

D8.2 Data Management Plan

Project Acronym:	PoliRural	
Project title:	Future Oriented Collaborative Policy Development for Rural Areas and People	
Grant Agreement No.	818496	
Website:	https://polirural.eu/	
Contact:	info@polirural.eu	
Version:	2.1	
Date:	31 April 2021	
Responsible Partner:	Plan4all	
Contributing Partners:	21c, MAC	
Reviewers:	Zuzana Pálková (SUA), Nikolaos Marianos (NP), Milan Kalas (KAJO), Patrick Crehan (CKA), Antoni Oliva Quesada (22sistema), Petra Korkiakoski (HAMK), Christian Hartman (JIIP)	
Dissemination Level:	Public	X
	Confidential - only consortium members and European Commission Services	
Keywords:	Data management plan, personal data protection, security, data repository, metadata catalogue.	

Revision History

Revision no.	Date	Author	Organization	Description
1.1	16/10/2019	Tomas Mildorf	P4All	Table of contents and responsibilities
1.2	05/11/2019	Tomas Mildorf	P4All	Introduction
1.3	11/11/2019	John O' Flaherty	MAC	Draft DMP table & Conclusions
1.4	12/11/2019	Pavel Kogut	21c	Data summary
1.5	26/11/2019	Tomas Mildorf	P4All	Final draft
1.6	29/11/2019	Tomas Mildorf	P4All	Incorporation of comments from internal reviewers, final version for submission
1.7	21/10/2020	John O' Flaherty	MAC	New section 3 subsection in response to Review 1 request for revision "DMP ... does not give significant levels of detail ... use of personal data"
1.8	24/11/2020	John O' Flaherty	MAC	New subsection in Section 3 to add use of the newly announced Open Research Europe facility update to the DMP.
1.9	30/11/2020	Tomas Mildorf	P4All	Added Section 3.5 and Section 3.6 in response to the Review 1 comment on data security and data management.
2.0	31/03/2021	Tomas Mildorf	P4All	Minor changes throughout the document.
2.1	30/04/2021	Tomas Mildorf	P4All	Incorporation of internal reviewers' comments

Responsibility for the information and views set out in this publication lies entirely with the authors.

Every effort has been made to ensure that all statements and information contained herein are accurate, however the PoliRural Project Partners accept no liability for any error or omission.

Table of Contents

Executive Summary	3
1 Introduction.....	4
1.1 Background.....	4
2 Data Summary	6
2.1 Information flow within the project.....	6
3 PoliRural Data Management Plan	9
3.1 Zenodo.....	10
3.2 Open Research Europe	11
3.3 PoliRural Data Management Plan Overview	13
3.4 PoliRural Protection of Personal Data.....	17
3.5 PoliRural Data Security	20
3.6 Project Repository and Metadata Catalogue	21
4 Conclusions.....	22
Annex 1 Responses to the monitors' comments	23

List of Tables

Table 1 Information collected/generated per work package	9
Table 2 PoliRural DMP.....	17

List of Figures

Figure 1 Information flow in PoliRural	6
Figure 2 PoliRural in the DMPonline Tool	10
Figure 3 The PoliRural ZENODO Community.....	11
Figure 4 Open Research Europe	12

Executive Summary

As participants in the Open Research Data Pilot, PoliRural has created this Data Management Plan (DMP), which will evolve during the project lifetime. This PoliRural DMP was developed based on the “Initial DMP” template using the DMPonline tool¹, as recommended by the OpenAIRE project², and saved as PoliRural. The project is providing its research results as immediate open access on the “PoliRural_Community” on Zenodo³ and Open Research Europe⁴. More elaborated versions of this DMP will be upgraded during the remainder of the project. P4All as the leader of this PoliRural DMP and CULS as the project coordinator will continually consult with all partners to examine their dissemination plans and decide whether the effective protection or commercial exploitation of knowledge could be at risk and would adversely affect the commercial exploitation of the PoliRural outputs.

¹ <https://dmponline.dcc.ac.uk>

² <https://www.openaire.eu>

³ <https://www.zenodo.org/>

⁴ <https://open-research-europe.ec.europa.eu>

1 Introduction

This report describes the PoliRural Data Management Plan (DMP) for the Horizon 2020 Open Access mandate for the project's publications and its participation in the Open Research Data Pilot. This requires that all publications arising from the work funded in the project be published in Open Access (free online access). The project will manage and share its data during the project and maintain the DMP updated. This will facilitate collaboration with other relevant H2020 projects to avoid duplication and through synergy to achieve a greater impact and better communications.

The Annex defines the abbreviations used in this document.

1.1 Background

The European Commission is running a flexible pilot under Horizon 2020 called the Open Research Data Pilot (ORDP)⁵. The ORDP aims to improve and maximise access to and re-use of research data generated by Horizon 2020 projects and takes into account the need to balance openness and protection of scientific information, commercialisation and Intellectual Property Rights (IPR), privacy concerns, security as well as data management and preservation questions.

While open access to research data thereby becomes applicable by default in Horizon 2020, the European Commission also recognises that there are good reasons to keep closed some, or even all, research data generated in a project.

The ORDP applies primarily to the data needed to validate the results presented in scientific publications. Other data can also be provided by projects on a voluntary basis, as stated in their Data Management Plans.

This Data Management Plan (DMP) is a document providing key elements of data management in the PoliRural project. This DMP describes the data management life cycle for the data to be collected, processed and/or generated by the PoliRural project.

In order to make research data findable, accessible, interoperable and reusable (FAIR), this DMP includes information on:

- the handling of research data during and after the end of the project,
- what data will be collected, processed and/or generated,
- which methodology and standards will be applied,

⁵ http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-dissemination_en.htm

-
- whether data will be shared/made open access and
 - how data will be curated and preserved, including after the end of the project.

This document is the first version of the PoliRural DMP which will be updated over the course of the project whenever significant changes arise, such as:

- new data being generated,
- changes in consortium policies,
- changes in consortium composition and external factors (e.g. new consortium members joining or old members leaving).

The DMP will be updated at least for each of the project technical reviews by the Commission.

This DMP is a living document that will serve to all PoliRural consortium members to adhere with the conventions to manage data generated in PoliRural in a way that data can be found, accessed, made interoperable and reusable.

2 Data Summary

2.1 Information flow within the project

Figure 1 depicts the information flows within PoliRural.

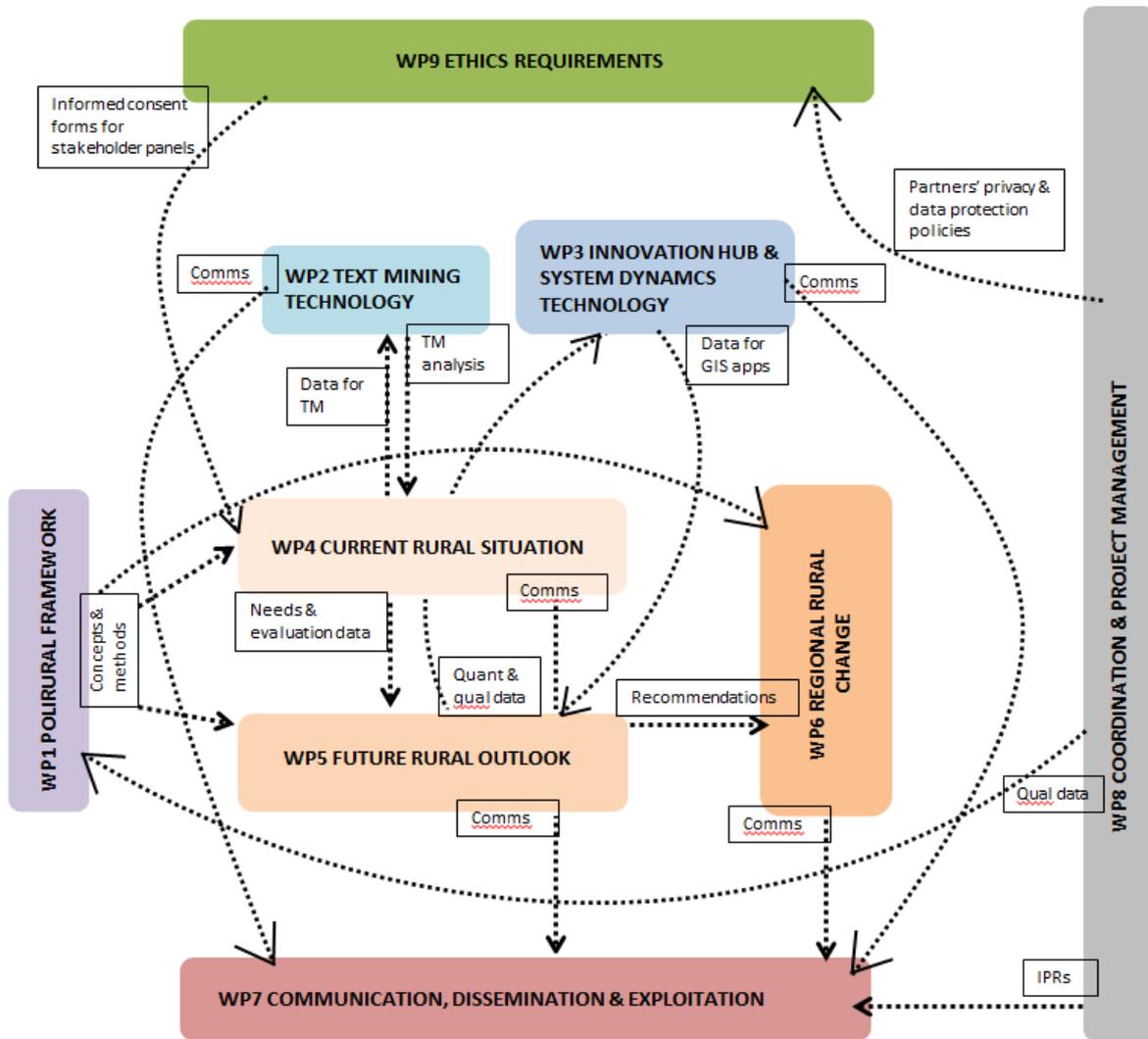


Figure 1 Information flow in PoliRural

WP1 interacts with many work packages but its primary purpose is to provide conceptual and methodological support to pilots in the form of guides, templates and reports. WP1 not only sends but also receives information to generate output. For example, ideas gathered at the kick-off (WP8) helped to shape the content of the first vision deliverable. Some research data are therefore featured in WP1 but mostly limited to consortium members; no third parties are involved.

WP2 builds a text mining solution (TM) to facilitate the understanding of rural needs and perceived effectiveness of regional policy measures. As such, the interaction aims to both

guide and collect information from pilot partners. Guidance takes the form of templates, reports and webinars, whereas data collection is carried out by means of questionnaires and online reporting tools.

WP3 oversees the development and integration of PoliRural's two other technical outputs: the innovation hub and the system dynamics (SD) tool. The hub serves as a platform for i) publishing all PoliRural results for further exploitation by the community and ii) creating new GIS services from spatial data. The hub's content, functionalities, look and feel are largely determined by pilot needs (user requirements) which are collected through questionnaires. The technical team has also collected a list of data sources that either belong to partners or third parties and which might eventually appear on the platform subject to sharing conditions. Another type of data to be collected by the hub team will come from testing activities.

The role of SD tool is to help to evaluate policy impacts under different scenarios. It will be accessible through the hub for greater convenience and sustainability. Information needed for SD development includes qualitative and quantitative data; the former to define the master model and the underlying dynamics; the latter to populate individual modules within a model. The actual data will therefore come from pilots and statistical databases.

WP4 is the first of two stages in the foresight process that investigates the current state of rural attractiveness in 12 study regions. Most of the data to be collected will come from two streams: social research and text mining. Examples of the former are surveys, interviews, focus groups, literature reviews. Text mining, for its part, will draw on web sources collected by pilots to identify certain keywords and sentiment toward a particular measure.

In WP5, the focus shifts from present to the future. The WP5 team will conduct literature reviews, collect information from pilots by means of internal and external data collection tools, and will use the SD tool to support forward looking activities.

WP6 minimises research-practice gap by targeting rural populations and recent or potential newcomers with tailored, mission-oriented interventions. Some engagement activities with external stakeholders are planned for this stage, which means our pilots will be collecting quantitative and qualitative data from research participants such as regional panel members.

No specific information will be collected as part of WP7 apart from news and events that are relevant to the project. These can be events run by pilots, other project partners or wider developments taking place in rural development.

WP8 manages the day-to-day business of the consortium, coordinating its members, managing risks and looking after the intellectual property. Most of the data to be collected is private and confidential. Some exceptions are deliverables like this one and the one on IPRs which will be made public.

WP9 will collect information on partners' data protection and privacy policies, and will implement the requirement for informed consent by asking pilots to translate and circulate the necessary forms before getting them signed by research participants.

Table 1 provides a summary of the collected or generated information per work package and the way of publication.

Work package	Information collected/generated	Information published
WP1 PoliRural Framework	<ul style="list-style-type: none"> • Research data from partners e.g. post-its from workshop, questionnaire responses • Literature: academic, blogs, news etc. 	<ul style="list-style-type: none"> • Reports
WP2 TM Technology	<ul style="list-style-type: none"> • Local data sources from pilots in the form of URLs that point to publications, social media, online media etc. that discuss rural issues 	<ul style="list-style-type: none"> • Reports • Text mining output, e.g. trained language models (to be updated during the course of the project)
WP3 Innovation Hub & SD technology	<ul style="list-style-type: none"> • List of data sources (partners' and third parties') • Hub testing data • Qualitative data about the region • Statistical databases 	<ul style="list-style-type: none"> • Datasets • Web services e.g. maps, charts
WP4 Current Rural Situation	<ul style="list-style-type: none"> • Data on rural area and policies collected through social research • Data on rural area and policies collected through text mining • Research participant data collected through Informed Consent Forms 	<ul style="list-style-type: none"> • Reports with aggregated survey and TM results
WP5 Future Rural Outlook	<ul style="list-style-type: none"> • Qualitative data e.g. foresight studies • Internal feedback from pilots in various formats • Social research directed at external stakeholders 	<ul style="list-style-type: none"> • Reports • System dynamics models

	<ul style="list-style-type: none"> • Statistical databases 	
WP6 Regional Rural Change	<ul style="list-style-type: none"> • Social research data 	<ul style="list-style-type: none"> • Reports
WP7 Communication, dissemination & exploitation	<ul style="list-style-type: none"> • Secondary sources e.g. summary of regional news on rural development • Input from pilots on their engagement activities 	<ul style="list-style-type: none"> • Online content: social media, posts on the PoliRural website and websites and social networks of PoliRural partners etc. • Promotional material • Capacity building material e.g. MOOC <ul style="list-style-type: none"> • Scientific papers and presentations.
WP8 Coordination & project management	<ul style="list-style-type: none"> • Private data e.g. contact information, financial reporting • Intellectual property rights 	<ul style="list-style-type: none"> • Reports, minutes, templates etc. <ul style="list-style-type: none"> • Public Deliverables (as in the DoA).
WP9 Ethics requirements	<ul style="list-style-type: none"> • Privacy, confidentiality and data protection policies 	<ul style="list-style-type: none"> • Informed consent form and ethics deliverables

Table 1 Information collected/generated per work package

3 PoliRural Data Management Plan

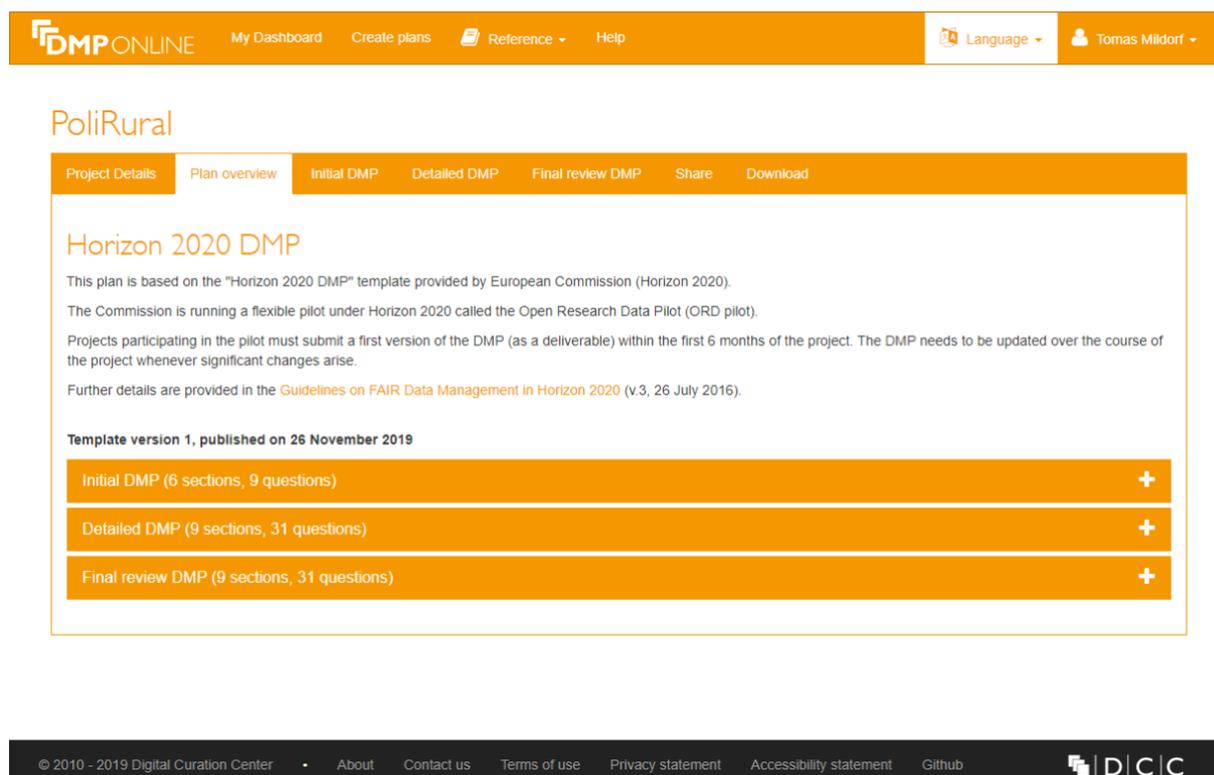
The PoliRural Data Management Plan (DMP), uses the template in Annex 1 of the document “Guidelines on FAIR Data Management in Horizon 2020”, V3, provided by the European Commission⁶, with the results from the online tool shown in the following sub-sections. The DMP is described in the following subsections, and summarised in Table 2.

This PoliRural DMP was developed based on the “Initial DMP” template using the DMPonline tool⁷ (Figure 2), as recommended by the OpenAIRE project⁸, and saved as PoliRural.

⁶ http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

⁷ <https://dmponline.dcc.ac.uk/>

⁸ www.openaire.eu



Project Details | Plan overview | Initial DMP | Detailed DMP | Final review DMP | Share | Download

Horizon 2020 DMP

This plan is based on the "Horizon 2020 DMP" template provided by European Commission (Horizon 2020).
The Commission is running a flexible pilot under Horizon 2020 called the Open Research Data Pilot (ORD pilot).
Projects participating in the pilot must submit a first version of the DMP (as a deliverable) within the first 6 months of the project. The DMP needs to be updated over the course of the project whenever significant changes arise.
Further details are provided in the [Guidelines on FAIR Data Management in Horizon 2020](#) (v.3, 26 July 2016).

Template version 1, published on 26 November 2019

Initial DMP (6 sections, 9 questions)	+
Detailed DMP (9 sections, 31 questions)	+
Final review DMP (9 sections, 31 questions)	+

© 2010 - 2019 Digital Curation Center • About Contact us Terms of use Privacy statement Accessibility statement Github 

Figure 2 PoliRural in the DMPonline Tool

3.1 Zenodo

The PoliRural project is providing its research results as immediate open access on the “PoliRural Community” (Figure 3) on the EU Commission/CERN Zenodo Open Access Repository⁹.

Zenodo is a research data general-purpose open access repository¹⁰. It was created by the OpenAIRE EU project and CERN¹¹ to provide a place for researchers to deposit datasets. It was launched in 2013, allowing researchers in any subject area to upload files up to 50GB. Zenodo is integrated with GitHub¹² to make code hosted in GitHub citable.

The PoliRural community collection at ZENODO can be accessed at:

<https://www.zenodo.org/communities/polirural/>

⁹ <https://www.zenodo.org/>

¹⁰ https://en.wikipedia.org/wiki/Open_access_repository

¹¹ <https://en.wikipedia.org/wiki/CERN>

¹² <https://github.com>

The following upload address will automatically ensure people who use it will have their record added to your community collection:

<https://www.zenodo.org/deposit/new?c=polirural>

The following address links to a OAI-PMH feed, which can be used by other digital repositories to harvest this community:

https://www.zenodo.org/oai2d?verb=ListRecords&set=user-polirural&metadataPrefix=oai_dc

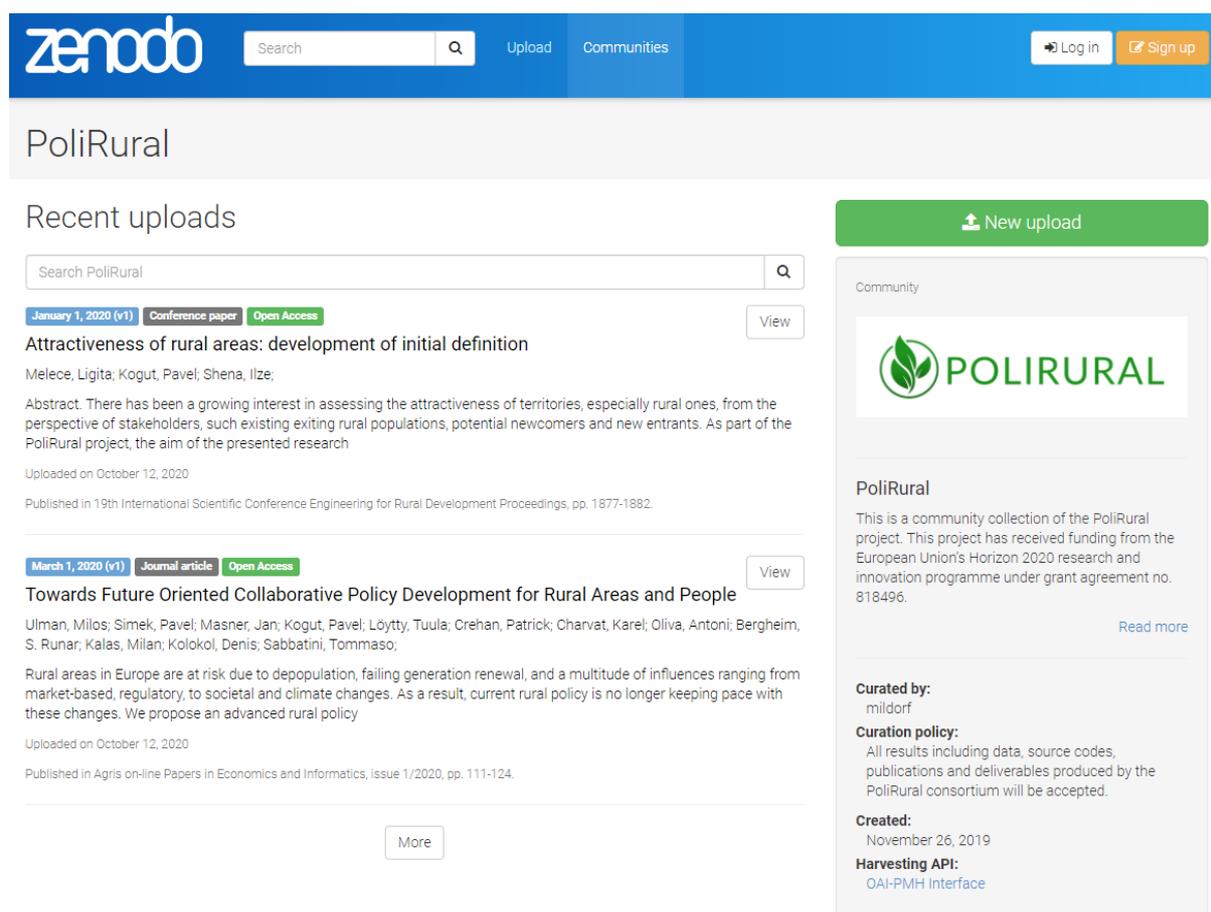


Figure 3 The PoliRural ZENODO Community

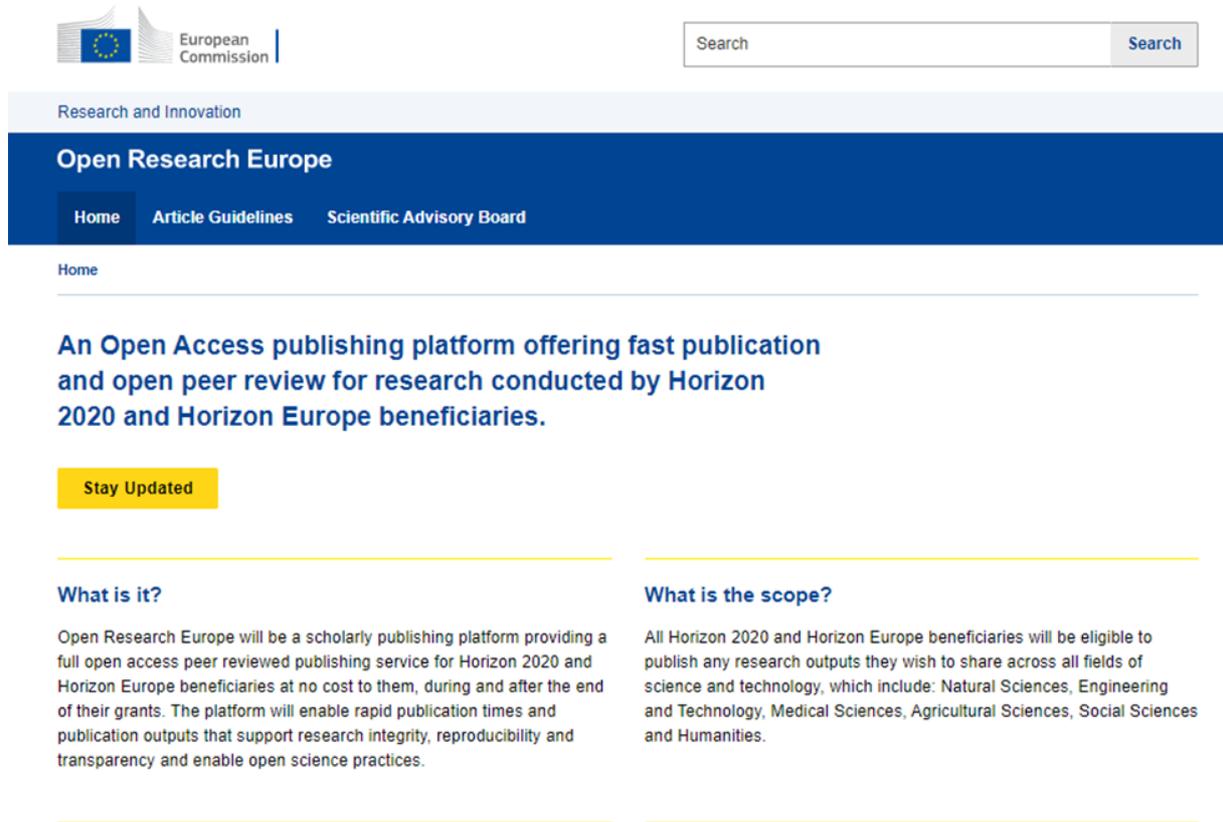
In addition to Zenodo, a project repository for data, services and related metadata was developed as part of the PoliRural Hub. See Section 3.6 for more details.

3.2 Open Research Europe

The project will also provide its research results on Open Research Europe¹³ (Figure 4), the new European Commission scientific publishing service, which will provide Horizon 2020 and

¹³ <https://open-research-europe.ec.europa.eu>

Horizon Europe beneficiaries with a venue to publish their results in full compliance with its open access policies when it is launched in 2021¹⁴.



The screenshot shows the homepage of the Open Research Europe website. At the top left is the European Commission logo. To its right is a search bar with the text 'Search' and a 'Search' button. Below the logo is the text 'Research and Innovation'. The main header is a dark blue bar with the text 'Open Research Europe' in white. Below this bar are three navigation links: 'Home', 'Article Guidelines', and 'Scientific Advisory Board'. Below the navigation bar is a 'Home' link. The main content area features a large blue heading: 'An Open Access publishing platform offering fast publication and open peer review for research conducted by Horizon 2020 and Horizon Europe beneficiaries.' Below this heading is a yellow button labeled 'Stay Updated'. The page is divided into two columns by a horizontal line. The left column has a heading 'What is it?' and a paragraph: 'Open Research Europe will be a scholarly publishing platform providing a full open access peer reviewed publishing service for Horizon 2020 and Horizon Europe beneficiaries at no cost to them, during and after the end of their grants. The platform will enable rapid publication times and publication outputs that support research integrity, reproducibility and transparency and enable open science practices.' The right column has a heading 'What is the scope?' and a paragraph: 'All Horizon 2020 and Horizon Europe beneficiaries will be eligible to publish any research outputs they wish to share across all fields of science and technology, which include: Natural Sciences, Engineering and Technology, Medical Sciences, Agricultural Sciences, Social Sciences and Humanities.'

Figure 4 Open Research Europe

¹⁴ https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/other/comm/open-research-europe_horizon-h2020_en.pdf

3.3 PoliRural Data Management Plan Overview

The PoliRural DMP is shown in Table 2.

DMP component	Issues addressed	PoliRural DMP
1. Data summary	<ul style="list-style-type: none"> • State the purpose of the data collection/generation • Explain the relation to the objectives of the project • Specify the types and formats of data generated/collected • Specify if existing data is being re-used (if any) • Specify the origin of the data • State the expected size of the data (if known) • Outline the data utility: to whom will it be useful 	<p><i>Purpose:</i> Project development, input to deliverables, input to scientific publications, dissemination, commercial advertisement.</p> <p><i>Relations:</i> real data from analysis, simulation, surveys and pilots' field-trial results</p> <p><i>Types and formats:</i> real numbers (with time tags), data files, text files, excel tables, topologies, ontologies, maps, charts and models.</p> <p><i>Existing data:</i> public policy documents and data from all levels (EU, national, regional, local), foresight/outlook studies, pilots' related websites, social media and online groups, standards and parameters, descriptions, algorithms.</p> <p><i>Origin:</i> Global, EU, national, regional, local public sources, partners, other EU projects (Ruralization, Newbie, Perceive, etc.), and pan-European initiatives (e.g. European Rural Parliament and European Rural Community Alliance, EIP-AGRI and ENRD).</p> <p><i>Expected size of data:</i> 1000 GB</p> <p><i>Utility:</i> all PoliRural Partners to reach project objectives, future collaboration, developments, widespread dissemination and commercialization after the end of the PoliRural project.</p>
2. FAIR Data		
2.1. Making data findable, including	<ul style="list-style-type: none"> • Outline the discoverability of data (metadata provision) • Outline the identifiability of data and refer to standard 	<p><i>Outline the discoverability of data (metadata provision):</i> descriptive filename, agreed data model (described within the file or in accompanying data documentation).</p>

<p>provisions for metadata</p>	<p>identification mechanism. Do you make use of persistent and unique identifiers such as Digital Object Identifiers?</p> <ul style="list-style-type: none"> ● Outline naming conventions used ● Outline the approach towards search keyword ● Outline the approach for clear versioning ● Specify standards for metadata creation (if any). If there are no standards in your discipline describe what type of metadata will be created and how 	<p><i>Code Versioning:</i> via Git Code Versioning System.</p> <p><i>Typical naming convention for documents:</i> <title> <version> <sortable date format (like yyyy-mm-dd)>, as described in D8.1 (Quality & Risk Plan)</p> <p><i>Outline the identifiability of data and refer to standard identification mechanism. Do you make use of persistent and unique identifiers such as Digital Object Identifiers:</i> different approaches will be used within the project (during, after) externally. The plan is to follow a standard. Digital Object Identifier (DOI)¹⁵ is not a suitable solution for most data types. The Metadata model should be able to "link" different file types, e.g. simulation results and analytics (e.g. architectures, parameters, operating conditions) used to produce the data. For spatial data, standardised metadata schemas such as the INSPIRE, ISO and Geo-DCAT AP will be used.</p> <p>For the developed code, APIs (Application Programming Interfaces) and other Interfaces and basic documentation will be written. Documentation will be provided in commonly used formats such as Word, PDF, Markdown (used in Git).</p> <p><i>Outline naming conventions used:</i> Will be chosen during the project. Preference will be given to standard methods.</p> <p><i>Outline the approach towards search keyword:</i> <i>Outside the project:</i> relevant keywords will be defined for every deliverable and checked for consistency. <i>Within the project:</i> the metadata, file-naming and internal folder structure is sufficient. <i>Outside the project:</i> for scientific publications, the project will use the keyword list of the publisher (e.g. IEEE keyword Taxonomy)</p> <p><i>Outline the approach for clear versioning:</i> part of the file naming convention</p> <p><i>Specify standards for metadata creation:</i> ISO and OGC standards. See Section 3.6 for more details.</p>
--------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

¹⁵ See https://en.wikipedia.org/wiki/Digital_object_identifier and www.doi.org

<p>2.2 Making data openly accessible</p>	<ul style="list-style-type: none"> Specify which data will be made openly available? If some data is kept closed provide rationale for doing so Specify how the data will be made available Specify 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)? Specify where the data and associated metadata, documentation and code are deposited Specify how access will be provided in case there are any restrictions 	<p><i>Which data will be made openly available:</i> public deliverables and dissemination material.</p> <p><i>How the data will be made available:</i> on the project website at https://polirural.eu/ and the PoliRural Digital Innovation Hub including the metadata catalogue. See Section 3.6 for more details.</p> <p><i>What methods or software tools are needed:</i> web browsers</p> <p>Other questions are not applicable.</p>
<p>2.3. Making data interoperable</p>	<ul style="list-style-type: none"> Assess the interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate interoperability. Specify whether you will be using standard vocabulary for all data types present in your data set, to allow interdisciplinary interoperability? If not, will you provide mapping to more commonly used ontologies? 	<p>Standard metadata vocabularies such as AGROVOC will be used where appropriate.</p> <p>The use of the INSPIRE registry will be encouraged.</p> <p>Data collected or generated by PoliRural will be made interoperable based on international standards, e.g. INSPIRE, ISO and OGC. In domains where hardly any standards exist (e.g. the text mining community), the most commonly used formats or formats supported by most commonly used tools will be sought.</p> <p>Developed code, APIs (Application Programming Interfaces) and other Interfaces, basic documentation will be written. Documentation will be provided in commonly used formats such as Word, PDF, Markdown (used in Git).</p>

<p>2.4. Increase data re-use (through clarifying licences)</p>	<ul style="list-style-type: none"> Specify how the data will be licenced to permit the widest reuse possible Specify when the data will be made available for reuse. If applicable, specify why and for what period a data embargo is needed Specify whether the data produced and/or used in the project is useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why <ul style="list-style-type: none"> Describe data quality assurance processes Specify the length of time for which the data will remain re-usable 	<p>Licensing is as defined in the Consortium Agreement for after the end of the Project.</p> <p>Restrictions will be defined to preserve the commercial exploitation of the project's outputs.</p>
<p>3. Allocation of resources</p>	<ul style="list-style-type: none"> Estimate the costs for making your data FAIR. Describe how you intend to cover these costs <ul style="list-style-type: none"> Clearly identify responsibilities for data management in your project Describe costs and potential value of long term preservation 	<p>Costs and responsibilities will be covered with the project budget and work allocated to WP8 "Coordination & Project Management".</p> <p>Potential value of long term preservation will be to maintain the rural attractiveness findings for policy makers and new entrants to the pilots and rural areas generally after the project ends.</p>
<p>4. Data security</p>	<ul style="list-style-type: none"> Address data recovery as well as secure storage and transfer of sensitive data 	<p>Each partner is responsible for the recoverability of their own generated data (backup plans are in place according to the institution or company practice). Access to project results is only provided to project members registered by the Coordinator. Standard security measures (authentication) are applied. See Section 3.5 PoliRural Data Security for more details.</p>

		<p>Data Protection issues are defined in Section 3.4 and in the deliverables of WP9 “Ethics requirements”.</p> <p>Privacy is a very important consideration for the project as each instance of policy and these will be implemented as part of the initial setup of the PoliRural components and services. The Partners are very aware that this data will need to be treated properly. So, while PoliRural will be very open and dynamic, there will be specific instructions and procedures, to avoid the unnecessary collection and use of personal data. And in any case where it is required, the source of the data will be clearly identified, and whether it was previously collected specifically for that purpose. In all cases, consent will be sought from the users and stakeholders involved. For more details on the PoliRural data protection policy, please refer to section 3.4 “PoliRural protection of personal data”.</p>
5. Ethical aspects	<ul style="list-style-type: none"> To be covered in the context of the ethics review, ethics section of DoA and ethics deliverables. Include references and related technical aspects if not covered by the former 	<p>As defined in the deliverables of WP9 “Ethics requirements” the Partners are committed to comply with the Ethics section of the DoA and “European Code of Conduct for Research Integrity, Revised Edition 2017”¹⁶.</p>
6. Other	<ul style="list-style-type: none"> Refer to other national/funder /sectorial /departmental procedures for data management that you are using (if any) 	<p>No specific other procedures.</p>

Table 2 PoliRural DMP

3.4 PoliRural Protection of Personal Data

As outlined in the above PoliRural DMP protection of personal data is at its core. So, while PoliRural is very open and dynamic, the following specific instructions and procedures are

¹⁶ https://ec.europa.eu/research/participants/data/ref/h2020/other/hi/h2020-ethics_code-of-conduct_en.pdf

being implemented to avoid the unnecessary collection and use of personal data, as described in D9.2.

The PoliRural management team takes lead responsibility on a day-to-day basis to ensure that data collection, processing, data storage and destruction is in line with the D9.2 guidelines and complies with current legislation. The Coordinator ensures that the research across the project follows these guidelines. Data protection requirements are included in the Steering Committee agenda and are discussed with the Steering Committee during the regular monthly calls. Each partner submits a brief report during internal reporting (every 6 months) to ensure compliance with legislation and the guidelines provided in the D9.1 and D9.2 deliverables.

The procedures implemented in PoliRural are as follows:

1. Justification for the processing of sensitive personal data

- a) PoliRural largely collects non-personal information and focuses on the participants' views related to policy and rural areas. PoliRural researchers carefully avoid sensitive issues (other than collecting sensitive data absolutely necessary for the research, such as ethnic background), and if any personal information is given, researchers take all necessary measures to ensure that the opinions cannot be directly associated to individuals. All Partners also adhere to the data protection requirements.
- b) When PoliRural presents research findings to the public, no personal information is shared; thus, it is not possible to identify individual participants afterwards. The treatment is statistical and not per individual. The collected data must be anonymised.
- c) Each beneficiary must have a Data Protection Officer (DPO) or at least adhere to the detailed data protection policy for the project in Annex 1 of D9.2.
- d) The project does not conduct any activities that can be considered unlawful according to European and international law. PoliRural fully respects the fundamental principles of research integrity as set out, for instance, in the European Code of Conduct for Research Integrity: European Code of Conduct for Research Integrity of ALLEA (All European Academies)¹⁷.

2. 'Data minimisation' principle and technical and organisational measures implemented to safeguard the rights and freedoms of the data subjects

- a) The Partners ensure that all data which is intended for processing is relevant and limited to the purposes of the research project. The data minimisation principle is crucial, and special attention is given to avoid collecting, processing or further

¹⁷ https://ec.europa.eu/research/participants/data/ref/h2020/other/hi/h2020-ethics_code-of-conduct_en.pdf

processing personal data, if it is not necessary. The amount of personal data, which is collected and used is kept to a minimum. Only the data from those participants, who have given their informed consent to the use of their data for the project research, can be processed.

b) Each Partner,

- i. Implements appropriate technical and organisational measures to ensure that, by default, only personal data, which are necessary for each specific purpose of the processing, are processed. That obligation applies to the amount of personal data collected, the extent of their processing, the period of their storage and their accessibility.
- ii. Destroys or erases any personal data without undue delay, if the personal data is no longer necessary, or which becomes no longer required during the project in relation to the purposes for which it was collected or otherwise processed.
- iii. Keeps personal data only for the time necessary to complete the research and to allow time for final publications. The data will be irreversibly destroyed by 31 May 2025 at the latest. The data will be securely destroyed or deleted in its entirety together with any back-ups making sure that it cannot be recovered. This includes third party service providers.

3. Security measures implemented to prevent unauthorised access to personal data or the equipment used for processing

Each Partner ensures the confidentiality and security of the processing of personal data under its control, as follows:

- a) The data is stored in password-protected files, on password-protected, encrypted digital devices with the passwords being changed on a regular basis to prevent any unauthorised access. The digital devices are stored in locked offices and the data is stored on secure servers at the Partner institution responsible for data collection, processing and analysis. All signed Informed Consent Forms are stored securely in lockable storage cabinets in locked offices of the researcher. For communication encrypted email, SMS and secure 'voice over IP' platforms are used.
- b) In case of breach of security leading to the accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to personal data, the Partner must inform the relevant research participant of the breach without undue delay, including

a summary description of the potential impact and a recommendation on measures to mitigate the possible adverse effects of the breach.

4. Anonymisation/pseudonymisation techniques to be implemented

When PoliRural presents research findings to the public, no personal information is shared; thus it is not possible to identify individual participants afterwards. The treatment is statistical and not per individual. As follows:

- a) Each Partner anonymises the collected data. Personal data such as name, gender, ethnic background, and contact details are stored securely. During the stages of data processing and analysis of the research findings, data is anonymised.

5. Additional measures implemented for the Innovation Hub and text mining

- a) For the Innovation Hub an email address is required as a unique identifier for the user. All other data fields are optional: users can, for example, enter a made up nickname or their real name, if desired. All data is protected and processed in accordance with the General Data Protection Regulation (GDPR) (EU 2016/679), and privacy statement is available on the Innovation Hub site.
- b) For Text Mining all data processing techniques follow a “privacy-by-design” approach, and compliance with ethical and data regulations such as General Data Protection Regulation (GDPR) (EU 2016/679) are ensured.

6. Further processing of previously collected personal data (secondary use)

- a) PoliRural does not use or make subject's data available for future research projects. PoliRural does not plan to use the data in multiple projects or for purposes other than PoliRural research.
- b) If in the course of the research project PoliRural needs to make any significant changes to the methodology or processing arrangements that have a bearing on the data subjects' rights or the use of their data, the relevant PoliRural Partners will make the data subjects aware of the intended changes, and seek and obtain their express consent. This will be implemented before any changes are made.

3.5 PoliRural Data Security

PoliRural works predominantly with open data. However, the PoliRural Innovation Hub which is based on Liferay is equipped with industry standard, government-grade encryption technologies¹⁸.

¹⁸ <https://help.liferay.com/hc/en-us/articles/360018175371-Liferay-DXP-Security-Overview>

In cases when personal or other sensitive data are collected, data are securely stored in Redmine¹⁹ protected by properly configured TLS/SSL and administered by the Czech Centre for Science and Society (CCSS). The access to the PoliRural project within this Redmine instance is provided by the project coordinator to project partners only. For each document, rights can be set in order to enable access only to certain persons.

3.6 Project Repository and Metadata Catalogue

In order to make research data more findable, accessible, interoperable and reusable (FAIR), a standardised project repository including a metadata catalogue was deployed as part of the PoliRural Innovation Hub. The final release of the hub is described in detail in D3.4 Innovation Hub - Final Release.

The project repository is composed of multiple tools for storing data and map services. This includes tools for importing additional data sources, creating map compositions or sharing documents and media.

Open Micka is used as a metadata catalogue. The catalogue is based on latest ISO and OGC standards for metadata and catalogue services and ensures interoperability with other catalogues, hubs and platforms. See Section 6.7 of D3.4 document for more details.

The data and related metadata can be accessed at <https://hub.polirural.eu/>.

In addition to the repositories mentioned above (PoliRural hub, PoliRural website, Redmine), project results related to text mining are stored at <https://semex.io/>. The architecture, authorisation, authentication and other details of this tool is described in D2.3 Final Text Mining Solution.

¹⁹ <https://redmine.ccss.cz/>

4 Conclusions

As participants in the Open Research Data Pilot, the project has created its PoliRural Data Management Plan, which will evolve during the project lifetime.

The project will continue to provide its research results as immediate open access on the “PoliRural_Community” on Zenodo. In addition, the project will provide its research results on Open Research Europe²⁰, the new European Commission scientific publishing service.

P4All as the leader of this PoliRural DMP and CULS as the PoliRural coordinator will continually consult with all partners to examine their dissemination plans and decide whether the effective protection or commercial exploitation of knowledge could be at risk and would adversely affect the commercial exploitation of the PoliRural outputs.

This latest version of the PoliRuralDMP will be updated at each periodic evaluation/assessment of the project²¹. A final version will be produced in time for the final project review. In addition, it is planned that new versions of this DMP will be created whenever important changes to the project occur due to inclusion of new data sets, changes in consortium policies or external factors.

²⁰ <https://open-research-europe.ec.europa.eu>

²¹ As recommended at www.openaire.eu

Annex 1 Responses to the monitors' comments

Comment made by the EU reviewers	Explanation
<p>DMP is basic, it does not give significant levels of detail and therefore is overly generic. For example data security 'Access to project results is only provided to project members registered by the Coordinator' or 'while PoliRural will be very open and dynamic, there will be specific instructions and procedures, to avoid the unnecessary collection and use of personal data'.</p>	<p>The DMP has been updated as follows:</p> <ul style="list-style-type: none"> - added Section 3.4 PoliRural Protection of Personal Data detailing the procedures in place to ensure personal data protection in all project activities with direct link to the D9.2 report. - added Section 3.5 PoliRural Data Security referring to the data security of sensitive or personal data collected within the project - added Section 3.6 Project Repository and Metadata Catalogue including the references to the project data repository and related tools such as the metadata catalogue - minor changes throughout the document