



DATA MANAGEMENT PLAN Update 1 (D7.2)

Oriol Biosca

Marta Carbonés Camps

Daniel Pi-Suñer

Andreu Ulied



CC BY-NC: This license allows users to distribute, remix, adapt, and build upon the material in any medium or format for non-commercial purposes only, and only so long as attribution is given to the creator.

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No GA 101061051. The work developed under this project is also supported by Innovate UK through the Horizon Europe Guarantee scheme. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or Innovate UK. Neither the European Union nor the granting authority can be held responsible for them.





Funded by
the European Union



UK Research
and Innovation

DATA MANAGEMENT PLAN Update 1 (D7.2)

Deliverable Number	WP / T
D 7.2	WP 7
Lead Beneficiary	Deliverable Authors
MCRIT	Oriol Biosca, Marta Carbonés Camps, Daniel Pi-Suñer, Andreu Ulled
Beneficiaries	Deliverable Co-Author(s)
-	-
Planned Delivery Date	Actual Delivery Date
29.02.2024	29.02.2024

Type of deliverable	R	Document, report (excluding periodic and final reports)	<input checked="" type="checkbox"/>
	DATA	Data sets, microdata, etc.	<input type="checkbox"/>
	DMP	Data management plan	<input type="checkbox"/>

Dissemination level	PU	Public, fully open, e.g. web (Deliverables flagged as public will be automatically published in CORDIS project's page))	<input checked="" type="checkbox"/>
	SEN	Sensitive, limited under the conditions of the Grant Agreement	<input type="checkbox"/>





INDEX

1.	Data Summary.....	1
1.1.	Purpose of the data collection.....	1
1.2.	Description of the data	1
1.3.	Data utility.....	3
1.4.	Data re-use	3
2.	FAIR Data.....	4
2.1.	Making data findable, including provisions for metadata	4
2.1.1.	Identifiability of data	4
2.1.2.	RUSTIK naming conventions	4
2.1.3.	Search keywords	5
2.1.4.	Metadata.....	5
2.2.	Making data accessible	6
2.2.1.	Repository for scientific publications	6
2.2.2.	Data and software repository	7
2.2.3.	Metadata.....	9
2.3.	Making data interoperable	10
2.3.1.	Interoperability.....	10
2.3.2.	Project-specific ontologies and vocabularies	10
2.4.	Increasing data re-use	11
2.4.1.	Data licensing.....	11
2.4.2.	Data re-use by third parties	11
3.	Other research outputs	12
4.	Allocation of resources.....	13
4.1.	Costs associated to fair data.....	13
4.2.	Data management accountability	13
4.2.1.	Long-term preservation.....	15
5.	Data security	16
5.1.	Long term preservation.....	16
6.	Ethics.....	17
7.	Other issues	18
BIBLIOGRAPHY.....		19
Annex 1: General outputs metadata		20
Annex 2: Geographical outputs metadata.....		21
Annex 3: Categories for datasets.....		22
Annex 4: RUSTIK Informed Consent Form.....		23





Action Number: 101061051

Action Acronym: RUSTIK

Action title: Rural Sustainability Transitions through Integration of Knowledge for improved policy processes

Date: 29/02/2024

Data Management Plan version: 2.0

Project Overview:

RUSTIK (Rural Sustainability Transitions through Integration of Knowledge for improved policy processes) is a four-year transdisciplinary research project. The project aims to enable rural communities' actors and policymakers to design better strategies, initiatives and policies fostering sustainability transitions of rural areas.

The project, funded by the Horizon Europe programme, undertakes an analysis of current adaption requirements and the support of effective rural policy-making processes for a better understanding of the different rural functionalities and characteristics as well as the potentials and challenges of rural areas. Environment, climate-energy, socio-economic and digital are the key transition pathways studied in the project.

Living Labs are placed in 14 Pilot Regions throughout 10 European countries are the central element to generate new insights into rural diversity and societal transformations. RUSTIK's Living Labs will work on the identification of new data and methods of data collection, combined with current data sources to set up relevant policy indicators. The project also focuses on data integration and dissemination, to make information and analysis accessible and valuable for actors and policymakers; and to improve rural impact assessment.

The RUSTIK information system is a web interface that provides access to relevant data sets, including indicator sets for functional rural areas, rural transitions, and resilience of rural areas to major shocks at European scale (Core European Dataset) as well as at Living Lab local/regional scales.

The final goal of the project is to enhance policy strategies and governance structures.





Executive Summary:

In accordance with the EU guidelines, RUSTIK has produced this Data Management Plan (DMP), a living document to be updated throughout the project. In accordance with guidelines for FAIR Data Management, the project aims to ensure that research data are Findable, Accessible, Interoperable and Re-usable. This first update of the Data Management Plan describes what data the RUSTIK project will generate, whether and how it will be exploited or made accessible for verification and re-use and how it will be curated and preserved after the end of the project. The data will be systemically managed and stored for the use in the project in general as well as for future research activities.

The RUSTIK Data Management Plan (DMP) aims to provide a strategy for managing data generated and collected during the project and to optimise access to and re-use of research data. RUSTIK will require the collection, processing and storage of a large variety of data sets. Specially, substantial amounts of data will be collected in WPs 1-4.

MCRIT is responsible for preparing the present document and for reviewing and updating it regularly.

While the focus of the first version of the DMP was mainly on data to be collected, this second version also reports on data produced in the context of the project and on non-sensitive data that can be made publicly available in open data repositories and registered at relevant catalogues.





1. Data Summary

1.1. Purpose of the data collection

What is the purpose of the data generation or re-use and its relation to the objectives of the project?

The primary aim of RUSTIK's data collection is to support the overarching goal of the project. This goal centers on enabling actors within rural communities and policymakers to craft more effective strategies, initiatives, and policies that foster sustainability transitions in rural areas. This is achieved by advancing our comprehension of various rural functionalities, characteristics, and future scenarios, encompassing both the potentials and challenges inherent in these regions.

In addition to its core objective, the data collection process also serves instrumental purposes. One key facet involves the exploration of how data from diverse scales can be harmonized, aiming to enhance the coherence and synergy of information across different levels. Another instrumental objective is the examination of how collected data can be effectively utilized in the design of strategies and the process of rural proofing. These dual instrumental pursuits contribute significantly to the overall efficacy of the project, ensuring that the collected data not only informs our understanding of rural dynamics but also lays the groundwork for practical and impactful interventions.

1.2. Description of the data

What types and formats of data will the project generate or re-use? What is the expected size of the data that you intend to generate or re-use? What is the origin/provenance of the data, either generated or re-used?

Part of the data will be collected from the official statistical and cartographic institutes, including European agencies as Eurostat, the European Environment Agency amongst others, and national and regional institutes, land registries or geographic institutes, and other local sources. RUSTIK will also retrieve data by new methods of data collection, including the use of big data (e.g. mobile phones, social media, etc.), citizen science and participatory methods (e.g. via Maptionnaire).

Each dataset will be easily identifiable and machine-readable through an efficient nomenclature used throughout all the project (RUSTIK naming conventions). A DOI (Digital Object Identifier) will be given to all public outputs created by RUSTIK (Identifiability of data).

The tables below set out the structure for the datasets, format and origin that will be generated and collected by RUSTIK. This second version of the DMP identifies more clearly the public data created or published by RUSTIK, as listed below:

Table. Data produced by the different WPs

Dataset name	WP Producer	Format	Source	Diss. level
Interviews with key actors	WP1	Video recording	Primary	Public (YouTube Channel)
New/improved typology of EU functional rural	WP1	PDF file	Primary	Public





Dataset name	WP Producer	Format	Source	Diss. level
areas in relation to transition pathways				
Catalogue including guidance and best practices of new sources of data and new methods of data collection	WP2	PDF file	Primary	Sensitive
Standardised structured data base at EU level	WP2	WMS, Shapefile, Raster file	Secondary (European Agencies & Global Data Provides)	Public
Structured data base in 14 pilot regions	WP2 / WP3	Shapefile, Raster file	Primary / Secondary	Public (restrictions may apply)
Improved methods for rural proofing and policy impact assessment	WP4	PDF file	Primary	Public
New place-based and enabling process model(s) for intervention logic, strategy and transition responses design at local and regional level	WP4	PDF file	Primary	Public
User-oriented tools based on the overview of operational use of data, indicators and evidence	WP5	PDF file	Primary	Public (restrictions may apply)
Synthesis report and Policy briefing on functional rural areas	WP5	PDF file	Primary	Public
4 press releases and communication campaigns	WP6	PDF file	Primary	Public
14 general infographics (one per LL)	WP6	PDF file	Primary	Public
14 short videos for Social Media	WP6	Video	Primary	Public





Dataset name	WP Producer	Format	Source	Diss. level
Webpage and related content	WP6	Varied digital material	Primary	Public
Periodical technical progress reports	WP7	PDF file	Primary	Sensitive
Periodical project accounts prepared and submitted to the commission RPI	WP7	PDF file	Primary	Sensitive
Events (internal meetings)	All	Varied digital material	Primary	Sensitive

1.3. Data utility

To whom might your data be useful ('data utility'), outside your project?

The data generated within the RUSTIK project holds significant utility for a diverse range of stakeholders, encompassing policy makers, public funders, academics, researchers, as well as urban and rural planners. The insights derived from this data will play a crucial role in facilitating a comprehensive understanding of challenges associated with rural sustainability transitions.

The information distilled from the project's data sets will not only be valuable but will also be presented in a user-friendly and intuitive manner. This user-centric approach ensures accessibility for a broad audience. Moreover, the findings will serve as essential inputs for the formulation of informed policies at various levels – be it regional, national, or European. By providing a nuanced perspective on rural sustainability, the data is poised to contribute to evidence-based decision-making processes and foster positive societal and environmental impacts.

1.4. Data re-use

Will you re-use any existing data and what will you re-use it for? State the reasons if re-use of any existing data has been considered but discarded.

Data re-use is a cornerstone of our project, serving dual purposes by contributing to the enrichment of EU-level databases and supporting the activities of the 14 Living Labs. While the creation of numerous datasets is anticipated, emphasizing data re-use remains a fundamental tenet of our project.

The specific datasets utilized by our diverse partners vary, given their origins from local, regional, or national sources, each tailored to the thematic focus of the respective Living Lab transition. For instance, project partners may leverage datasets encompassing satellite imagery, climatic conditions, elevation information, transportation details, and housing statistics. Additionally, statistical datasets covering demographics, labour market dynamics, economic indicators, and societal trends, along with census data such as the 2021 census or agricultural census, may be incorporated into the project. This broad spectrum of data sources ensures a comprehensive and contextualized approach to address the unique requirements of each Living Lab and underscores the importance of data re-use in achieving the objectives of our initiative.





2. FAIR Data

The RUSTIK Data Management Plan applies the Findable, Accessible, Interoperable, Re-usable (FAIR) approach for the project's results. The general data management policy that is presented in the subsequent chapters has been developed in accordance with Horizon 2020 guidelines for FAIR principles and for open access requirements. It applies mainly to new results that are expected to be produced by RUSTIK and that are to be made available by the project consortium as open source, open science, and open data.

All the consortium has signed before the start of the project an Informed Consent Form regarding data protection, confidentiality, risks and sharing, aligned with the principles of FAIR data ([see Annex 4](#)).

2.1. Making data findable, including provisions for metadata

2.1.1. Identifiability of data

Will data be identified by a persistent identifier?

All datasets will include metadata covering the characteristics and scale of the dataset, its origin, the methodology by which it was obtained, whether it is already used in a publication (and if so which one), and links(s) to other similar datasets within the project repository as well as external databases.

For enhanced accessibility and citation, all open data, publications, and open-source software generated by RUSTIK will be assigned a Digital Object Identifier (DOI). This DOI ensures a seamless and unique method for citing and locating content. RUSTIK leverages the services of Zenodo for DOI assignment, as this process is exclusively facilitated by DOI registrants through a registration agency.

Regarding open RUSTIK results deposited in various repositories, such as institutional repositories, scientific publishers' repositories, or other data and research repositories, they will be identified by a persistent Uniform Resource Identifier (URI) at the minimum. In cases where the institution holding the data is a DOI registrant with an agreement from a DOI registration agency, a DOI will also be assigned.

The assignment of unique identifiers (DOI, Personally Identifiable Information – PII, International Standard Serial Number – ISSN, etc.) to scientific publications depends on the chosen open-access strategy determined by the editors, the respective scientific publisher, and the selected research repository. This flexibility ensures alignment with varied publication practices while maintaining a commitment to clear and persistent identification of RUSTIK's scholarly contributions.

2.1.2. RUSTIK naming conventions

Each dataset will be easily identifiable, machine-readable, and citable. Names will include information on the covered area, its source, the year of collection, category of the data and the name of the dataset itself.

[Living Lab]_[Administrative unit]_[Transition]_[Topic]_[Year]_[Number]





Where:

- **Living Lab** is a two character code referring the specific Living Lab
- **Administrative unit** is a two values numeric code. It is specific for each Living Lab and in all cases, LAU is 10, being lower numbers the units below LAU units and bigger numbers the administrative units over this level
- **Transition** is a two character code that refers to the Socio-Economic (SE), Environmental (EN) and Digital (DG) transitions
- **Topic** is a numeric code that refers to the specific topics inside each of the transitions
- **Year** provides the dataset creation year, in four numbers
- **Number** is 01 for this first dataset matching all the previous stages and increases accordingly as a unique ID.

Example:

- GA_10_SE_03_2020_01 is a dataset for the Galicia Living Lab that is representing data at LAU level for the Socio-Economic transition and falls on the topic of Land Fragmentation. It represents data for the year 2020 and is the first dataset matching all these criteria.

Statistical themes are first defined from EUROSTAT Statistical Sub-themes (see Categories for datasets). The theme is defined by the data collector once naming the file, and in cases where none of the themes represents the topic, additional topics are allowed.

2.1.3. Search keywords

Will search keywords be provided in the metadata to optimize the possibility for discovery and then potential re-use?

Search keywords will be provided in the metadata to optimize the possibility for discovery and potential re-use.

All open RUSTIK data deposited in a repository will have search keywords as part of their metadata. Keywords for open data will be selected by the individuals publishing the data, five being the optimum number.

2.1.4. Metadata

Will rich metadata be provided to allow discovery? What metadata will be created? What disciplinary or general standards will be followed? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how. Will metadata be offered in such a way that it can be harvested and indexed?

Comprehensive metadata will be meticulously curated to enhance the discoverability of data within the RUSTIK project. The emphasis on rich metadata underscores our commitment to providing accurate and detailed information, a crucial factor in optimizing the findability and re-usability of the data outputs. Given the diverse nature of data and the multitude of metadata standards, finding a universal fit may be challenging.

To address this challenge, RUSTIK adopts a pragmatic and feasible approach. The strategy involves establishing a common and minimal catalogue metadata schema for all datasets published in public catalogues and data repositories. This collaborative decision ensures consistency and coherence, facilitating effective data discovery while acknowledging the unique characteristics of each dataset. This approach not only streamlines the metadata creation





process but also aligns with best practices, promoting accessibility and usability for a broad audience.

Metadata descriptions will follow EUROSTAT Metadata criteria (see <https://ec.europa.eu/eurostat/web/main/data/metadata>):

METADATA	
ESS Reference Metadata Reporting Standards	A set of international standards for exchange of statistical information between organisations
Classifications	International statistical classifications and nomenclatures
Code lists	Code lists are predefined, organised sets of items that describe one or more statistical concepts
Legislation and methodology	EU legal acts and methodological manuals relating to statistics
Concepts and Definitions	CODED (Eurostat's Concepts and Definitions Database) and other online glossaries relating to survey statistics
National metadata	Links to national methodology webpages: concepts, classifications, survey methodologies, etc.

For Geospatial datasets, a schema shaped after the INSPIRE Metadata Implementing Rules, based on EN ISOs 19115 and 19119 will be used. For any other dataset, a schema based on the DataCite Metadata Schema Documentation v4.46 will be used. The schemas can be found in the Annexes [1](#) and [2](#).

For data collected or produced by the project but not made available in a public catalogue or data repository (e.g. for licensing or privacy reasons), a schema with minimal metadata according to the internal needs of the RUSTIK project and partners will be used.

2.2. Making data accessible

RUSTIK open results will be made accessible according to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020. All open results (data, software, scientific publications) of the project will be openly accessible at Zenodo (Open Access repository). Specifically, research data needed to validate the results in the scientific publications will be deposited in a data repository at the same time as a publication. The project deliverables will be made available on the project website after their submission to the European Commission.

2.2.1. Repository for scientific publications

Providing open access to peer-reviewed scientific publications can be ensured either by publishing in green or gold open access journals with or without author processing fees. Any scientific publications from RUSTIK and the related bibliographic metadata must be made available as open access and announced on the project website. To automate the process of reporting scientific publications and related research data in OpenAIRE, the publication should be deposited in an OpenAIRE-compliant repository, either by the authors of the publication or by a scientific publisher. While additional forms of disseminating open access papers, including academic social network sites such as ResearchGate are possible, an electronic copy of the publication has to be deposited in a suitable open access repository in the first place.





According to the European Research Council's Guidelines on Open Access, "Venues such as ResearchGate or Academia.edu that require users to register in order to access content do not count as repositories. The posting of publications on a personal, institutional or project-specific webpage or the deposit in a publically accessible Dropbox account is not sufficient to satisfy the requirements either."

If the chosen repository is not fully OpenAIRE compliant, the publications or data must be linked at <https://www.openaire.eu/participate/claim> with the respective funding agency (European Commission). The journal's visibility and prestige (translated in the Impact Factor) of the journal, together with the speed of publication, should be considered when choosing a journal for publication of a manuscript. According to the EC recommendation, authors of the publication are encouraged to retain their copyright and grant adequate licences to publishers.

2.2.2. Data and software repository

Will the data be deposited in a trusted repository?

To ensure that data management procedures are unified across the project, a common default Open Access repository for open data and open-source software generated within the project has been chosen. In the end, it is up to the owner of the data or software to decide whether he wants to use an institutional repository or the RUSTIK default repository. Regarding datasets used in the Living Labs experiments, they will be published (if there are no privacy concerns) in the RUSTIK System.

Repositories other than this default ones should be OpenAIRE-compliant and issue a DOI (see 2.2.1). The default repository of the RUSTIK project for depositing publications, open data and open-source software is Zenodo. It is an EC-co-funded, multidisciplinary repository, for publications and data.

Does the repository ensure that the data is assigned an identifier? Will the repository resolve the identifier to a digital object?

A DOI is automatically assigned to all Zenodo files, which can be uploaded in any file format. Zenodo allows researchers to deposit both publications and research data, while providing means to link them. Data is stored in the CERN cloud infrastructure. Zenodo is compliant with the open data requirements of the EU Research and Innovation funding programme and OpenAIRE. Furthermore, a RUSTIK project page (community) is expected to be set up.

Have you explored appropriate arrangements with the identified repository where your data will be deposited?

There is no need for such an arrangement. Zenodo is OpenAIRE's recommended "catch-all" repository for projects like RUSTIK without ready access to an organized data centre.

Will all data be made openly available? If certain datasets cannot be shared (or need to be shared under restricted access conditions), explain why, clearly separating legal and contractual reasons from intentional restrictions. Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if opening their data goes against their legitimate interests or other constraints as per the Grant Agreement. Will the data be accessible through a free and standardized access protocol?

The different types of data that are generated during the project are open by default with the following general exceptions:

- **Copyright and permissions for re-using third-party data sets:** Processing and combining input data from many different sources may lead to unclear Intellectual Property Rights





(IPR) situations regarding the generated output data, therefore such repurposed data (e.g. model output data) can only be made open if any of the underlying data (e.g. model input data) is open, too.

- **Personal data treatment and confidentiality issues:** Datasets referring to the quality and quantity of certain elements at risk, such as people and critical infrastructures, are not open by default as their publication may pose privacy, ethical or security risks.
- **Data-driven business model:** Data that is exploited commercially will not be made open.
- **User-generated content:** Data related to individual adaptation scenarios (e.g. adaptation options, performance indicators, criteria, etc.) that is generated by (external) end users will only be made open with explicit permission from the end user.

If there are restrictions on data needed to validate the results presented in scientific publications, access to individuals with legitimate interest will be granted on request.

Currently, the Data Manager has no record of a beneficiary having the intention to opt-out from the present DMP provisions, neither concern on confidentiality nor permissions have been raised by the project partners. If the situation changes, the DMP will be updated accordingly.

If an embargo is applied to give time to publish or seek protection of the intellectual property (e.g. patents), specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.

Research data needed to validate the results of the scientific publications will be made available as open access at the same time as the publication. If an embargo period is imposed by the publisher, the publication and the related data are not made openly accessible until the embargo period has expired. In Horizon 2020, the embargo period must be shorter than 6 months. Information (metadata) about the publication and the related data will be made available at the same time as the publication, regardless of whether an embargo period has been imposed. Details of when the publication and the data will become available will be included in the metadata.

If there are restrictions on use, how will access be provided to the data, both during and after the end of the project?

When a restriction on open access to research data is necessary, attempts will be made to make data available under controlled conditions to other individual researchers. In the case where restricted or embargoed data is stored in the Zenodo repository, information about the restricted data will be published in the repository, and details of when the data will become available will be included in the metadata.

Metadata for both open, closed, embargoed and restricted records are always publicly available in Zenodo. Data files and data sets for restricted access records are only visible to their owners and to those the owner grants access.

Restricted access allows researchers to upload a dataset and provide the conditions under which they grant access to the data. Researchers requesting access must provide a justification proving they fulfil these conditions. The owner of the dataset gets notified for each new request and can decide to either accept or reject the request. If the request is accepted, the requester receives a secret link which usually expires within 1-12 months.





Is there a need for a data access committee (e.g. to evaluate/approve access requests to personal/sensitive data)?

In case there are any issues regarding the restricted access to research results, RUSTIK's quality assurance and ethics board can act as data access committee and seek clarification.

2.2.3. Metadata

Will metadata be made openly available and licenced under a public domain dedication CCO, as per the Grant Agreement? If not, please clarify why. Will metadata contain information to enable the user to access the data?

Metadata will be made openly available. Documentation of (open source) software needed to access the data and developed by RUSTIK will be made available on the RUSTIK website and the respective source code and release repositories.

How long will the data remain available and findable? Will metadata be guaranteed to remain available after data is no longer available?

The availability and findability of data in the RUSTIK project will be upheld throughout its lifecycle and beyond. The commitment to FAIR Data Management principles ensures that the research data will remain accessible, discoverable, and usable.

The data, along with its accompanying metadata, will be systematically managed and stored for the duration of the project. Post-project, efforts will be made to curate and preserve the data for future research activities, aligning with the project's commitment to ensuring the longevity of its research outputs.

Regarding the duration of data availability, the data will remain available and findable for at least five years after project end. The metadata will be guaranteed to remain available even after the data is no longer accessible, providing enduring documentation and context for the research outputs.

It is noteworthy that the RUSTIK Information System, designed primarily for the Living Labs, holds the potential to extend its impact beyond the project's immediate scope. In this context, collaboration with the Rural Observatory is deemed strategic. This collaboration aligns with the flagship initiative launched in 2022, based on the European Commission's Long-Term Vision for Rural Areas (LTVRA). Developed by the Joint Research Centre (JRC) in coordination with DG AGRI and DG REGIO, the Rural Observatory platform serves as a comprehensive resource offering data about rural areas. It incorporates data and knowledge generated by EU-funded projects, including RUSTIK, and provides statistics, indicators, and analyses drawn from various sources, such as JRC data, ESPON, Eurostat, and Horizon Europe projects.

The ongoing collaboration with the JRC concerning the Rural Observatory, initiated with two meetings during the summer of 2023, underscores RUSTIK's commitment to broader data exchange and accessibility. By aligning with established initiatives and platforms, RUSTIK aims to ensure the availability and discoverability of its data throughout its lifecycle and beyond.

Will documentation or reference about any software be needed to access or read the data be included? Will it be possible to include the relevant software (e.g. in open source code)?

In case there is open-source software developed by RUSTIK, it will be hosted in public GitHub repositories. A source code management system will be made available together with the respective open data in a repository. Moreover, since the source code of the software is stored on GitHub, releases can automatically be published in a supported repository.





2.3. Making data interoperable

2.3.1. Interoperability

What data and metadata vocabularies, standards, formats or methodologies will you follow to make your data interoperable to allow data exchange and re-use within and across disciplines?

To ensure the interoperability of our data for seamless exchange and re-use across diverse disciplines, RUSTIK is committed to adopting standardized vocabularies, methodologies, and formats. Interoperability is fundamental for fostering collaboration among researchers, institutions, organizations, and countries. This involves adhering to established standards for data formats, with a focus on compatibility with available (open) software applications. The goal is to enable effortless recombination with datasets originating from various sources.

It is essential to note that our data collection is an ongoing process. However, we anticipate that the resultant datasets will be disseminated in the NetCDF format. NetCDF, characterized by its self-describing and machine-independent data formats, provides a robust foundation for exchanging and re-using scientific data. Moreover, a variety of (open source) software applications are available, capable of both reading and generating NetCDF datasets. This strategic choice enhances accessibility and promotes versatility, supporting collaborative endeavours that transcend disciplinary boundaries and geographic borders.

Metadata interoperability

The following standard vocabularies will be used in the default metadata schema for all types of open data:

- License: Open Definition (<http://opendefinition.org/>)
- Funders: FundRef (<https://www.crossref.org/services/funder-registry/>)
- Grants: OpenAIRE (<http://api.OpenAIRE.eu/>)

Web interoperability

Web interoperability is a crucial factor for the functionality and usability of any tool. Proper design, development, testing, and maintenance are required to provide a compatible website or portal independent of platforms, browsers, operating systems and devices. International interoperability standards for web design, established by the World Wide Web Consortium (W3C), will be followed by RUSTIK to deliver compliant tools and services. We will use state-of-the-art development and maintenance procedures to ensure full compatibility for all users. We will additionally follow the guidelines set out in 'A Conformance to Web Content Accessibility Guidelines 2.0'.

Our software engineering team will conduct thorough stress testing on the developed pages, ensuring compatibility across a spectrum of internet browsers and devices. This proactive approach guarantees that the tools and services provided by RUSTIK are robust, reliable, and accessible to a diverse audience, fostering a positive and inclusive online experience for all users.

2.3.2. Project-specific ontologies and vocabularies

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? Will you openly publish the generated ontologies or vocabularies to allow re-using, refining or extending them?

RUSTIK will contribute to a refined understanding of rural diversity and dynamic across Europe by policymakers and rural actors, by conceptualising functional rural areas, rural transitions, and resilience, and by developing new methods of data collection and analysis, and RUSTIK Information System suitable to capture the concepts. The RUSTIK project will result in a catalogue with detailed explanations, guidance and best practices of new data sources and data collection





methods, that are expected to fill existing gaps of information on the rural diversity in Europe. Further information will be included in the third version of this DMP, as the project evolves.

2.4. Increasing data re-use

2.4.1. Data licensing

Will your data be made freely available in the public domain to permit the widest re-use possible? Will your data be licensed using standard re-use licenses, in line with the obligations set out in the Grant Agreement?

Owners of open results arising from the RUSTIK project are encouraged to release their work under a Creative Commons license, preferably Creative Commons Attribution CC-BY-4.0. Authors of scientific publications arising from the RUSTIK project are encouraged to seek an agreement with the scientific publisher of the publication that allows the authors to retain the ownership of the copyright for their work and deposit the publication in an Open Access repository.

How will you provide documentation needed to validate data analysis and facilitate data re-use (e.g. readme files with information on methodology, codebooks, data cleaning, analyses, variable definitions, units of measurement, etc.)?

Readme files will be included in the case of software development. In the case that other clarifications are needed, they will be reviewed in future versions of this DMP.

2.4.2. Data re-use by third parties

Will the data produced in the project be useable by third parties, in particular after the end of the project? Will the provenance of the data be thoroughly documented using the appropriate standards? Describe all relevant data quality assurance processes.

Open results produced by the project and deposited in a respective repository are usable by third parties after the end of the project. If confidentiality, security, personal data protection obligations or IPR issues related to specific research data that is needed to validate a scientific publication forbid open access, the data will be deposited in a restricted repository and access may be granted only upon request and under the conditions of a restricted license.

Further to the FAIR principles, DMPs should also address research outputs other than data, and should carefully consider aspects related to the allocation of resources, data security and ethical aspects.

Public documentation produced by the project will be posted both in the Zenodo repository and, when considered relevant, in the project website (rustik-he.eu).





3. Other research outputs

In addition to the management of data, beneficiaries should also consider and plan for the management of other research outputs that may be generated or re-used throughout their projects. Such outputs can be either digital (e.g. software, workflows, protocols, models, etc.) or physical (e.g. new materials, antibodies, reagents, samples, etc.).

Beneficiaries should consider which of the questions pertaining to FAIR data above, can apply to the management of other research outputs, and should strive to provide sufficient detail on how their research outputs will be managed and shared, or made available for re-use, in line with the FAIR principles.

No outputs of this kind are expected to be produced as part of the RUSTIK project. Additional details will be reported, if needed, in future versions of the DMP.





4. Allocation of resources

4.1. Costs associated to fair data

What will the costs be for making data or other research outputs FAIR in your project (e.g. direct and indirect costs related to storage, archiving, re-use, security, etc.)?

There are no immediate costs anticipated to make the open results generated in RUSTIK. Especially no costs are foreseen for storing open results in the project's default repository (Zenodo). Additional details will be reported, as needed, in future versions of the DMP.

How will these be covered? Note that costs related to research data/output management are eligible as part of the Horizon Europe grant (if compliant with the Grant Agreement conditions)

As part of the Living Lab activities in WP3, each of the Pilot Region Partners (PRP) (except for Galicia: LLC) has a budget of 30,000€ to subcontract service providers for data collection. Any unforeseen costs related to open access to research data in Horizon Europe are eligible for reimbursement during the duration of the project under the conditions defined in the Grant Agreement.

4.2. Data management accountability

Who will be responsible for data management in your project?

Data management activities concern the whole project and need to be coordinated and monitored both at project and work package level. Data management is also linked to publication of project results and thus dissemination activities. Therefore, the following roles and responsibilities can be identified:

Project Data Manager (MCRIT)
Developing the data management plan and policy in cooperation with the project management in WP7 and the technical partners.
Coordinating the technical realisation (data survey, data repositories, metadata catalogues ...).
Monitoring data management activities (both collection and publication) and deadlines.
Providing support to WP data managers, a role assigned to the WP Leaders.
Providing solutions for specific issues in accordance with project management.





Work package Data Managers (WP Leaders)
Implementing the data management policy in their respective WPs.
Monitoring data management activities and deadlines and sending reminders to partners.
Offering customized help and further guidance for filling out the WP data surveys.
Asking partners for missing information or clarifications.
Providing input to the data management plan by analysing and summarising the WP-specific data surveys.
Offering customized help and further guidance for publishing open data and open-source software.
Monitoring that open results (data and software) are deposited in the default repository or a complementary OpenAIRE-compliant repository and sending reminders to partners.
Contacting the quality assurance and ethics committee in case of questions and ethical and privacy issues that may forbid a publication of the data.

Dissemination Manager (FEUGA)
Offering assistance in choosing the right publication path.
Offering customized help and further guidance for publishing scientific publications
Ensuring that the open access policy of the journal complies with the H2020 open data requirements before the researcher submits a manuscript.
Monitoring that green access (self-archiving) publications are deposited in repositories and sending reminders to partners.
Monitoring that metadata about publications is made available in the R&I Participant Portal (preferably automatically through OpenAIRE) and on the RUSTIK website.
Monitoring that research data related to a publication is made available in repositories and linked to respective publication.





Quality Assurance and Ethics Manager (IfLS)

Performing a quality assurance and ethics assessment of open data before their publication.

Keeping contact with data managers and decide together with the project management on critical issues.

Data Provider/Scientist (Partners uploading data)

Informing the data & dissemination managers when new open data/papers ready for publication are available.

Describing the data (by means of appropriate metadata) or scientific publication in accordance to the RUSTIK data management and with help of the tools provided by the project.

Depositing (publishing into a repository) the data or scientific publication in accordance to the RUSTIK data management policy and with help of the tools provided by the project.

4.2.1. Long-term preservation

How will long-term preservation be ensured? Discuss the necessary resources to accomplish this (costs and potential value, who decides and how, what data will be kept and for how long)?

No immediate costs are anticipated for open data that is stored for long-term preservation. Additional details will be reported, if needed, in future versions of the DMP.





5. Data security

The data will be protected and managed to the General Data Protection Regulation of the EU. (The task will furthermore include management of knowledge and IPR issues according to the rules and agreements stated in the Consortium Agreement.) Critical risks will be assessed by IfLS and their handling and mitigation will be agreed and regularly updated in the RUSTIK risk register.

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

RUSTIK will follow and respect the European regulation on data protection throughout the contract and adhere to the Guidelines concerning web services. Hence, careful consideration will be given to Data Protection Regulation (EC) No 45/20011, on the protection of individuals with regard to the processing of personal data, applicable until 25 May 2018, General Data Protection Regulation (EU) 2016/6792, on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, applicable after 25 May 2018 and Guidelines of the European Data Protection Supervisor (EDPS) on Web Services.

In the case of data published in Zenodo, it is stored in CERN's EOS service in an 18 petabytes disk cluster. Each file copy has two replicas located on different disk servers. For each file, two independent MD5 checksums are stored. The servers are managed according to the CERN Security Baseline for Servers.

Security is a critical requirement since a site/portal installation must be secured from all types of attacks (intrusion, denial of service, interception of content etc.) threatening the system availability and the data integrity. For this reason, security will be ensured in three different layers, i.e. the user session (securing data exchanged between the server and the users), the application (restricting the user access depending on the corresponding user rights) and the data (securing access to the content repositories).

These provisions will apply for data retrieved at European level but also for data retrieved and generated in the Living Labs.

5.1. Long term preservation

Will the data be safely stored in trusted repositories for long-term preservation and curation?

Results that are openly shared and deposited in the Zenodo repository are subjected to secure, long-term preservation practices. This ensures the continued availability and accessibility of these valuable outcomes over an extended period. The Zenodo repository, known for its commitment to data preservation, serves as a reliable and enduring storage solution for the project's open results.

Additionally, the project's website will be maintained for a minimum of five years after the conclusion of the project. During this period, users will have continued access to the RUSTIK system, and this access will be facilitated through the same project website. This sustained online presence guarantees that stakeholders, researchers, and the wider community can revisit and engage with the RUSTIK system, accessing pertinent information and resources through the familiar platform for an extended duration.





6. Ethics

Are there, or could there be, any ethics 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).

To safeguard sensitive data, the RUSTIK project implements security measures. Ethical considerations, particularly regarding privacy, are paramount, and our data handling practices comply with relevant regulations and ethical standards.

- **Informed Consent:** where applicable, individuals contributing data to the RUSTIK project – and specifically in the RUSTIK Information System – are provided with clear and transparent information about how their data will be used. Informed consent mechanisms are in place to ensure that individuals understand and agree to the terms of data usage.
- **Anonymization and De-identification:** whenever feasible, sensitive information is anonymized or de-identified to protect the privacy of individuals. This involves removing or encrypting personally identifiable information to prevent the identification of specific individuals.
- **Compliance with Regulations:** our data handling practices adhere to relevant regulations, such as GDPR (General Data Protection Regulation) in the European Union, and other applicable ethical standards. This ensures that our procedures align with legal requirements governing the protection of personal data.
- **Data Minimization:** we practice data minimization, only collecting and retaining data that is essential for the intended purposes. Unnecessary or excessive data is not collected to reduce the risk associated with handling sensitive information.

Additional details will be reported, if needed, in future versions of the DMP.

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

RUSTIK project partners will be informed about when and where approvals are needed and how to deal with personal data. Further on, it is communicated where informed consent is needed and how to get them, e.g. addressing informed consent procedure for communication with stakeholders.





7. Other issues

Do you, or will you, make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones (please list and briefly describe them)?

Currently, the project does not make use of procedures for data management other than those described in this data management policy.

Note: Public/governmental institutions and other RUSTIK partners may have their own obligations/policies for data management, but in the project (for data collected and produced by the project) we currently only follow the EC procedure.





BIBLIOGRAPHY

European Commission (2021). Data Management Plan Template. <https://enspire.science/wp-content/uploads/2021/09/Horizon-Europe-Data-Management-Plan-Template.pdf>

European Research Council Executive Agency (2016). H2020 Programme Guidelines on FAIR Data Management in Horizon 2020. https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

European Research Council Executive Agency (2017). Guidelines on Implementation of Open Access to Scientific Publications and Research Data in projects supported by the European Research Council under Horizon 2020 https://ec.europa.eu/research/participants/data/ref/h2020/other/hi/oa-pilot/h2020-hi-erc-oa-guide_en.pdf

DOI Foundation. What are registration agencies? <https://www.doi.org/the-community/what-are-registration-agencies/>

Eurostat. EU and National metadata. <https://ec.europa.eu/eurostat/web/international-trade-in-goods/methodology/eu-and-national-metadata>

European Commission; Directorate-General for Research & Innovation (2017). H2020 Programme Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020. https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf

Crossref. Funder Registry. <https://www.crossref.org/services/funder-registry/>

OpenAIRE Graph. The OpenAIRE. <http://api.OpenAIRE.eu/>

Open Knowledge Foundation. Creative Commons Attribution License (cc-by) <https://opendefinition.org/licenses/cc-by/>

RUSTIK: Rural Sustainability Transitions through Integration of Knowledge for improved policy processes (2023). <https://rustik-he.eu>

EDPS, European Data Protection Supervisor (2016). Guidelines on the protection of personal data processed through web services provided by EU institutions. https://edps.europa.eu/sites/default/files/publication/16-11-07_guidelines_web_services_en.pdf

CERN (2021). EOS Open Storage. <https://eos-web.web.cern.ch/eos-web/>

CERN Data Centre & Invenio. About Zenodo. <http://about.zenodo.org/infrastructure/>





Annex 1: General outputs metadata

Mandatory fields that the metadata of the RUSTIK outputs have to contain. Additional fields can be included, from DataCite’s metadata schema.

GENERAL OUTPUTS METADATA	
Title	Title of the deposition.
Description	Abstract or description for deposition.
Files	Deposition files identifiers, filenames, size of the files in bytes and MD5 checksum of files.
Upload type	Type of the deposition from a controlled vocabulary (publication, dataset, software ...).
Publication date	Date of publication in ISO8601 format (YYYY-MM-DD).
Creators	The creators/authors of the deposition.
License	Open license from controlled vocabulary “Open Definition Licenses Service”.
DOI	Digital Object Identifier assigned by the DOI registrant (e.g. Zenodo), also used for versioning.
Keywords	Free form keywords for this deposition.
Related identifiers	Persistent identifiers of related publications, datasets and software.
Communities	List of communities the deposition to appears in (https://zenodo.org/communities/clarity/).
Grants	List of European Commission FP7 grants which have funded the research for this deposition (730355). Needed to establish the relationship to RUSTIK.





Annex 2: Geographical outputs metadata

Mandatory fields that the metadata of the RUSTIK outputs have to contain. Additional fields can be included, from INSPIRE’s metadata schema.

GENERAL OUTPUTS METADATA	
Title	
Description	
Type	
Locator	
Unique Identifier	
Language	
Topic category	
Keywords	
Originated Controlled Vocabulary	
Geographic location: Bounding Box	
Reference System (EPSG)	
Temporal extent	
Date of publication	
Spatial Resolution	
Condition applying to access and use	
Responsible party (role)	
Contact	





Annex 3: Categories for datasets

Source of the categories: EUROSTAT Statistics Explained.

Transport	Trade	Environment and energy
Transport	Goods	Energy
	Services	Environment

Science, technology and digital society	Agriculture, forestry and fisheries	General and regional statistics/EU policies
Digital economy and society	Agriculture	Non-EU countries
Science and technology	Fisheries	Regions and cities
	Forestry	Sustainable development goals
		Policy indicators

Economy and finance	Industry and services	Population and social conditions
Balance of payments	Short-term business statistics	Migration and asylum
Comparative price levels	Structural business statistics	Crime
Consumer prices	Business registers	Culture
Exchange rates and interest rates	Globalisation in businesses	Education and training
Government finance	Production statistics	Health
National accounts (incl. GDP)	Tourism	Labour market
		Living conditions
		Population and demography
		Social protection
		Sport
		Youth





Annex 4: RUSTIK Informed Consent Form

Informed Consent Form for participating in the project titled: Rural Sustainability Transitions through Integration of Knowledge for improved policy processes (RUSTIK)

Project partner: [name]

Name of sponsor: European Commission, Horizon Europe (Destination: Resilient, inclusive, healthy and green rural, coastal and urban communities).

This Informed Consent Form has two parts:

- Part I: Information Sheet (to share information about RUSTIK)
- Part II: Certificate of Consent (for signatures if you choose to participate)

You will be given a copy of the full Informed Consent Form.

Part I: Information Sheet

Introduction:

We approach you in name of the European RUSTIK project. This research and innovation action is funded by the European Commission's Horizon Europe programme. The project does not reflect the opinions of the European Commission. The overall objective of RUSTIK is to develop and apply a participatory, multi-actor process to (a) advance understanding of different characteristics and future scenarios of rural areas, and their potentials and challenges, that will then be used to (b) enable rural communities' actors and policy makers to design better strategies, initiatives and policies fostering sustainable transitions. The project is international and takes place over a four-year period across 12 European countries (Austria, Bulgaria, Germany, Spain, Finland, France, Italy, Poland, Sweden, Slovenia, Serbia, and the UK). We invite you to read this consent form as we would appreciate to discuss and share your specific experiences with and thoughts on the key topics of our project, about which you may talk to anybody you feel comfortable with. Please take some time to reflect on whether you would like to participate or not. If there's anything you don't understand in this Information Sheet, feel free to ask any questions at any time.

Participant selection

You are invited to contribute to RUSTIK due to your experiences as a representative of the [stakeholder description] and take part in [project activity description]. This experience may help us to understand [aspects].

Voluntary participation

Your participation in this research is entirely voluntary. You can choose to participate or not.

Procedures

The information collected is confidential. In the case of recorded information, the [audio/video] data will be kept safe in a locked filing cabinet. All audio recordings of interviews and any other tangible data (i.e. notes during the interview) will be destroyed within five years after the closure of the project.

Risks

The RUSTIK project activities might potentially include sensitive and personal issues (i.e. political opinions, cultural values). This kind of personal data shall be processed fairly and lawfully and shall be obtained only for [add a specified and lawful purpose], and shall not be further processed





in any manner incompatible with that/those purpose/s. Moreover, you do not have to answer any question if this makes you feel uncomfortable.

Benefits and reimbursement

You may benefit from the results of our RUSTIK project in the sense that its outcomes will provide more detailed insights into sustainable rural transitions, and as such aims to contribute to capacity building among rural actors, including policy-makers, academics and representatives of businesses and civil society for fostering such sustainable transitions. There will be no reimbursement for your contribution.

Confidentiality

All collected data material will be stored in a way that will guarantee anonymity. Nobody will be named at the analysis of data, although direct quotes from the interviews might be used in reporting activities and publications. Also other forms of communication of project findings may set certain limits to confidentiality. However, this research does not deal with any sensitive subjects, so the likelihood of such experiences is very small.

Data storage

All partners, including Pilot Region Partners, will store data for five years, counting from the end of the project, to make these data available to other research and practice partners at specific request.

Unexpected/incidental findings

In the event of unexpected or incidental findings, these will be jointly reflected by participants and researchers. For example, such reflection will be an integral part of joint evaluation phases foreseen in the RUSTIK Living Labs.

Sharing the results

The project findings are expected to be published in multiple ways, including events/conferences/fairs, (non-) scientific publications, dissemination workshops, a RUSTIK practice guide, and a RUSTIK final conference.

Who to contact

If you have any questions, please contact the RUSTIK partners in your national, regional or local setting [partner address + e-mail of team leaders]. If these are not responding adequately, feel free to contact the overall European project coordination:

Project leader: Simone Sterly

Institute for Rural Development Research at Goethe University Frankfurt (IfLS) Kurfürstenstraße 49, 60486 Frankfurt am Main, Germany
sterly@ifls.de

This proposal has been reviewed and approved by the [name of partner/responsible institution], whose task it is to make sure that research participants are protected from any harm.





Part II: Certificate of Consent

I, the undersigned, confirm that (please tick box as appropriate):

1.	I have read and understood the information about the RUSTIK project, as provided in the Information Sheet (Part I)	<input type="checkbox"/>
2.	I have been given the opportunity to ask questions about the project and my participation.	<input type="checkbox"/>
3.	I voluntarily agree to participate in the project.	<input type="checkbox"/>
4.	I understand I can withdraw at any time without giving reasons and that I will not be penalised for withdrawing nor will I be questioned on why I have withdrawn.	<input type="checkbox"/>
5.	The procedures regarding confidentiality have been clearly explained (e.g. use of names, anonymisation of data, etc.) to me.	<input type="checkbox"/>
6.	The use of the data in sharing, archiving, dissemination and publications has been explained to me.	<input type="checkbox"/>
7.	I, along with the RUSTIK project partner, agree to sign and date this informed consent form.	<input type="checkbox"/>

Participant:

Name Signature Date

Researcher:

Name Signature Date

