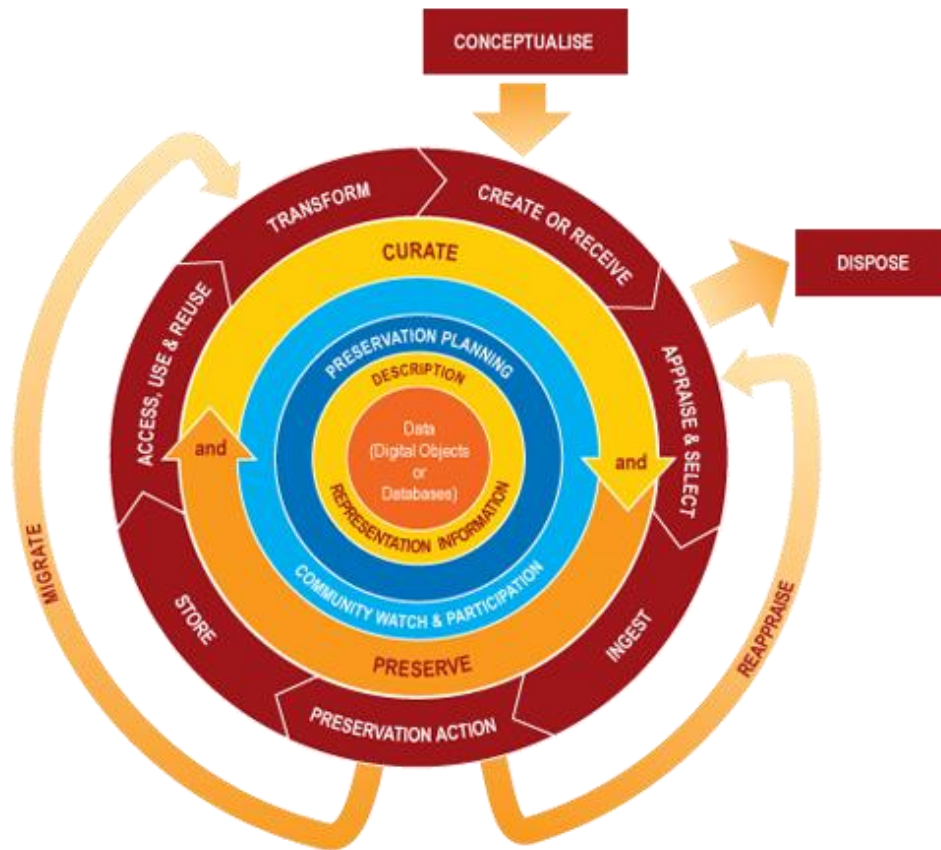


Figure 1. Research Data Management (RDM) life cycle. Retrieved from University of Southampton (2016)

Research Data Management (RDM) ensures that the information is responsibly used, preserved and curated under legal, ethical, technical and regulatory frameworks. RDM involves planning the RD structure and management strategies to follow (Figure 2). However, this process is independent of the Ethical assessment of any issues regarding the collection, processing, sharing and storage of data during the project.

Usually, the researchers are responsible for identifying ethical concerns during the research process. They must elevate them to the Principal Investigator, and then the express consult is passed to the Ethical Board for final assessment and recommendations. In the Let’s Care project, TLX is the Consortium partner for reference about the ethical and legal consultations, and additionally, an external **Independent Ethics Board (IEB)** has been appointed to supervise the critical ethical dimension of the project in **M1**¹.

¹ D8.1.

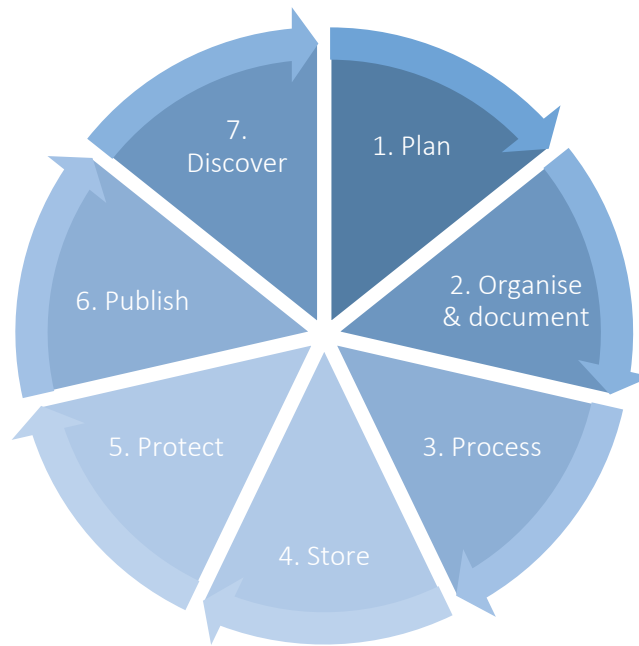


Figure 2. Research Data Management (RDM) process. Retrieved from CESSDA (2017- 2019)

The **FAIR Guiding Principles** for scientific data management and stewardship structure the data management process (Wilkinson et al., 2016): Findability, Accessibility, Interoperability, and Reusability (Table 1). The emphasis on adopting these principles is to make research outputs readily available and serving for both human-driven and machine-driven activities, and so to maximise the benefits of research results to both the scientists and society.

FAIR Principle	Implementation indicators
To be Findable	<ul style="list-style-type: none"> F1. (meta)data are assigned a globally unique and persistent identifier F2. data are described with rich metadata (defined by R1 below) F3. metadata clearly and explicitly include the identifier of the data it describes F4. (meta)data are registered or indexed in a searchable resource
To be Accessible	<ul style="list-style-type: none"> A1. (meta)data are retrievable by their identifier using a standardised communications protocol <ul style="list-style-type: none"> A1.1 the protocol is open, free, and universally implementable A1.2 the protocol allows for an authentication and authorisation procedure, where necessary A2. metadata are accessible, even when the data are no longer available



To be Interoperable	<p>I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.</p> <p>I2. (meta)data use vocabularies that follow FAIR principles</p> <p>I3. (meta)data include qualified references to other (meta)data</p>
To be Reusable	<p>R1. meta(data) are richly described with a plurality of accurate and relevant attributes</p> <p>R1.1. (meta)data are released with a clear and accessible data usage license</p> <p>R1.2. (meta)data are associated with detailed provenance</p> <p>R1.3. (meta)data meet domain-relevant community standards</p>

Table 1. Description of implementation indicators of the FAIR guiding principles. Retrieved from Wilkinson et al. (2016)

3. Data management strategy

The Let’s Care project aims at understanding how a safe attachment can positively affect early school leaving (ESL), school disengagement and educational (under)achievement in Europe, as well as design, tailor and pilot intervention tools for interventions that improve the caring dimension of educational systems in Europe. The strategy designed for the **Data Management Plan (DMP)** elaboration has followed a step driven process (Figure 3) inspired by the ERC Open Research framework² and the CESSDA Data Management Expert Guide (DMEG) (2017-2019). This strategy progressively specifies a summary of the expected data to be produced during the project’s life and builds upon the procedures to make data findable, openly accessible and interoperable; the mechanisms established to increasing data re-use and the final allocation of resources with the data security procedures.

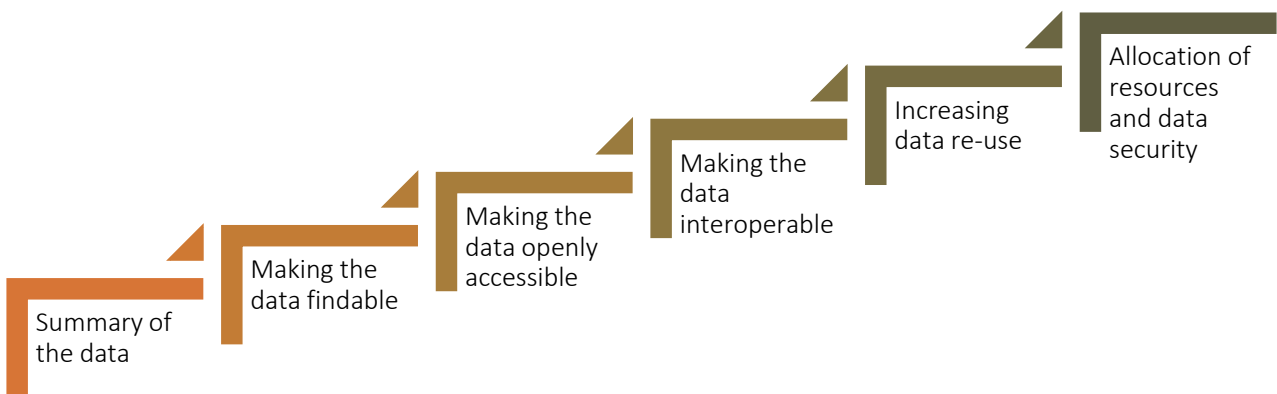


Figure 3. Summary of planned steps in the data management process.

² <https://erc.europa.eu/manage-your-project/open-science>



3.1. Summary of the data in the Let's Care project.

The Let's Care project will involve the collection, management and storage of technical research data, scientific publications and personal data when gathering becomes essential to enable the execution of the project.

Research data represents all data collected, processed and analysed with scientific research purposes to underpin determinants affecting student security of school underachievement, (under)engagement and early dropout. These data include:

- a. *Technical and methodological documentation*: refers to the needed documentation of the protocols stabilised and the materials elaborated within and for the implementation of the project e.g. questionnaires, data collection protocols, piloting materials, variable codebooks, communication and dissemination documents, etc.
- b. *Datasets*: refers to the actual information collected during project implementation. A data-level description will enable the needed processing to make the data apprehensible, cleaned and if necessary transformed to improve its quality and interoperability. The expected datasets to be collected are:
 - i. *Re-used data*: in principle, the project aims to collect primary data in the work packages specifically dedicated to obtain information directly from participants and stakeholders at educational settings. However, when possible, the use of other secondary sources from previous (particularly, European) research projects will be encouraged. The aim is building upon previous quality research findings and to adhere to the Open Science Horizon Europe strategy. The research within WP2 will help to identify any possible data set in secondary sources that might be relevant for Let's Care. This is not definitive and further updates might be done at a later stage of the project.
 - ii. *Qualitative data*: qualitative data will be collected as part of T3.1, T3.2 and T3.3. The design and implementation of data collection will take place between M4 and M30 of the project. Joint analysis of qualitative and quantitative data will take place as part of T3.5, from M19 to M40 of the project. The reason to collect these datasets is first, to develop scientific knowledge (theoretical, methodological and empirical) to build and validate a Safe Education model that explains how the caring dimension of educational systems influence early school leaving and academic (under)achievement and, second, providing a logical model for educational interventions aimed at improving these educational outcomes.



- iii. *Quantitative data*: qualitative data will be collected as part of T3.4. The design and implementation of research will take place between M13 and M30 of the project. Joint analysis of qualitative and quantitative data will take place as part of T3.5, from M19 to M40 of the project. The reason to collect these data sets is first, to develop scientific knowledge (theoretical, methodological and empirical) to build and validate a Safe Education model that explains how the caring dimension of educational systems influence early school leaving and academic (under)achievement and, second, providing a logical model for educational interventions aimed at improving these educational outcomes.

DATASET REFERENCE	DATASET NAME	WP	SOURCE ORIGIN	DATA TYPE	DATA SIZE AND FORMAT ¹
SECONDARY DATA					
SECONDS01	Re-used dataset 01	WP3	<i>To be evaluated during project implementation</i>	<i>To be evaluated during project implementation</i>	<i>To be evaluated during project implementation</i>
PRIMARY DATA					
QUAL01	Qualitative dataset 01	WP3	Let's Care data collection with families and children	QUALITATIVE. 24 in-depth interviews	Set of small files collectively small (megabytes to less than 100Gb)
QUAL02	Qualitative dataset 02	WP3	Let's Care data collection with children	QUALITATIVE. 12 life histories	Set of small files collectively small (megabytes to less than 100Gb)
QUAL03	Qualitative dataset 03	WP3	Let's Care data collection with teachers	QUALITATIVE. 12 focus groups	Set of small files collectively small (megabytes to less than 100Gb)
QUAL04	Qualitative dataset 04	WP3	Let's Care data collection in schools	QUALITATIVE. 24 ethnographies	Set of small files collectively small (megabytes to less than 100Gb)
QUAL05	Qualitative dataset 05	WP3	Let's Care data collection with policy-makers and stakeholders	QUALITATIVE. 24 Semi-structured interviews	Set of small files collectively small (megabytes to less than 100Gb)
QUAL06	Qualitative dataset 06	WP4 WP5	Data collection for the development and piloting of the	QUALITATIVE. Workshops and pilots	Set of small files collectively small



			Safe Teaching Training program		(megabytes to less than 100Gb)
QUAL07	Qualitative dataset 07	WP4 WP5	Data collection for the development and piloting of the Safe School Label	QUALITATIVE. Workshops and pilots	Set of small files collectively small (megabytes to less than 100Gb)
QUANT01	Quantitative dataset 01	WP3	Let's Care data collection in schools	QUANTITATIVE. Survey data from 120 schools	Small file (megabytes to less than 100Gb)
QUANT02	Quantitative dataset 02	WP3	Let's Care data collection with students	QUANTITATIVE. Survey data from 18000 students	Small file (megabytes to less than 100Gb)
QUANT03	Quantitative dataset 03	WP3	Let's Care data collection with teachers	QUANTITATIVE. Survey data from 2400 teachers	Small file (megabytes to less than 100Gb)
QUANT04	Quantitative dataset 04	WP4 WP5	Data collection for the development and piloting of the Safe Learning profile	QUANTITATIVE. Online questionnaire data	Set of small files collectively small (megabytes to less than 100Gb)
QUANT05	Quantitative dataset 05	WP4 WP5	Data collection for the development and piloting of the Safe Teaching Training program	QUANTITATIVE. Online/offline questionnaire data	Set of small files collectively small (megabytes to less than 100Gb)
QUANT06	Quantitative dataset 06	WP4 WP5	Data collection for the development and piloting of the Safe Teaching Training program	QUANTITATIVE. Online/offline questionnaire data	Set of small files collectively small (megabytes to less than 100Gb)

Note: ¹ to be re-assessed during project implementation

Table 2. List of Let's Care envisioned datasets


- c. Diagnosis and intervention tools: As part of WP4 and WP5, the Let's Care tools will be generated and piloted. The work will take place between M30 and M46 of the project implementation. It will involve the use of the data collected within WP3 and further data collected for tailoring, test and validate the final tools, involving workshops, piloting and evaluation with schools, teachers and institutions. The reason to collect these data is to make possible the definition of the intervention model and tools aimed at improving early school leaving and academic (under)achievement. The size of these tools is still to be determined but will represent a small file or set of files (megabytes to less than 100 GB).



TOOL	WP	ORIGIN	DATA TYPE	PUBLICATION VENUE PLANNED
Let's Care Safe Education Indexes	WP4	Primary testing, uptake, and evaluation in Extremadura	Software	Let's Care Hub
Let's Care Safe Learning e-profile	WP4	Primary testing, uptake, and evaluation in Extremadura	Software	Let's Care Hub
Let's Care Safe Teaching Training program	WP4	Primary testing, uptake, and evaluation in Portugal, Italy and Spain	Educational materials (text)	Let's Care Hub
Let's Care Safe Teaching toolkit (safe-peer observation tool & wiki database)	WP4	Data collection with teachers in WP2 and WP3.	Educational materials (text) and web	Let's Care Hub (in the section 'Safe teaching Lab')
Let's Care Safe School Label	WP4	Primary testing, uptake, and evaluation with institutions	Software	Let's Care Hub

Table 3. List of Let's Care tools

Scientific publications and other reports: The project will adhere to the Horizon Europe (2021-2027) obligations related to scientific publications:

REQUIREMENT	DESCRIPTION
Acknowledgement of the European Research Council (ERC) funding	<p>All publications related to the project must include the acknowledgement to the European Union (EU) support by including the following formula and the EU emblem:</p> <p><i>"Funded by the European Union (ERC, acronym, project number). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them."</i></p> 
Accountability of contractual obligation when selecting the publishing venue of the authors' choice	<p>The scientific publications produced within the project must consider the contractual obligation and, as long as these are granted, the authors can chose the publishing venue of their choice to submit their works.</p> <p>In this regard, the authors must retain sufficient intellectual property rights with an appropriate licence (whether by CC BY or equivalent). The ERC suggests useful resources (SHERPA/RoMEO service; Directory of</p>



	Open Access Journals ; list of compliant book publishers ; Open Access Toolkit).
At publication time, the publications must be immediately deposited in a trusted open access repository	Two critical aspects must be considered: (1) provide the specific information and metadata of the publication required by the GA and (2) use a trusted repository. In this regard, either a golden, green or a hybrid route can be undertaken as long as they preserve appropriate licence (whether by CC BY or equivalent).
Open metadata and information needed for the validation of the conclusions must be provided	All the project outputs must be public in order to validate the conclusions reached. In this regard the metadata of deposited publications must be (1) open under a Creative Commons Public Domain Dedication (CC 0) or equivalent, (2) in line with the FAIR principles (Findable, Accessible, Interoperable and Re-usable) and (3) appoint the ERC grant number and the grant acronym.

Table 4. Review of contractual obligation of the Horizon Europe (2021-2027) framework

The Let’s Care project has the compromise to produce several scientific documents and articles:

SCIENTIFIC OUTPUT	WP	DISSEMINATION LEVEL	PUBLICATION VENUE PLANNED
3 Doctoral thesis	N/A	Open access. Non-commercial exploitation	Public repositories of the corresponding university (VERITATI for UCP, DSpace repository for COM, OPUS for FHV, AIK Repozytorium)
Green/White papers with recommendations	WP6	Open access. Non-commercial exploitation	Hub library
15 Scientific articles	N/A	Open access. Non-commercial exploitation	Publishing venue of the authors’ choice

Table 5. Committed scientific outputs in the Project’s Grant Agreement

Personal data refers to any information that can identify a person directly or indirectly. A combination of a few pieces of information can identify a person and so an adequate control and processing of this type of data is paramount for an ethical project execution that complies with the legal requirements. Additionally, **sensitive personal data** is a particular category of personal data that requires extra protection and includes data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs or trade union membership, genetic data, biometric data, health data, data concerning a natural person's sex life or sexual orientation. Data may be sensitive for several reasons, including legal (e.g. GDPR, contractual), commercial, security or ethical reasons. TLX, as leaders of T7.3 and T7.4 of the project, will provide GDPR and legal and ethical support to all partners. In this regard, they organised a first **legal and ethical workshop** on **01/02/2023** to extend



this support to the partners and establish a clear communication channel for consultation before any research activity occurs.

The collection of primary data might involve gathering personal data at some point. Identifying data sets containing personal data is also relevant to consider whether sharing the data (in the current format) is possible. A good overview of data sets managed in the project will allow for verifying that GDPR checks have been carried out for processing activities related to these data sets. The data sets in the project that are sensitive will be marked as such, and it will be assessed whether they can be shared at all, and if so, in what format or with which adaptations and security measures to be applied to respect the principle of open data as much as possible, while also acknowledging legitimate concerns that attach to sharing data containing sensitive elements. When personal data is present in a data set, *pseudonymisation* techniques and *anonymisation* techniques will be used where appropriate. Any *contractual arrangements* under the GDPR will be prepared and measures and safeguards will be proposed when needed. For the data processing through the Hub, a **Data Protection Impact Assessment (DPIA)** will be carried out.

Additionally, personal data might be gathered to manage usual administration as well as the assistance and registration to communication and dissemination activities. In these cases, the partners in charge of data management will provide the complete information about the particular purposes for the collection of these data procedures asking for consent to process data and making available alignment or combination of the rights to restriction, erasure or destruction of personal data.

3.2. Involvement and roles in the data management process

The partner teams will implement technical, information security management and organisational measures to ensure privacy and data protection rights in the project (Table 6). All partners will adhere and follow GDPR regulations, being each team responsible for their own data control and processing, and having a designated DPO for consultation. Further clarification on the particular roles and definition on the responsibilities and ownership of data will be carried out benefiting from the work of T7.3 and T.4 regarding legal and ethical requirements and GDPR compliance and D6.2 defining the preliminary exploitation strategy of the project.

N°	ROLE	SHORT NAME	LEGAL NAME	COUNTRY
1	COO	COM	UNIVERSIDAD PONTIFICIA COMILLAS	ES



2	BEN	CID	CIDALIA CONSULTORIA TECNICA EN DIVERSIDAD SLL	ES
3	BEN	FHV	FACHHOCHSCHULE VORARLBERG GMBH	AT
4	BEN	POLO	ISTITUTO COMPENSIVO DI BOSCOCHIESANUOVA	IT
5	BEN	PROMA	FUNDACION PROMAESTRO	ES
6	BEN	TLX	TIME.LEX	BE
7	BEN	JEX	CONSEJERIA DE EDUCACIÓN Y EMPLEO - JUNTA DE EXTREMADURA	ES
8	BEN	ARID	STOWARZYSZENIE ARID	PL
9	BEN	PRSC	PANEVEZIO RAJONO SVIETIMO CENTRAS	LT
10	BEN	ZIC	ZABALA INNOVATION CONSULTING, S.A.	ES
11	BEN	UCP	UNIVERSIDADE CATOLICA PORTUGUESA	PT
12	BEN	AIK	AKADEMIA IGNATIANUM W KRAKOWIE	PL
13	BEN	IPA	STICHTING INTERNATIONAL PARENTS ALLIANCE	NL
14	BEN	KITE	KITE	BG

Table 6. List of project partners and their roles in the Let's Care project.

3.3. Compliance of the FAIR principles

3.3.1. Making the data Findable

For the **project-level technical documentation**, each team will follow their institutional *internal protocols of data organisation* to make daily practice operative. Regarding the data sharing among partners for project execution, the Consortium reached an agreement on using **Nextcloud**³ hosting in the cloud as an intranet space for file management to share technical documentation and materials. This decision was based on the functionalities of collaborative edition of files, remote storage on cloud, securitised and trustworthy hosting with restricted access, and open-source community development. ZIC as part of the Project Office has been designated as the partner in charge of

³ <https://cloud.europole.org/>



managing the organisation and denomination of files in the intranet. The folders within this cloud had been predefined and structured by WPs and their corresponding tasks; nonetheless, the partners can create ad hoc folders when necessary within each task folder following a logical structure. The documentation to be shared in Nextcloud include for instance internal documentation and minutes, agenda, calendars, etc.

The **data-level documentation** will be organised in simple files (whether at internal local archives or the shared intranet in Nextcloud) and only more complex databases will be used in the case of bigger datasets derived from research activities. Data quality monitoring will be part of the methodological design of each research tasks ensuring specific planning and assessment. Folder structure, file naming and organisation of variables in a codebook will be decided upon completion of the designing stage of each empirical activity. The leading team in each research activity will be in charge of defining the strategy concerning the versioning of data files (and scripts) at each task. After confirming that the data complies with the data quality assurance procedures, the final version of shareable data will be storage in the Nextcloud corresponding folder.

The teams will follow the **metadata standards** for general research data established by the Dublin Core Metadata Initiative (DCMI)⁴ including the following fields:

- Title
- Creator
- Subject
- Description
- Publisher
- Contributor
- Date
- Type
- Format
- Identifier
- Source
- Language
- Relation
- Coverage
- Rights

⁴ <https://dublincore.org/>



The DMB will follow up on the updated recommendations under current development from the CESSDA in this regard, and adapt to them if deemed necessary in future updates of this DMP. The metadata of project-level and data-level documentation will be converted to machine-readable files following the Data Documentation Initiative for Social Sciences (DDI Alliance, 2017).

Finally, each file or set of data will have assigned specific **Persistent Data Identifiers (PDI)** – like **DOIs** –, to prevent link rot. Link rot means that the hyperlink defined for a specific data stops referring to the original source due to deletion or movement of files. Without PDIs, the data would not be findable, less to say accessible.

3.3.2. *Making the data Accessible*

Research data accessibility relates to where, when and to whom the data will be made available outside the Consortium. For this purpose, Let's Care foresees three main storage locations: the Hub, the public repository for Open Science in Zenodo and the individual institutional repositories of academic partners.

The Let's Care Hub will be design and developed as a Content Management System (CMS), which means that it will be a space online to find all the information, content and tools and to further exchange and engage with the project community. Therefore, the Hub aims are to facilitate the communication and exchange; the establishment of a lively community; and act as a storage system for project products and results, and so it will serve as communication and content management platform that will be digital, online, open-source, free access, collaborative and participative. The elements composing the Hub are:

- A library containing public deliverables, newsletters, publications, and training material; documentation of data collection (methodology, ethical protocols, questionnaires, datasets); safe education database of policies, programs, and projects; project tools; and links to external tools.
- A data collection platform (quantitative data).
- An interaction space (for the project target groups) with registration through the project website.
- A platform for secure data upload/download: the Let's Care Hub gives access to approved contacts and hub subscribers who will be able to do data uploads and downloads.

Initially, the Hub was appointed as the hosting system of the Consortium intranet as well, but, as mentioned in the previous section, Nextcloud was elected as the hosting intranet of the consortium due to its many advantages in terms of content management and cost-efficiency. D1.5 – LETS CARE Hub platform will be submitted by **M12** of the project.



The Let's Care repository in Zenodo is aimed at storing the scientific publications and datasets of the project. Zenodo was appointed in the project's proposal to automatically becoming part of OpenAIRE, an initiative supported by the European Commission to promote Open Access policies. Zenodo works with open, free and universally implementable protocols that facilitate data retrieval. In this way, the limitations on how the data can be accessed will be kept to a minimum, only applying what is absolutely necessary.

Storage in Zenodo will allow creating the unique PDIs of each dataset and research documentation to be made available. As stated in the previous section, the metadata will follow the current highest standards to make the research results readily and easily accessible to the scientific community in the long-term. In fact, as stated in its website "Zenodo uses JSON Schema as internal representation of metadata and offers export to other popular formats such as Dublin Core or MARCXML".

Open Access policies will be followed for peer-reviewed and accepted publications, conference proceedings, and workshop presentations. Published versions of the scientific publications elaborated in the framework of the Let's Care project will be the only data to be made openly available by default in Zenodo and the corresponding institutional repositories of the academic partners (VERITATI for UCP, DSpace for COM, OPUS for FHV, Repozytorium for AIK)

The future publications will be licensed preferably under Creative Commons Attribution International Public License (CC BY) or in the case that this is not possible, under another license that gives the same rights. The beneficiary of the published articles (including the authors) must keep enough ownership rights to follow open-access rules, and therefore the Creative Common Public Domain Dedication (CC 0) will be the preferential type of licencing. For books and other long texts, the license may not allow commercial use or changes to be made (e.g. CC BY-NC, CC BY-ND).

3.3.3. Making the data Interoperable

With the aim of making the intended data interoperable, the use of the following formats for the research data is proposed based on the expert advice of the UK Data Service recommendations⁵ (Table 7). The project will strive for the interoperability of the data and its integration with other systems by using standardised and familiar formats.

Data type	Recommended formats	Acceptable formats
Tabular	.csv, .tab, .sav, .dta	.txt, .xls/.xlsx, .por, .mdb/.acddb
Geospatial	.shp, .gml	.mdb,.mif
Textual	.rtf, .txt	.html, .docx

⁵ <https://ukdataservice.ac.uk/learning-hub/research-data-management/format-your-data/recommended-formats/>



Image	.tif, .m	.jpeg, .jpg, .png
Audio	.flac	.mp3, .wav
Video	.mp4, .ogv	.avchd, .mov, .wmv
Scripts	.rtf, .pdf, .htm	.txt, .doc/.docx, .xls/.xlsx, .xml

Table 7. List of recommended and acceptable formats for the Let's Care research data.

Metadata will facilitate the use of the data by other users, its combination with information from other sources, while promoting transparency. Metadata of deposited data will be freely available under the Creative Common Public Domain Dedication (CC 0).

The information provided will include things like the author(s), title, and date of publication. It will also include funding information from Horizon Europe or Euratom and details about the project name, acronym, and number. Licensing terms and persistent identifiers for the publication, authors, organisations, and grants will also be included. If applicable, there will be identifiers for any additional research output or tools used to confirm the findings in the publication.

3.3.4. Making the data Reusable

The project aims at developing high impact data and publications to be reused as much as possible. In the previous sections, we set the bases for making the data Findable, Accessible and Interoperable:

1. By following legal, ethical and technical standards during data collection, production and processing
2. By ensuring quality assessment procedures, internal data organisation and versioning.
3. By providing sufficient description of both the data and metadata (containing, for instance, the date associated with data collection, any data particularities or if data is raw or processed).
4. By following the agreed standards on metadata and choosing standardised and familiar formats
5. By storing data, tools and scientific publication in trusted storage solutions and repositories
6. By protecting the usage rights linked to the data under a clear and accessible license

On top of these standards, Let's Care teams will include information about the data origin, how to cite it to ease the promotion of the project's results. The strategy for the communication, dissemination and exploitation of the project results has been described in "D6.2 Plan for Dissemination and Exploitation including Communication Activities" that is to be submitted in M6.

In this regard, the project's webpage is the public face of the project, containing information on the project consortium and ambitions, its public deliverables (with download links when they become available), news, external links of related sites, policy, research and practice. Data re-usability will



therefore be promoted through this venue as much as possible during the project and with view to the exploitation and dissemination of its results once it is finished.

4. Ethics and data security

Relevant project partners will carry out an ethical appraisal before any of the primary data collection can take place, including data collection through the Hub. Informed consents and information sheets will be prepared and ethical clearance be obtained where this is required. Research partners are responsible for this but can count on the support of legal and ethical partner TLX, who also provides written guidance on these topics to the consortium (workshop, legal and ethical guidance report). Hence, ethical elements of the design of the research data collection and further use are checked by the relevant research partner, the internal legal and ethical partner, and elements of this will also be checked with the external Independent Ethics Board, creating layers of ethics protection. In addition, where ethical commission clearance is needed for the research, this will be obtained from the competent ethics committee.

Data security and safety is an important element of these checks and partners will make explicit how the collect, store, send and share the data in a safe manner, so this can be checked by the internal legal and ethical partner, the ethical commission where applicable, and the external Independent Ethics Board where relevant. When data sets contain personal data, a GDPR check of these elements will be required as well. As stated before, for any data collection through the Hub, this will be covered by a DPIA.

References

- CESSDA Training Team (2019). *Data Management Expert Guide*. <https://doi.org/10.5281/zenodo.3820473>
- DDI (2017). *Data Documentation Initiative Alliance metadata scheme*. <https://ddialliance.org/>
- Dublin Core (2023). *Dublin Core Metadata Initiative*. <https://www.dublincore.org/>
- European Research Council (n.d.). *Open Science*. Retrieved 20 January 2023, from <https://erc.europa.eu/manage-your-project/open-science>
- Scott, M., & Cox, S. (2016). *Introducing Research Data*. <https://doi.org/160:da6022521420>
- Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., ... Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3(1), 160018. <https://doi.org/10.1038/sdata.2016.18>