R3PACK - REDUCE, REUSE, RETHINK PACKAGING TOWARDS NOVEL FIBRE-BASED PACKAGING AND REUSE SCHEMES

Grant Agreement No. 101060806

Work package: WP1 - Project Management

Task: 1.4 - Data Management

Deliverable: 1.2 - Data Management Plan

Dissemination level: Public

Issue date: 30th of November 2022

Lead beneficiary: (RE)SET

DOCUMENT INFORMATION

HISTORY OF CHANGES		
Version	**Publication date**	**Changes**
1.0	30.11.2022	Initial version

ADDITIONAL AUTHOR(S) CONTRIBUTION	
Name	**Organization**
Elodie Schott	(RE)SET

DOCUMENT DISTRIBUTION		
Issue	**Date**	**Distributed to**
V1	30.11.2022	European Commission

VERIFICATION AND APPROVAL		
	Date	**Distributed to**
	10.11.2022	Antoine Covillers, (RE)SET
	25.11.2022	Antoine Covillers, (RE)SET

Funded by the European Union

This project has received funding from the European Union's Horizon Europe Research and Innovation programme under Grant Agreement No 101060806. This document reflects the views of the author(s) and does not necessarily reflect the views or policy of the European Commission. Whilst efforts have been made to ensure the accuracy and completeness of this document, the European Commission shall not be liable for any errors or omissions, however

TABLE OF CONTENT

Funded by the European Union

This project has received funding from the European Union's Horizon Europe Research and Innovation programme under Grant Agreement No 101060806. This document reflects the views of the author(s) and does not necessarily reflect the views or policy of the European Commission. Whilst efforts have been made to ensure the accuracy and completeness of this document, the European Commission shall not be liable for any errors or omissions, however

EXECUTIVE SUMMARY

A Data Management Plan (DMP) sets the framework for the handling of data produced in R3PACK from acquisition over curation to dissemination and shall thereby assure the implementation of best practice procedures for lifecycle management of R3PACK data during and beyond the lifetime of the project.

The following DMP is produced as part of the Research & Innovation project R3PACK, in line with Annex 5 of Grant Agreement about specific rules on Communication, Dissemination, Open Science and Visibility (Article 17).

The DMP is in line with the FAIR principles as required by the European Commission and will be deposited in a public repository to allow full access.

It describes the data that will be generated, how it will be managed and made openly accessible throughout the lifetime of the project.

It is a living document that will be updated as it evolves throughout the lifetime of the project. A new version, enriched with project outputs as the project evolves (new data sets, new publications, changes in data access or curation policies, etc.), will be submitted by M12 and M24.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe Research and Innovation programme under Grant Agreement No 101060806. This document reflects the views of the author(s) and does not necessarily reflect the views or policy of the European Commission. Whilst efforts have been made to ensure the accuracy and completeness of this document, the European Commission shall not be liable for any errors or omissions, however

INTRODUCTION

As required by Horizon Europe and according to project guidelines advocating for Open Science, R3PACK engages in not only making scientific publications publicly available, but also as advised:

- all data needed to validate the results presented in scientific publications
- any other data (described in the Digitial Output Management Plan – DOMP)
- other information about the tools used within the project to achieve current results (algorithms, surveys, analysis protocols)

The open access to R3PACK's data will allow to foster the cross-industry uptake of the newly developed solutions, as often the lack of dissemination slows down the innovation process within a sector. The data management will contribute to promoting a systemic approach to how to best design, process and commercialize new sustainable packaging.

1. DATA SUMMARY

1.1 PROJECT DESCRIPTION

R3PACK is a research and innovation project funded by the European Commission under the Grant Agreement 101060806, which aim is to reduce, reuse and rethink single-use plastic packaging. Within the given timeframe of the project the global objectives are:

- to develop sustainable fibred-based packaging solutions to substitute the existing solutions made with plastic and
- implement economically and environmentally viable reuse schemes to reduce plastic waste as well as extend packaging lifecycle.

R3PACK's consortium gathers 24 organizations from 7 different countries, bringing together key actors of the food value chain, from the packaging manufacturer to the retailer, combined with experts in the food sector, from companies providing innovative solutions to universities. With their combined expertise R3PACK will move from R&D to commercial real-life demonstration to secure fast and extensive uptake of industrially relevant, cross-sectorial, cost-effective technologies and reuse models allowing immediate substitution of complex multi-layer plastic packaging.

1.2 DATA REUSE AND GENERATION

Will you re-use any existing data and what will you re-use it for?

Funded by the European Union

This project has received funding from the European Union's Horizon Europe Research and Innovation programme under Grant Agreement No 101060806. This document reflects the views of the author(s) and does not necessarily reflect the views or policy of the European Commission. Whilst efforts have been made to ensure the accuracy and completeness of this document, the European Commission shall not be liable for any errors or omissions, however

Existing data from R3PACK's partners will be used as a starting point of the project, to build up on existing research and background knowledge. The data will mainly serve WP2 about market analysis consumer behaviours, WP3 about reuse models and WP4 looking into new fibre-based packaging solutions. The concerned data has been agreed upon and listed as "background" in the appendix of the Consortium Agreement for each partner.

The existing data may be the results from previous pilots conducted by an industrial to substitute their plastic packaging with an alternative solution or a previous experiment carried out by a laboratory or innovation center providing insights on the performance of a targeted solution within R3PACK.

What types and formats of data will the project generate or re-use?

The re-used data provides among other things key information about the industrial products specifications, the technical performance of the enablers' solutions for reusable packaging and fibre-based packaging and background knowledge on previous trials, consumer studies, experiments in shops with such packaging. It will allow to generate 24 deliverables contributing to the reduction of plastic packaging in the food sector; e.g. guidelines for promoting behaviour change and social innovation, protocols for washing centers and food safety standards, decision-support model for reuse logistic network design, decision-matrix based on materials, barrier combination and application performance evaluation, process methodology for producing high performance fibre-based, high-barrier packaging that can be upscaled for commercialization and life cycle analysis of the defined reuse model and fibre-based packaging.

The data will be both qualitative and quantitative, as the first part of the project will be dedicated to innovating (selecting, producing, iterating, and reporting) and the second part of the project to test in real-life conditions the determined solutions. The results will consist in text documents e.g., lab reports, recommendations, indicators, standards (word, pdf) and numeric documents e.g., technical reports, result tables, simulation setups (excel).

What is the purpose of the data generation and its relation to the objectives of the project?

The purpose of the data generation is to secure fast and extensive uptake of industrially relevant, cross-sectorial, cost-effective innovative technologies allowing immediate substitution of complex multilayer plastic packaging with high performing fiber-based packaging and economical, industrial, and environmental optimisation of reuse schemes demonstrated at large scale and transnationally in 2 EU countries by 2 major retailers, covering the needs of 13 different food product types.

Documenting the R&D and Demonstration process allows to close an information gap within the field of research for alternative packaging. The generated results, once openly published, can widely contribute to accelerate the food sectors transition towards zero single-use plastic, and by extension speed up the uptake process for other sectors (cosmetics, furniture, pharmaceuticals).

Funded by the European Union

This project has received funding from the European Union's Horizon Europe Research and Innovation programme under Grant Agreement No 101060806. This document reflects the views of the author(s) and does not necessarily reflect the views or policy of the European Commission. Whilst efforts have been made to ensure the accuracy and completeness of this document, the European Commission shall not be liable for any errors or omissions, however

Ultimately, the data by R3PACK serves the purpose of a European-wide adoption of sustainable fibre-based solutions and viable reuse schemes to generate a greater impact about reducing over-packaging, single-use plastics, and related microplastic pollution.

What is the expected size of the data that you intend to generate or re-use?

The expected size of data that will be generated or re-used is in a range of a few hundreds of megabytes to a few gigabytes.

What is the origin/provenance of the data, either generated or re-used?

The re-used data stems from each partner's individual background listed in annex of the Consortium Agreement. Depending on the nature of the Beneficiary, for instance Industrial or Scientist, the data can originate from scientifical research, literature reviews and laboratory tests.

Similarly, the data generated will be a product of the R&D and Demonstration phase: modelling tools, lab experiments and qualitative analysis. For instance, packaging manufacturers will perform machinability tests, experts will carry out resistance to barriers (O2, H20, …) tests with tools generating technical data sheets and lab reports.

All partners participate in the generation of data through the formalisation of deliverables presenting the results of R3PACK.

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

The data is primarily targeted towards:

- Food retailers,
- Food industrials,
- Packaging manufacturers

to foster the adoption of new fibre-based packaging solutions and optimised reuse model across the European Union, with the objective to reduce the use of single-use plastic packaging. The solutions might also be of interest to regulators, to disseminate and legally support the results, to researchers in the field of packaging, to certification offices, and the greater public (consumers).

2. FAIR DATA

As described by OpenAire in their guidelines for making data FAIR, "the increasing availability of online resources means that data need to be created with longevity in mind. Providing other researchers with access to data facilitates knowledge discovery and improves research transparency. [...] The FAIR principles describe how research outputs should be organized so they can

Funded by the European Union

This project has received funding from the European Union's Horizon Europe Research and Innovation programme under Grant Agreement No 101060806. This document reflects the views of the author(s) and does not necessarily reflect the views or policy of the European Commission. Whilst efforts have been made to ensure the accuracy and completeness of this document, the European Commission shall not be liable for any errors or omissions, however

be more easily accessed, understood, exchanged and reused." For that purpose produced data will be deposited in the trusted European repository Zenodo, hosted by CERN, an intergovernmental organization, as will be further explained in part 2.2.1 Repository Management.

2.1 MAKING DATA FINDABLE, INCLUDING PROVISIONS FOR METADATA

Findability criteria encompasses following key questions:

2.1.1.1. Will **rich metadata** be provided to allow discovery? What metadata will be created? What disciplinary or general standards will be followed?
2.1.1.2. Will **search keywords** be provided in the metadata to optimize the possibility for discovery and then potential re-use?
2.1.1.3. Will metadata be offered in such a way that it can be **harvested** and **indexed**?

Zenodo responds positively to all requirements by implementing following actions as described on their website *(https://about.zenodo.org/principles/)*:

- "Data are described with rich metadata (*defined by R1 section d.*)
  - Zenodo's metadata is compliant with DataCite's Metadata Schema minimum and recommended terms, with a few additional enrichements.

- Metadata clearly and explicitly include the identifier of the data it describes
  - The DOI is a top-level and a mandatory field in the metadata of each record.

- (Meta)data are registered or indexed in a searchable resource
  - Metadata of each record is indexed and searchable directly in Zenodo's search engine immediately after publishing.
    Metadata of each record is sent to DataCite servers during DOI registration and indexed there."

2.2 MAKING DATA ACCESSIBLE

2.2.1 Repository management

Will the data be deposited in a trusted repository?

The data will be deposited in the trusted European repository Zenodo, hosted by CERN, an intergovernmental organisation, founded in 1954. The repository is partly funded by the European Commission via the OpenAIRE projects (e.g. FP7: OpenAIRE (246686); Horizon 2020: OpenAIRE2020 (643410), ...)

Funded by the European Union

This project has received funding from the European Union's Horizon Europe Research and Innovation programme under Grant Agreement No 101060806. This document reflects the views of the author(s) and does not necessarily reflect the views or policy of the European Commission. Whilst efforts have been made to ensure the accuracy and completeness of this document, the European Commission shall not be liable for any errors or omissions, however

Cern's servers are stored in Europe, ensuring the respect of European laws, for instance with regards to GDPR.

Zenodo hosts the community "R3PACK", which allows to safely share and store the consortium's data and manage its visibility.

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

Zenodo is a user-friendly, reliable, and trustworthy platform based on the Open Science principles for storing, publishing and sharing research data. The repository is in line with the FAIR principles definition as referenced from: *Wilkinson, M. D. et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci. Data 3:160018 doi: 10.1038/sdata.2016.18 (2016).*

The reason to use this repository are the following:

- it allows researchers to deposit both scientific publications and any other type of data;
- data related to R3PACK can immediately be identified as such through the dedicated Community created in ZENODO, so all beneficiaries of the project can link the uploaded documents to the Community;
- the repository has backup and archiving capabilities;
- ZENODO assigns all publicly available uploads a unique Digital Object Identifier (DOI) for citation;
- the repository allows different access rights (open, embargoed, restricted, closed)

Will all data be made openly available?

As per the Grant Agreement, open peer review of R3PACK scientific publications will be made fully accessible. In addition, for the purpose of European-wide systemic change, R3PACK has agreed to grant access to the results foundations including:

- Main hypotheses of experimental work conducted in WP2, WP3, WP4 and primary statistical analyses will be pre-registered.

- Data and analysis code will be made publicly accessible after publication.

Full open access will be provided to 23 out of 25 of R3PACK's deliverables. All these deliverables will be fully accessible to enable a broader impact of the project's results. The only restricted deliverables will be D2.1 and D4.1. They will be registered as sensitive, and their access will be limited under the conditions of the Grant Agreement. The data will made available as quick as possible after its generation.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe Research and Innovation programme under Grant Agreement No 101060806. This document reflects the views of the author(s) and does not necessarily reflect the views or policy of the European Commission. Whilst efforts have been made to ensure the accuracy and completeness of this document, the European Commission shall not be liable for any errors or omissions, however

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

Two deliverables, D2.1 and D4.1, are confidential. They will not be openly published and will have a restricted access within the consortium.

Deliverable 2.1 consists in collecting key data from the retailers, industrials, and packaging manufacturers of the consortium, which will provide an overview of R3PACK's starting point in terms of packaging, but also contribute widely to other tasks. The collected data is highly sensitive as it contains trade secrets, which require full confidentially regarding the public and a high level of restriction for members of the consortium. Several partners of R3PACK are competitors, the leak of such information could break their trust and severely harm anti-competition law. Therefore, access to the deliverable 2.1 and related data will be limited to the partners which are not direct competitors, do not pose any conflict of interest and that do require the information for the purpose of their work for R3PACK. As coordinator, (RE)SET will be the intermediary between partners to ensure secrecy. (RE)SET has evaluated the need for information for each work package and defined accordingly the few authorized recipients (Unibo, Innovhub, Rise, …). The coordinator then collects the data directly from the concerned parties, centralizes it and protect the files with a secret code, before distributing it to the targeted people. The process avoids direct communication between the partners and thus, any breach of confidentiality.

The second deliverable, D4.1 is linked to the R&D activities of the project which are entitled to Intellectual Property due to their breakthrough nature for society. Produced data concerns alternative solutions for single-use plastic packaging, a paramount challenge not only for EU, but worldwide. At the moment, it is impossible to evaluate what will be subject to a patent, as the deliverable is only due month 18 (November 2023). Partners at the origin of the data, will be the owners and decide whether to license their technology. The DMP will be updated once a decision can be reached regarding the produced data within this deliverable. In any case, the end-results (excluding the data outputs produced to reach them) will be public under the form of a decision matrix (D4.2), allowing their wider use among industrials, retailers and other interested stakeholders. The temporary confidentiality of D4.1 will not prevent external actors to benefit from its results.

2.2.2 Accessibility requirements

Accessibility criteria encompasses following key questions:

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

2. Will the data be accessible through a **free and standardized access protocol?**

3. How will the **identity** of the person accessing the data be ascertained?

Funded by the European Union

This project has received funding from the European Union's Horizon Europe Research and Innovation programme under Grant Agreement No 101060806. This document reflects the views of the author(s) and does not necessarily reflect the views or policy of the European Commission. Whilst efforts have been made to ensure the accuracy and completeness of this document, the European Commission shall not be liable for any errors or omissions, however

4. Will metadata be made **openly available and licenced** under a public domain dedication CC0, as per the Grant Agreement?

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

Zenodo responds positively to all requirements by implementing following action as described on their website *(https://about.zenodo.org/principles/)*:

1. **"F1**: (meta)data are assigned a globally unique and persistent identifier
    - A DOI is issued to every published record on Zenodo.

2. (Meta)data are retrievable by their identifier using a standardized communications protocol
    - Metadata for individual records as well as record collections are harvestable using the OAI-PMH protocol by the record identifier and the collection name.
    - Metadata is also retrievable through the public REST API.
    - The protocol is open, free, and universally implementable

3. The protocol allows for an authentication and authorization procedure, where necessary
    - Metadata are publicly accessible and licensed under public domain. No authorization is ever necessary to retrieve it.

4. All metadata in Zenodo may be freely used under the CC0 waiver.
    - Zenodo registers DOIs for all uploads.
    - Zenodo allows including information on alternate persistent identifiers, as well as linking to related persistent identifiers.

5. Metadata are accessible, even when the data are no longer available
    - Data and metadata will be retained for the lifetime of the repository. This is currently the lifetime of the host laboratory CERN, which currently has an experimental programme defined for the next 20 years at least.

Metadata are stored in high-availability database servers at CERN, which are separate to the data itself.**"**

2.3 MAKING DATA INTEROPERABLE

Interoperability criteria encompasses following key questions:

1. 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? Will you follow community-endorsed interoperability best practices? Which ones?

Funded by the European Union

This project has received funding from the European Union's Horizon Europe Research and Innovation programme under Grant Agreement No 101060806. This document reflects the views of the author(s) and does not necessarily reflect the views or policy of the European Commission. Whilst efforts have been made to ensure the accuracy and completeness of this document, the European Commission shall not be liable for any errors or omissions, however

2. Will your data include **qualified references** to other data (e.g. other data from your project, or datasets from previous research)?

Zenodo responds positively to all requirements by implementing following actions:

1. **"**(Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
    - Zenodo uses JSON Schema as internal representation of metadata and offers export to other popular formats such as Dublin Core or MARCXML.

   (Meta)data use vocabularies that follow FAIR principles
    - For certain terms we refer to open, external vocabularies, e.g.: license (Open Definition), funders (FundRef) and grants (OpenAIRE).

2. (Meta)data include qualified references to other (meta)data
    - Each referrenced external piece of metadata is qualified by a resolvable URL." https://about.zenodo.org/principles/

2.4 INCREASE DATA RE-USE

2.4.1 Accessibility to reusable data

Will your data be made freely available in the public domain to permit the widest re-use possible?

Refer to section **b.i. Repository Management.**

Will the data produced in the project be useable by third parties, in particular after the end of the project?

The non-sensitive and non-confidential data (cf. deliverable 2.1 and 4.1) produced in the project will be useable by third parties during and after the end of the project.

How will you provide documentation needed to validate data analysis and facilitate data re-use?

The coordinator will set up a publication policy for the Partners when publishing on the designated repository Zenodo, that shall host all open access data of

Funded by the European Union

This project has received funding from the European Union's Horizon Europe Research and Innovation programme under Grant Agreement No 101060806. This document reflects the views of the author(s) and does not necessarily reflect the views or policy of the European Commission. Whilst efforts have been made to ensure the accuracy and completeness of this document, the European Commission shall not be liable for any errors or omissions, however

R3PACK, or other channels. It will require from the party to provide at least following information about the datasets for them to be identified properly:

- description of data,
- date of deposit,
- author(s),
- Horizon Europe funding,
- grant project name, acronym, and number,
- licensing terms (if applicable),
- persistent identifiers for the dataset,

In addition, to above mentioned criteria, the data type will be systematically indicated to facilitate evaluation and reuse of the data (unit of measurement, variable definitions, …)

2.4.2 Quality Assurance

A process for dissemination of the scientific publications, results, research outputs and any other data has been established to ensure the quality and of information made accessible to the public. It will make sure of the confidentiality status of the data, but also that it is presented in a comprehensive way to be easily re-usable by external parties. The process will first consist in evaluating the confidentiality of the data to prevent any leak of sensitive information, then data will be reviewed by the Coordinator to validate following criteria:

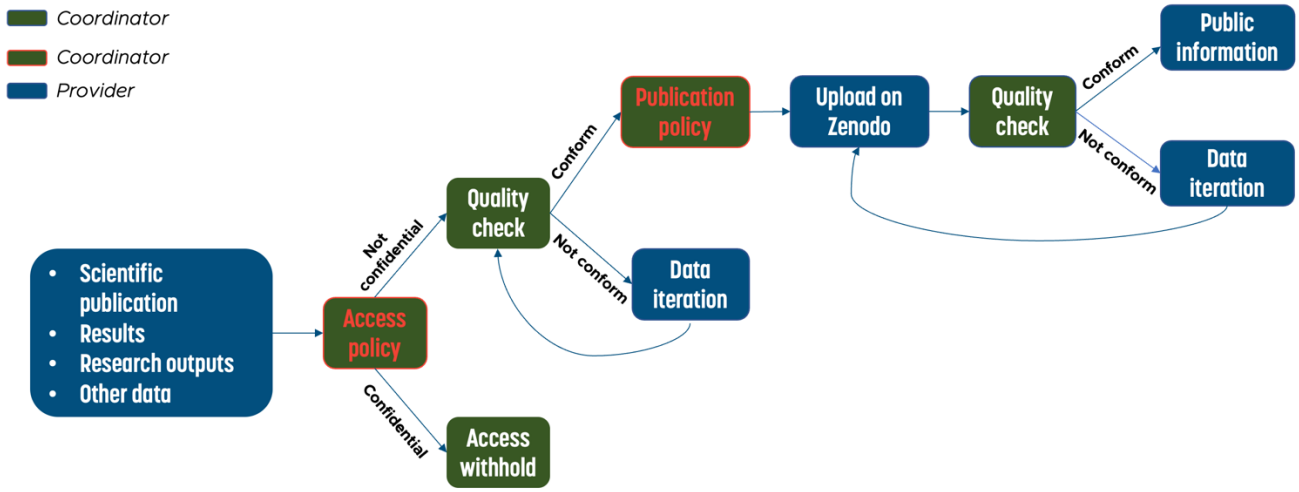
- Quality of delivered information
- Conformity with European Commission, Grant Agreement and Consortium Agreement rules
- Respect of standard layout
- Readability and understandability of information
- Inclusion of data outputs
- Inclusion of references, sources, …

In the case that partners do not meet these requirements, they will have to perform iterations on their datasets and re-submit the new version to the Coordinator until the expectations are fulfilled. Following the quality check, Partners will have to abide by the publication policy in order to publish their content. The Coordinator will perform a last quality check to ensure the set rules have been followed, which will lead to the open access publication of the data.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe Research and Innovation programme under Grant Agreement No 101060806. This document reflects the views of the author(s) and does not necessarily reflect the views or policy of the European Commission. Whilst efforts have been made to ensure the accuracy and completeness of this document, the European Commission shall not be liable for any errors or omissions, however

Figure.1 R3PACK Quality Assurance Process



2.4.3 Other reusability requirements

Reusability criteria encompasses following key questions:

1. Will your data be **licensed** using standard reuse licenses, in line with the obligations set out in the Grant Agreement?

To date the data that could be licensed is not yet produced by work package 4. When the possibility arises, the licensed data will follow our commitments to the European Commission and the DMP will be updated accordingly.

2. Will the **provenance** of the data be thoroughly documented using the appropriate standards?

Zenodo responds positively to the latter requirement by implementing following actions as described on their website *(https://about.zenodo.org/principles/)*:

- **"**(Meta)data are richly described with a plurality of accurate and relevant attributes
    - Each record contains a minimum of DataCite's mandatory terms, with optionally additional DataCite recommended terms and Zenodo's enrichments.

- (Meta)data are released with a clear and accessible data usage license
    - License is one of the mandatory terms in Zenodo's metadata and is referring to an Open Definition license.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe Research and Innovation programme under Grant Agreement No 101060806. This document reflects the views of the author(s) and does not necessarily reflect the views or policy of the European Commission. Whilst efforts have been made to ensure the accuracy and completeness of this document, the European Commission shall not be liable for any errors or omissions, however

- Data downloaded by the users is subject to the license specified in the metadata by the uploader.

- (Meta)data are associated with detailed provenance
  - All data and metadata uploaded is traceable to a registered Zenodo user.
  - Metadata can optionally describe the original authors of the published work.

- (Meta)data meet domain-relevant community standards
  Zenodo is not a domain-specific repository, yet through compliance with DataCite's Metadata Schema, metadata meets one of the broadest cross-domain standards available.”

3 OTHER RESEARCH OUTPUTS

The management of other research outputs that will arise from the action will be described in the Digital Output Management Plan incorporated in the Dissemination & Communication Plan, which first version will be delivered M6 of the project, as part of task 7.1.

4 ALLOCATION OF RESOURCES

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.)?

Zenodo is partly funded by the European Commission and donations from other partners. Cern hosting Zenodo cannot accept funds in exchange of a service, therefore Zenodo’s services are free of charge for R3PACK.

Who will be responsible for data management in your project?

The Coordinator of R3PACK, (RE)SET, is responsible for the data management of the project, as part of WP1 focusing on project management, task 1.3.

How will long term preservation be ensured?

The preservation of data is guaranteed for a time period of 20 years minimum according to Cern (see section **b.ii. Accessibility requirements)**

The coordinator together with the partners will decide collectively and monitor the accessibility of data over time. For instance, after a minimum time of exploitation, confidential results or data could be re-evaluated to make them publicly accessible. If costs arise, they will be discussed between the Coordinator and Partners to decide whether costs should split evenly.

Funded by the European Union

This project has received funding from the European Union’s Horizon Europe Research and Innovation programme under Grant Agreement No 101060806. This document reflects the views of the author(s) and does not necessarily reflect the views or policy of the European Commission. Whilst efforts have been made to ensure the accuracy and completeness of this document, the European Commission shall not be liable for any errors or omissions, however

5 DATA SECURITY

What provisions are or will be in place for data security?

Cern and Zenodo ensure maximal security through following protocols applied to the data centres, services, networks, and staff as described on their website *(https://about.zenodo.org/infrastructure):*

- **"CERN Data Centre**: Our data centres is located on CERN premises and all physical access is restricted to a limited number of staff with appropriate training and who have been granted access in line with their professional duties (e.g. Zenedo staff do not have physical access to the CERN Data Centre)

- **Servers**: Our servers are managed according to the CERN Security Baseline for Servers, meaning e.g. remote access to our servers are restricted to Zenodo staff with appropriate training, and the operating system and installed applications are kept updated with latest security patches via our automatic configuration management system Puppet.

- **Network**: CERN Security Team runs both host and network-based intrusion detection systems and monitors the traffic flow, pattern and contents into and out of CERN networks in order to detect attacks. All access to zenodo.org happens over HTTPS, except for static documentation pages which are hosted on GitHub Pages

- **Data**: Zenodo stores user passwords using strong cryptographic password hashing algorithms (currently PBKDF2+SHA512). Users' access tokens to GitHub and ORCID are stored encrypted and can only be decrypted with the application's secret key.

- **Application**: We are employing a suite of techniques to protect your session from being stolen by an attacker when you are logged in and run vulnerability scans against the application.

- **Staff**: CERN staff with access to user data operate under CERN Operational Circular no. 5, meaning among other things that staff should not exchange among themselves information acquired unless it is expressly required for the execution of their duties access to user data must always be consistent with the professional duties and only permitted for resolution of problems, detection of security issues, monitoring of resources and similar. Staff are liable for damage resulting from any infringement and can have access withdrawn and/or be subject to disciplinary or legal proceedings depending on seriousness of the infringement.**"**

Zenodo also covers data storage, ensuring data is safely store, archived and can be recovered in case an issue arises. The longevity of data related to R3PACK is therefore assured as described on their website

Funded by the European Union

This project has received funding from the European Union's Horizon Europe Research and Innovation programme under Grant Agreement No 101060806. This document reflects the views of the author(s) and does not necessarily reflect the views or policy of the European Commission. Whilst efforts have been made to ensure the accuracy and completeness of this document, the European Commission shall not be liable for any errors or omissions, however

(https://about.zenodo.org/infrastructure):

"All files uploaded to Zenodo are 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 we store two independent MD5 checksums. One checksum is stored by Invenio and used to detect changes to files made from outside of Invenio. The other checksum is stored by EOS and used for automatic detection and recovery of file corruption on disks.

Zenodo may, depending on access patterns in the future, move the archival and/or the online copy to CERN's offline long-term tape storage system CASTOR in order to minimize long-term storage costs.

EOS is the primary low latency storage infrastructure for physics data from the Large Hadron Collider (LHC) and CERN currently operates multiple instances totalling 150+ petabytes of data with expected growth rates of 30-50 petabytes per year. CERN's CASTOR system currently manages 100+ petabytes of LHC data which are regularly checked for data corruption.

Invenio provides an object store like file management layer on top of EOS which is in charge of e.g. version changes to files.**"**

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

Data will be safely stored in the world's largest general-purpose research repository, Zenodo hosted by Cern. As mentioned in section **b.ii. Accessibility requirements:**

(Me)tadata are accessible, even when the data are no longer available. Data and metadata will be retained for the lifetime of the repository. This is currently the lifetime of the host laboratory CERN, which currently has an experimental programme defined for the next 20 years at least.

6 ETHICS

Are there, or could there be, any ethics or legal issues that can have an impact on data sharing?

No ethical risk has been identified related to R3PACK's scope of action. The coordinator will look out for any throughout the lifetime of the project to avoid ethical or legal issues.

The project will be carried out in compliance with Article 14, Annex 5 of the Grant Agreement:

- ethical principles (including the highest standards of research integrity)
- applicable EU, international and national law, including the EU Charter of

Funded by the European Union

This project has received funding from the European Union's Horizon Europe Research and Innovation programme under Grant Agreement No 101060806. This document reflects the views of the author(s) and does not necessarily reflect the views or policy of the European Commission. Whilst efforts have been made to ensure the accuracy and completeness of this document, the European Commission shall not be liable for any errors or omissions, however

Fundamental Rights and the European Convention for the Protection of Human Rights and Fundamental Freedoms and its Supplementary Protocols.

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

Informed consent for data sharing and long-term preservation, following GDPR rules, will be included in questionnaires dealing with personal data.

REFERENCES

OpenAire website: https://www.openaire.eu/how-to-make-your-data-fair

European Commission H2020 manual: https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

Zenodo website: https://about.zenodo.org/principles/ ; https://about.zenodo.org/infrastructure/

Funded by the European Union

This project has received funding from the European Union's Horizon Europe Research and Innovation programme under Grant Agreement No 101060806. This document reflects the views of the author(s) and does not necessarily reflect the views or policy of the European Commission. Whilst efforts have been made to ensure the accuracy and completeness of this document, the European Commission shall not be liable for any errors or omissions, however