



Rural Environmental Monitoring via ultra wide-ARea networkS And
distriButed federated LEarning

Data Management Plan

Document Type	Deliverable
Document Number	D6.2
Primary Author	PDM
Dissemination Level	PU
Project Acronym	REMARKABLE
Project Title	Rural Environmental Monitoring via ultra wide-ARea networkS And distriButed federated LEarning
Grant Agreement Number	101086387
Project Website	https://remarkable.ulusofona.pt/
Project Coordinator	PDM
Version	1.00 Final

REMARKABLE received funding from the European Research Executive Agency (REA) under grant agreement No 101086387 under European Union's HORIZON Unit Grant research and innovation programme.



Authoring & Approval

Authors of the document

Name/Beneficiary	Position/Title	Date
Luis Miguel Campos / PDM	Project coordinator	14-04-2023
Paolo Calciati / PDM	Project coordinator	14-03-2024

Reviewers internal to the project

Name/Beneficiary	Position/Title	Date
João Carvalho / ULHT	Professor	18-03-2024
Nuno Pedrosa / PDM	Senior Software Architect	15-03-2024

Approved for submission to the REA by - representatives of beneficiaries involved in the project

Name/Beneficiary	Position/Title	Date
Luis Miguel Campos / PDM	Project coordinator	20-03-2024

Rejected by - representatives of beneficiaries involved in the project

Name/Beneficiary	Position/Title	Date

Document History

Version	Date	Status	Author	Justification
0.1	14/04/2023	Draft	Luis Miguel Campos	First draft
0.2	30/10/2023	Draft	Luis Miguel Campos	Removal of Zoetis and addition of Ionian University in the consortium
0.3	10/03/2024	Draft	Paolo Calciati	Updated content in Sections 3, 4 and 5
1.0	19/03/2024	Final	Paolo Calciati	Minor fixes. Final release

Copyright Statement

©2023 REMARKABLE Consortium. All rights reserved.

Abstract

This document presents the Data Management Plan (DMP) for the REMARKABLE project, outlining strategies for the collection, organization, storage, and sharing of project data and knowledge. The formulation of this plan draws upon various sources, including the REMARKABLE Grant Agreement, guidelines provided by the European Commission for data management in H2020/Horizon Europe projects, and insights from consortium members.

Included within this DMP are methodologies, tools, and repositories deemed appropriate for managing and disseminating project-generated information. This encompasses a wide array of data types, such as scientific publications from consortium members, Open-Source software components, user statistics, project deliverables, and data collected from European Higher Education Institutions alliances.

A significant emphasis is placed on ensuring compliance with Horizon Europe guidelines concerning Open Research Data. To this end, a descriptive template, adhering to the FAIR principles, has been established for all collected datasets, with provisions for regular updates as needed.

The selected publishing repositories must provide public accessibility and offer secure and reliable data storage environments. Additionally, preference is given to repositories well-recognized within their respective fields to enhance the dissemination goals of the REMARKABLE project.

It's important to note that this DMP is a dynamic document, subject to continuous refinement based on the evolving needs of the project's research and development objectives, and informed by ongoing input from consortium members and European Higher Education Institutions.

Table of Contents

Abstract	2
Table of Contents	3
1. Introduction.....	5
1.1. Applicable Reference Material.....	5
1.2. List of Acronyms.....	5
2. Project Introduction	7
2.1. The REMARKABLE project	7
2.2. Project key messages	8
2.3. Keywords.....	9
3. Methodological Framework for Data Management Plan.....	10
3.1. Data Management Process	13
3.2. Data Security and GDPR Compliance	14
3.3. Data Related to Scientific Publications	14
3.4. REMARKABLE Public Material	15
4. Archiving Data and Preserving Infrastructure	16
4.1. SharePoint / Microsoft 365 Platform	16
4.2. Project Website.....	16
4.3. Data and Document Repositories	16
Zenodo	17
ResearchGate	17
OpenAIRE.....	17
4.4. Code Repository.....	18
5. Datasets and Publications for DMP	20
5.1. Project Public Deliverables.....	20
5.2. Project Software Components	22
5.3. Project Research Datasets	23
5.4. Project Publications	24
Scientific Publications.....	24
Other Publications.....	25
6. Summary and Conclusions	27

List of Tables

Table 1. List of acronyms.....	6
Table 2. Summary of REMARKABLE project objectives.....	8
Table 3. The Data Management Process approach for each dataset.....	14
Table 4. REMARKABLE deliverables.....	20
Table 5. Description of REMARKABLE's public deliverables.....	22
Table 6. Description of REMARKABLE's software component.....	23
Table 7. Description of REMARKABLE's dataset.....	24
Table 8. Description of REMARKABLE's publication.....	25

List of Figures

Figure 1. Granting open access diagram.....	12
---	----

1. Introduction

This document presents the REMARKABLE's Data Management Plan, which is directly connected with the work performed in WP6. The contractual agreements for dissemination are provided in the Grant Agreement, Section 1.2, and Annex 5, Articles 16 and 17 [1].

Section 1.2 of the Description of the Action (DoA) states that data that will be generated during the project will follow FAIR principles and that the DMP will comply with the Guidelines on Data Management in Horizon Europe projects.

Data sets to be collected in the project will contain only technical system parameters; none of personal nor other sensitive information will be collected during experimentation and demonstration sessions. Nonetheless, DMP will comply with Directive 95/46/EC of the European Parliament and of the Council of 24/10/95 on the protection of individuals with respect to processing of personal data.

More specifically the Data Management Plan aims to:

1. Outline the responsibilities for data protection and sharing within an ethical and legal framework;
2. Ensure the protection of the intellectual property created by the project;
3. Support open access to the project's research outcomes and scientific publications;
4. Support the openness of data related to both the publications and the development processes of the project;
5. Define a documentation framework for the annotation of the collected knowledge towards increased discoverability and validation.

This document describes the methodological framework for handling research data collected or produced during the project (Section 3), the platforms and repositories chosen to archiving those data (Section 4), and the FAIR templates to describe data generated throughout the lifespan of the project (Section 5). Finally, Section 6 contains the conclusions of the document.

1.1. Applicable Reference Material

[1] Grant Agreement Number: 101086387 — REMARKABLE — HORIZON-MSCA-2021-SE-01.

1.2. List of Acronyms

Acronym	Definition
AI	Artificial Intelligence
DMP	Data Management Plan
DoA	Description of the Action
EC	European Commission
EU	European Union
EWP	Erasmus Without Paper
FAIR	Findable, Accessible, Interoperable, Re-usable
FL	Federated Learning
GA	Grant Agreement

GDP	Gross Domestic Product
GDPR	General Data Protection Regulation
ICT	Information and Communication Technologies
IoT	Internet of Things
ML	Machine Learning
ORD	Open Research Data
TDR	Trustworthy Digital Repository

Table 1. List of acronyms.

2. Project Introduction

2.1. The REMARKABLE project

Internet of Things (IoT) technology combined with complementary support for data analytics is the cornerstone of today's digital transformation. The societal and economic impact of IoT/machine learning (ML) systems in urban and suburban areas significantly outpaces the one in rural areas due to a limited reach of connectivity infrastructure. IoT technologies have a huge potential for improving the economy and quality of life in rural areas, both in developed and developing countries. For instance, about 30.6% of the EU population lives in rural areas, which cover over 83% of the total EU area. Nonetheless, the average GDP per capita in rural EU areas is only 75% of the EU average [5]. Moreover, even though mobile networks cover more than 99% of the population in some European countries (such as the UK), they cover only about 79% of their landmass, thus leaving more than 20% of deep rural country areas without signal coverage [6]. To reverse further widening of the urban-rural gap, one needs to bring efficient and affordable IoT/ML solutions to deep rural areas, reaching out to applications and use cases ranging from wildlife management, rural tourism, livestock monitoring, water and air pollution control, and others.

REMARKABLE is an interdisciplinary project comprising experts from computer science, communication engineering, life sciences, environment and management. These experts come from diverse organisations in the UK, Europe, and Africa. The project's vision is to bring IoT/ML systems a step closer to seamless, energy efficient and secure deployment targeting use cases in deep rural areas. This will be done by identifying main gaps in connectivity and affordable data analytics and through interleaved research, development and validation in a real-world setting. The project is centred on an IoT/ML-based technological platform that will be adapted and demonstrated in the context of use cases applied in environmental monitoring, management, and conservation.

In short, the REMARKABLE project emphasises a necessity of bringing rural areas into the reach of IoT/ML technologies. Its ultimate goal is to facilitate a reduction of urban-rural gap which is currently increasing. Making advanced information and communication technologies (ICT) such as IoT/ML systems a rural commodity will play a crucial role in reversing the rural depopulation trends due to an expanded range of economic opportunities through empowering and modernising traditional rural ecosystems. The added value of deploying IoT/ML in deep rural areas is in reaching out to new streams of data sources that could prove invaluable in tackling and better understanding the growing environmental concerns, ranging from local and regional (such as pollution monitoring) to global ones (such as climate change).

REMARKABLE considers several objectives in terms of research and innovation that will also have a large environmental and societal impact, summarised in Table 2.

	Objective	Outcome	WP-M	Deliverable	Respons. Partner
Research and Innovation Objectives	Secure and Trustworthy Sensing, Localisation and Digital Twins	Design of robust, secure, trustworthy, and traceable IoT platform suitable for deep rural applications	M44	D.1.1-D.1.4	ULHT
	Connecting the Unconnected – Ultra Wide-Area IoT Networks	Provide a solution for connectivity of IoT devices deployed in deep rural areas beyond the reach of current wireless cellular network infrastructure	M44	D.2.1-D.2.4	UNS-FTN
	Secure and Frugal Distributed Data Analytics for Rural IoT	Develop a novel data analytics platform based on privacy-preserving distributed ML methods that are frugal secure and scalable	M44	D.3.1-D.3.4	MMU
	Demonstration, Validation and Assessment	Demonstrate, validate, and assess developed solutions in uses cases in real-life conditions, across European and African countries	M48	D.4.1-D.4.4	UA
Environmental and Societal Impact	Health and vitality monitoring of livestock in real-time	Enable quick animal treatment and prevent spreading illness, increase food production, track animals, identify grazing patterns, prevent desertification			
	Wildlife monitoring	Support tracking of endangered animals, reduce their poaching, improve tourist experiences in wildlife parks and reservations			
	Soil and agronomic management	Support automated irrigation and increase all-season production of food products			
	River pollution and air quality monitoring	Prevent health risk to humans, protect aquatic ecosystems from collapse and prevent the proliferation of phytoplankton			

Table 2. Summary of REMARKABLE project objectives.

2.2. Project key messages

REMARKABLE offers at least four main high-level messages that are foreseen for the principal findings produced by the projects, which are focused on the following concepts:

- The REMARKABLE project is centred on developing an IoT/ML-based technological platform that will be adapted and demonstrated in the context of use cases applied to environmental monitoring, management, and conservation.

- REMARKABLE uses and assesses innovative methodologies based on statistical data processing and decentralised federated learning methods specifically designed for different use case implementations and demonstrations.
- The REMARKABLE project places a specific focus on rural environments and, in particular, on the African continent due to the huge potential of the number of IoT applications in Africa and the lack of traditional connectivity options.
- REMARKABLE strives at developing various added-value services ranging from wildlife management, rural tourism, livestock monitoring, water and air pollution control and others.

2.3. Keywords

Distributed federated learning, Internet of Things, Rural environmental monitoring, Statistical data processing, Ultra-wide area networks.

3. Methodological Framework for Data Management Plan

The role of a Data Management Plan (DMP) is to establish a framework for managing research data generated or acquired during the project's lifespan, as well as after its completion. This framework encompasses various aspects, including the nature of the data, data collection strategies, intended beneficiaries of the data, metadata usage for enhanced retrievability, standardization, open-access considerations, and data storage and preservation protocols.

In alignment with the Open Research Data Pilot, the REMARKABLE DMP adheres to the H2020 guidelines for ensuring data is Findable, Accessible, Interoperable, and Reusable (FAIR). To achieve this, the DMP follows the FAIR template provided by the European Commission.

The components included in FAIR are the following:

- Data summary;
- FAIR data principles
 - Making data findable, including provisions for metadata;
 - Making data openly accessible;
 - Making data interoperable;
 - Increase data re-use (through clarifying licences);
- Allocation of resources;
- Data security;
- Ethical aspects;
- Other issues refer to the national/ funder / sectorial / departmental procedures for data management that is being used (if any).

The Data Summary and the FAIR data principles will be addressed separately for each dataset that is expected to be generated during the project. However, given the nature of REMARKABLE, the datasets we plan to collect and process are mostly related to information collected during the experimental phase.

Allocation of material resources for storage and archiving is not foreseen since we will use the resources already available for the existing joint degrees where the consortium members participate.

As regards publication and other open access costs, they will be provided under the project¹.

By default, Horizon Europe projects and ERASMUS participate in the Open Research Data Pilot and they must deposit the following data in a research data repository:

1. All data needed to validate the results presented in scientific publications, including the metadata that describe the research data deposited. This is called the “underlying data”. These data must be deposited as soon as possible.
2. Any other data (for instance curated data not directly attributable to a publication, or raw data), including the associated metadata, as specified and within the deadlines laid down in the DMP – that is, according to the individual judgement by each project/grantee.
3. Projects should also provide information via the chosen repository about the tools that are needed to validate the results, e.g. specialised software or software code, algorithms, and analysis

¹ http://ec.europa.eu/research/participants/data/ref/h2020/other/hi/oa-pilot/h2020-hi-erc-oa-guide_en.pdf

protocols. Where possible, they should provide these instruments themselves, or alternatively, provide direct access to them.

In the Guidelines on FAIR Data Management in Horizon 2020, the European Commission states: “Where will the data and associated metadata, documentation and code be deposited? Preference should be given to certified repositories which support open access where possible.”

Researchers, information managers and other stakeholders can rely on a framework of various international certification standards for digital repositories to assess and improve the quality of their work processes and management systems. “Trustworthy Digital Repository” (TDR) is a term often used in this respect.

Beneficiaries must also provide open access, through the repository, to the bibliographic metadata that identify the deposited publication. The purpose of the bibliographic metadata requirement is to make it easier to find publications and ensure that EU funding is acknowledged. Information on EU funding must therefore be included as part of bibliographic metadata so that Horizon 2020 can be properly monitored, statistics produced, and the programme's impact assessed.

To monitor any embargo periods, the publication date and embargo period must be provided. The persistent identifier (for example a Digital Object Identifier) identifies the publication. It enables a link to be provided to an authoritative version of the publication.

Open Access is one of the main principles of Horizon Europe; by Open Access we mean the provision of free of charge online access to scientific information for any user. The beneficiaries' obligation to granting open access is differentiated between scientific publications and research data.

- Scientific publication: Publication of academic and research work, most often in the form of an article, research paper and otherwise, in scientific journals or in other forms (e.g. textbook, conference proceedings, etc.).
- Research data: This refers to the recorded factual material commonly accepted in the scientific community as necessary to validate research findings. Research data generated for the REMARKABLE project could be generated from questionnaires, algorithms, methodologies, source code, etc.

All participating projects' beneficiaries are required to ensure open access for their peer-reviewed scientific publications relating to their results, as defined in Article 29.2 of the H2020 - General MGA².

There are two routes to open access for scientific publications³:

1. Gold open access / open access publishing - the practice of immediately publishing in open access mode (in open access journals or in ‘hybrid’ journals), shifting the payment of publication costs from readers' subscriptions to author fees. These costs are usually borne by the researcher's university or research institute or the agency funding the research.
2. Green open-access / self-archiving – the practice of depositing of a published article or a final peer-reviewed manuscript in an open-access online repository (by the author or a representative).

² H2020 Multi-Beneficiary General Model Grant Agreement v5.0, available at: http://ec.europa.eu/research/participants/data/ref/h2020/mga/gga/h2020-mga-gga-multi_en.pdf

³ http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-Access-data-management/open-Access_en.htm

A 6-12-month embargo period before the data is granted open-access may be considered appropriate by some scientific publishers.

Therefore, the open access to publications process is as follows:

1. Publications are deposited in online repositories.
2. Open access route is selected.
3. Open access is granted to publications.

Note that the steps mentioned above are not strictly successive, but may occur simultaneously, depending on the selected open-access route and a possible embargo period set by the consortium.

Regarding research data for projects participating in the ORD pilot, it is obligatory to ensure open access to all data needed for result validation⁴. Whether other parts of data will be made open access, is left to the discretion of the beneficiaries, as they must ensure that the main objective of the project will not be jeopardised by the publicity. Ethical and privacy concerns raised by publication of particular data, as well as protection of Intellectual Property Rights (IPR) are also a great deterrent to granting open access. Justification for excluding particular parts of data from being open access must be included in the DMP. The open-access research data must be deposited in online repositories, available for access, mining, exploiting, processing, and disseminating, free of charge for any user, accompanied by the appropriate information — via the repository — regarding the specific tools and instruments that beneficiaries have at their disposal, considered to be necessary for validating the results. Where possible, these tools or instruments should be provided.

The procedures previously described are summarised in Figure 1.

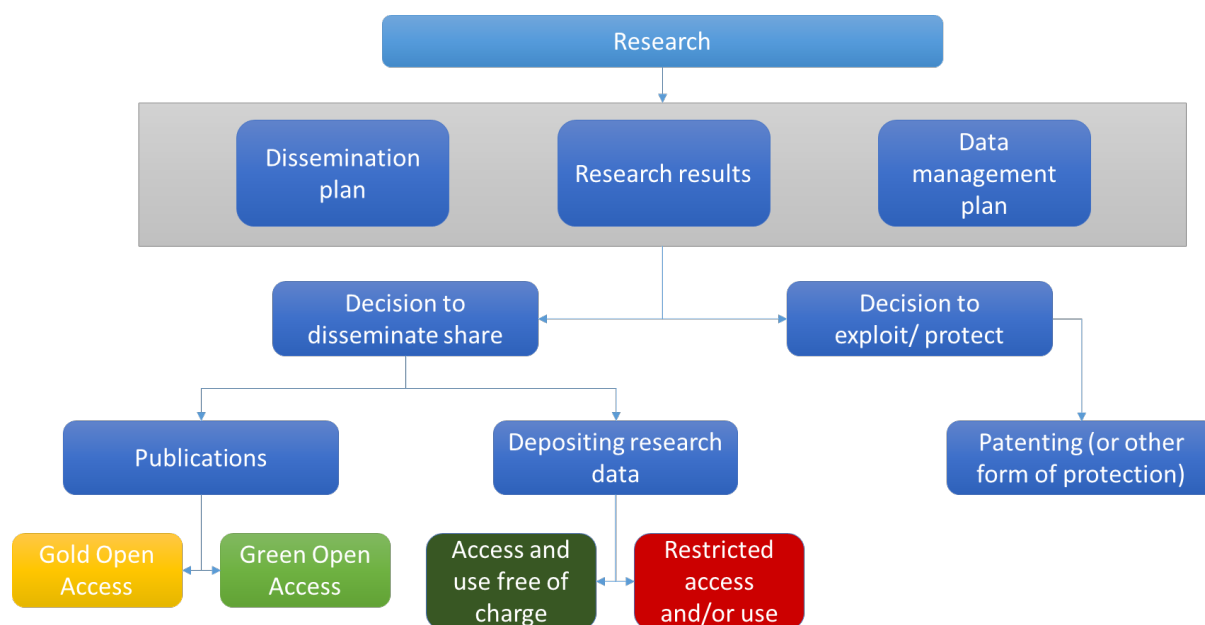


Figure 1. Granting open access diagram

⁴ Article 29.3, H2020 Multi-Beneficiary General Model Grant Agreement

Provision for the GDPR⁵, the newly enacted EU regulation about data, is also included, as described in the GDPR compliance section of the REMARKABLE DMP.

GDPR is a unified regulatory framework that took effect from May 25th, 2018, and is implemented across the European Union. Due to its regulatory nature, it is directly binding and applicable for all EU members and does not require an individual enabling state legislation. Intended to replace the 1995 EU Data Protection Directive⁶ in the light of the growing need for clarifying and safeguarding the “digital rights” of all natural persons in EU, GDPR harmonises the relevant national EU States legislations and opens up the scope to cover even companies without physical presence in the EU.

Organisations acquiring and/or processing data of natural persons are required to adopt more robust data management and security systems. At the same time GDPR empowers citizens, by enhancing monitoring and control over their own data. As stated in Article 1 of GDPR:

1. This Regulation lays down rules relating to the protection of natural persons with regard to the processing of personal data and rules relating to the free movement of personal data.
2. This Regulation protects fundamental rights and freedoms of natural persons and in particular their right to the protection of personal data.
3. The free movement of personal data within the Union shall be neither restricted nor prohibited for reasons connected with the protection of natural persons regarding the processing of personal data.

As previously noted, significant changes on data, which may arise during the project and the development of the platform, are to be reported in the form of additions to the present deliverable. As this is the only planned release of the deliverable, it will be made available to the EU commission at the end of the project with all changes it was subject to.

3.1. Data Management Process

The REMARKABLE Data Management Plan (DMP) outlines a series of steps aimed at classifying the various datasets to be collected and analysed based on the analysis presented in the previous section. Each step of the process poses a question requiring a reply, with the response defining the actual status and respective handling for each dataset generated or acquired during the project. Data storage, preservation, and sharing were deliberately excluded from the questionnaire, as all datasets produced and utilized in the REMARKABLE project will be stored, preserved, and shared through the selected platforms (MS Teams), rendering these aspects of data management common for all datasets.

The questions presented in Table 3 were selected to classify the datasets.

Issues to be addressed for dataset	Positive Answer (yes)	Negative Answer (no)
Needed for result validation?	Public	Private
Produces added value to third parties?	Public	Private

⁵ General Data Protection Regulation, available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679&from=EN>

⁶ Data Protection Directive, OJ 1995 L 281, available at: <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:31995L0046>

Can the created data - which may be derived from third-party data - be shared?	Public	Private
Contains personal data as referred to in GDPR - Article 4?	Private	Public
Contains data back traceable to private individuals?	Private	Public
Contains data that could be used in activities raising ethical issues or constitute a danger to the society?	Private	Public
Contains sensitive data or a security threat for one or more partners of the project (e.g. confidential information)?	Private	Public
Either a Licence restriction or an embargo is applied?	Private	Public
Contains data jeopardising a project patent?	Private	Public

Table 3. The Data Management Process approach for each dataset

3.2. Data Security and GDPR Compliance

Throughout the duration of the project, REMARKABLE will undertake the identification and design of data subjects and the GDPR framework. This process will involve providing an analytical description of how the project implements GDPR regulations. This will encompass detailing the data subjects involved in data collection activities, such as surveys and questionnaires, the types and sources of data (including personal data) to be collected and processed, the roles of data controllers and data processors, as well as the processing activities related to personal data. The goal is to inform the design and implementation of common digital management systems that comply with the principles of Erasmus Without Paper (EWP) and facilitate the issuance of joint diplomas.

According to Article 4 of the EU GDPR, the Data Controller, the Data Processor and the Data Recipient are defined as follows:

- Controller – “means the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data”.
- Processor – “means a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller”.
- Recipient – “means a natural or legal person, public authority, agency or another body, to which the personal data are disclosed, whether a third party or not”.

3.3. Data Related to Scientific Publications

REMARKABLE will publish scientific publications in conferences and journals as part of the planned dissemination activities. Following the EC Mandate on Open Access, REMARKABLE adheres to the Open Access policy choosing the most appropriate route for each case. Whenever possible, REMARKABLE favours the ‘green’ open access route, in which the published article or the final peer-reviewed manuscript will be deposited in an online repository, before, at the same time as, or after publication, ensuring that the embargo period requested by certain publishers has elapsed.

Scientific publications’ data are made available often using accessible PDF files. The metadata to be used will be compliant with the format requested by OpenAIRE as well as the one requested by the repository where the papers are to be deposited.

3.4. REMARKABLE Public Material

All public information and materials, including public deliverables, brochures, and posters, will be readily accessible on the project website in the form of downloadable PDF files. In instances where Intellectual Property Rights (IPR) of foreground knowledge need protection, corresponding disclosures will be promptly published.

Each deliverable will be accompanied by a set of keywords and a brief description to facilitate indexing and searchability in search engines. These keywords aim to highlight the main topics addressed in the document, whether it is a report or a software-related document.

The audience for REMARKABLE's public deliverables varies, ranging from general audiences interested in project activities to more specialized stakeholders. These deliverables provide valuable insights into the benefits of REMARKABLE, particularly through the experiences gained from existing joint European degrees.

4. Archiving Data and Preserving Infrastructure

The subsequent section incorporates concise descriptions of the platforms and repositories designated for REMARKABLE's data storage and dissemination. It outlines their structure and functionalities pertaining to open access, storage, backup, and charging policy, thus rationalizing their selection. Each platform and repository have been carefully chosen to align with the requirements outlined in the FAIR data principles and the Open Research Data (ORD) pilot.

4.1. SharePoint / Microsoft 365 Platform

SharePoint and OneDrive in Microsoft 365 are cloud-based services designed to facilitate content, knowledge, and application management within organizations. These platforms empower teamwork, streamline information retrieval, and enable seamless collaboration across the organization. They simplify the process of capturing, sharing, and retrieving information among virtual teams, enhancing productivity while reducing network bandwidth requirements and email volumes among project team members. Hosted by PDM, the platform offers a wide array of features including document locking, online preview and editing, and version control.

REMARKABLE partners utilize Microsoft 365 as the primary document repository for all files exchanged within the consortium. This includes intermediate versions of deliverables, meeting materials (such as agendas, notes, presentations, demos, and minutes), and any other documents used for gathering inputs from project partners. Access to Microsoft 365 material requires credentials, with platform usage restricted solely to the REMARKABLE consortium and, upon request, to the European Commission (EC).

4.2. Project Website

EMARKABLE's website serves as the primary online public information hub for the project and can be accessed via the URL (<https://remarkable.ulusofona.pt/>). The website features both static and dynamic textual content, including an overview of the project's concept, goals, and proposed approach, as well as real-time updates on events, news, and blog posts.

A dedicated section, titled "Published Materials," is designated for all public data related to the project and its progress. This section hosts various documents, including deliverables, research publications, press releases, and software documentation, primarily in the portable document format (PDF). Additionally, documents may be available in the Office Open XML format for ease of re-use, such as in the case of press releases. In instances where a file is deposited on a social media or data repository platform, a link to the respective source is provided, along with simple metadata information such as the title, a brief description, and the document type.

Public information on the REMARKABLE website is accessible without restrictions, allowing any visitor to access the content without the need to create an account or provide personal data. Furthermore, all webpage-related data is regularly backed up to ensure continuity and data integrity.

4.3. Data and Document Repositories

There are multiple different services which can be used for archiving project data and documents – Zenodo, ResearchGate and OpenAIRE - which are presented in the following sections.

Zenodo

Following the EC recommendation, REMARKABLE will utilize the Zenodo platform as its designated data and document repository. Zenodo is a free, open research data repository developed by OpenAIRE and CERN and launched in 2013 and is fully compliant with the open data requirements of Horizon 2020 and the EU Research and Innovation funding programme, as well as the Open Access policies of the European Union. The platform offers support for publishing scientific papers, white papers, and structured research data in various formats, including XML, and facilitates collaboration with open-source code repositories such as GitHub. All uploaded data and documentation are structured using metadata and licensed under CC license (Creative Commons 'No Rights Reserved'). Notably, the property rights or ownership of a data asset remain unchanged upon uploading it to Zenodo.

Zenodo ensures both security and availability of data. Data files are stored in CERN Data Centres, primarily in Geneva, with replicas in Budapest. Nightly backups of data files and metadata are conducted, and files are regularly checked against their checksums (using MD5 algorithm) to ensure file content integrity. In the event of repository closure, Zenodo guarantees seamless integration of all content into other suitable repositories without impacting citations and links.

Regarding REMARKABLE's data management strategy, all public results generated or collected during the project will be uploaded to Zenodo for open access, long-term storage, and dissemination. This includes public deliverables, software documentation (potentially in the case of a common digital management system), research papers, presentations, and datasets. At the project's conclusion, consideration will be given to utilizing an institutional research data repository to further disseminate the project outcomes.

Publications regarding the REMARKABLE project have already been issued on Zenodo with the following DOI: [10.5281/zenodo.10844869](https://doi.org/10.5281/zenodo.10844869).

ResearchGate

In addition to establishing Zenodo as the primary repository for data and documents, REMARKABLE will leverage ResearchGate to enhance the dissemination of scientific publications. ResearchGate, launched in 2008, serves as a vibrant networking platform for scientists and researchers, boasting over 3 million users worldwide. It offers a plethora of services including publication sharing, networking with colleagues, seeking and providing expert insights, and fostering global collaborations. By creating a ResearchGate project profile, REMARKABLE aims to provide a dedicated space for showcasing research results pertaining to the project's objectives. Partners will actively upload their research publications, enriching them with necessary metadata, and engaging with the wider scientific community by addressing inquiries and fostering discussions.

Through ResearchGate, REMARKABLE aims to amplify the visibility and impact of its research outputs while fostering meaningful interactions within the scientific community. Through strategic utilization of ResearchGate, REMARKABLE seeks to maximize the impact and reach of its research outcomes, ultimately contributing to advancements in the field and fostering a culture of open science and collaboration.

OpenAIRE

OpenAIRE, standing for Open Access Infrastructure for Research in Europe, is a robust and dynamic initiative aimed at facilitating and promoting open access to scholarly publications and research data

across Europe and beyond. Founded in 2008, OpenAIRE operates as a collaborative effort among various stakeholders in the research community, including universities, research institutions, libraries, and funding agencies. At its core, OpenAIRE serves as a comprehensive infrastructure providing a range of services to support the open access ecosystem. One of its primary functions is the aggregation and indexing of research outputs, including peer-reviewed articles, conference papers, and datasets, sourced from diverse repositories across different disciplines. By consolidating this wealth of research outputs, OpenAIRE offers researchers, policymakers, and the general public a centralized platform for discovering and accessing scholarly information freely.

Moreover, OpenAIRE plays a pivotal role in advancing open science practices by promoting transparency, reproducibility, and collaboration within the research community. Through its initiatives and partnerships, OpenAIRE actively advocates for open access policies and standards, facilitates the adoption of open access publishing models, and provides guidance and support to researchers navigating the complexities of open research practices.

In addition to its role in promoting open access to research outputs, OpenAIRE also offers valuable services for archiving and preserving project data and documents. By leveraging OpenAIRE's infrastructure, researchers and project consortia can securely deposit datasets, project reports, and related documents, ensuring their long-term accessibility and citability within the scholarly community. OpenAIRE's repositories adhere to digital preservation standards, guaranteeing the integrity and authenticity of archived data over time. Furthermore, these repositories support interoperability with other archival systems, facilitating seamless integration with existing project workflows.

Overall, OpenAIRE provides a comprehensive solution for project consortia to archive and preserve their data and documents, enhancing their visibility, accessibility, and impact while ensuring their enduring availability for future research endeavors.

4.4. Code Repository

REMARKABLE does not anticipate the development of any products; however, it will engage in software development within the fields of AI and IoT. In the event that software development occurs, certain partners have agreed that it will adhere to the Open Source paradigm. Any developed software will be deposited in a public Git-based web repository, making it available to the community for utilization and expansion.

When considering Git-based hosting providers, REMARKABLE has two primary options: GitLab and GitHub.

GitLab⁷ is a comprehensive, Git-based software development platform that seamlessly integrates a multitude of essential tools for software development, deployment, and project management:

- Code hosting in repositories with version control.
- Issue Tracker for managing new implementations, bug reports, and feedback.
- Issue Boards for organizing and prioritizing tasks.
- Code review through Merge Requests with live-preview changes per branch, facilitated by Review Apps.

⁷ <https://about.gitlab.com>

- Built-in Continuous Integration, Continuous Deployment, and Continuous Delivery support to automate building, testing, and deploying applications. GitLab's CI pipeline is triggered for each pull or push, executing a series of jobs in parallel stages (batches); if one of the jobs fail, the next stage is usually not executed. The pipeline typically consists of four stages: build, test, staging, and production. The status of current and historical pipelines is conveniently visualized in a dedicated Pipelines tab to aid users in monitoring the deployment process.
- Integration with Docker via GitLab Container Registry, enabling each project to have its own space for storing Docker images.

In addition to GitLab, the consortium is exploring the possibility of utilizing GitHub⁸ for collaborative software development of open-source components. GitHub, also a Git-based online repository hosting provider, boasts a significant market share with 24 million connected developers as of March 2018. While GitHub is primarily recognized for code storage, its platform supports various formats and features. Its services align closely with those offered by GitLab, making it a viable option to meet the project's requirements.

⁸ <https://github.com/about>

5. Datasets and Publications for DMP

In this section, we present all anticipated data to be generated throughout the lifespan of the REMARKABLE project. These data are categorized into four distinct groups: public deliverables, open-source software components, research datasets, and publications. Each category, as well as every individual dataset generated during the project, will be described using the adapted FAIR Template.

Guidelines for determining the points covered in this template are derived from FAIR principles themselves, which outline the standards for research data. This ensures that ultimately, data should be easily Findable, Accessible, Interoperable, and Re-usable.

5.1. Project Public Deliverables

Public deliverables of the REMARKABLE project are presented in Table 4.

Deliverable	Title	Due Date (month)
D1.1	Secure and trustworthy ML-based rural IoT platform architecture	18
D1.2	Digital Twins models for Rural IoT System Design	36
D1.3	ML-based security and trustworthiness for Rural IoT platforms	44
D1.4	Multi-modal secure localization methods for Rural IoT in GPS-denied environments	44
D2.1	Coverage capabilities and limitations of existing LP-WAN technologies	18
D2.2	IoT connectivity in deep coverage holes using HAPS/LEO satellites	36
D2.3	Coverage extension for rural IoT connectivity using UAVs	44
D2.4	Sustainable Ultra-Wide Area IoT Network WP2 Architecture	44
D3.1	Novel feature extraction and dimensionality reduction methods for unstructured IoT sensor data	24
D3.2	ML models for data and computation constrained edge/fog nodes	36
D3.3	A peer-to-peer federated learning algorithm for model parameter aggregation in edge/fog nodes	44
D3.4	A frugal, secure and distributed fog-based platform for smart applications	44
D4.3	Inter-Disciplinary Outputs and Pilots Demonstration	24
D4.4	Evaluation and User Acceptance Report	48
D6.1	Dissemination Exploitation and WP6 Communication Plan	6
D6.2	Data Management Plan	6

Table 4. REMARKABLE deliverables

The FAIR template has been modified, in line with the FAIR data principles, to accommodate the required details for all the public deliverables. The result is presented in Table 5, which contains a general description of the REMARKABLE Public Deliverables, addressing all relevant issues indicated originally. Management of all public deliverables generated in the project will be implemented using this table.

1. Public deliverable summary

Purpose	The purpose of the public deliverable
Relation to the objectives of the project	Relation of the public deliverable to the project objectives
Types/Formats	All public deliverables are (or are accompanied by) reports in the cross-platform portable document format (PDF).
Re-use of any existing data	Source of re-used datasets, IRP issues etc.
Origin	How the included data was generated (or mention source, if collected)
Size	Size of the public deliverable
Utility for others	To whom and how the deliverable will be useful.

2. FAIR Data

2.1. Making public deliverables findable, including provisions for metadata

Metadata provision	Metadata is added manually and includes name, author, all consortium partner organisations, relevant keywords
Metadata standards	No specific metadata standard used
Unique identifier	The public deliverables are assigned URLs by upload on the official REMARKABLE website
Naming conventions	<ul style="list-style-type: none"> Naming convention used: REMARKABLE_[Deliverable Code]-[Deliverable Title]-vA.B. Version numbering convention used: vA.B, where A is a major version of the deliverable (Submission to European Commission) and B is minor version of the deliverable for updates during the preparation phase.
Search keywords	Metadata keywords serve as search keywords
Version control	All changes reported in the document history section.

2.2 Making public deliverables openly Accessible

Classification	Confidentiality level: PU (public)
Sharing and access regimes	Before submission: available only to consortium partners through the Microsoft 365 (Sharepoint) platform After submission: publicly available through the official REMARKABLE website
Needed method/software	No special software needed for the PDF format
Repository	Microsoft 365 (Sharepoint) platform and REMARKABLE's official website
Access authorisation	Before submission: Access only to authorised consortium partners. After submission: upload on the website, no authorisation needed

2.3. Making public deliverables interoperable

Data/metadata vocabularies and other I/O standards -

Mapping to common ontologies -

2.4. Increase re-use of public deliverables (through clarifying licences)

Licence	No licence needed
Re-use availability schedule	After submission: immediately granted free open Access for mining, exploiting, processing and disseminating
Re-use by third parties	After submission: Accessible and re-usable from third parties. No access and time limitations apply
Quality assurance	Internal quality audit control by the Quality Manager and two assigned reviewers (consortium partners)
Availability period	No time limitation scheduled after the end of the project

Table 5. Description of REMARKABLE's public deliverables

5.2. Project Software Components

Software components of the REMARKABLE project are presented in this section. Each software component is reported with all related information displayed in Table 6, independently of it being an open- or closed-source component (this will be decided individually by every partner).

1. Software Component summary

Purpose

Relation to the objectives of the project

Types/Formats

Re-use of any existing data

Origin

Size

Utility for others

2. FAIR Data

2.1. Making component findable, including provisions for metadata

Metadata provision

Metadata standards

Unique identifier

Naming conventions

Search keywords

Version control

2.2 Making component openly Accessible

Classification

Sharing and access regimes

Needed method/software

Repository

Access authorisation

2.3. Making component interoperable

Data/metadata vocabularies and other I/O standards -

Mapping to common ontologies -

2.4. Increase re-use of component (through clarifying licences)

Licence

Re-use availability schedule

Re-use by third parties

Quality assurance

Availability period

Table 6. Description of REMARKABLE's software component

5.3. Project Research Datasets

This section refers to datasets that will be generated during the lifetime of the project and will be used for academic research purposes. Datasets will be reported as they become available using the relevant FAIR analysis table, as shown in Table 7.

1. Dataset summary

Purpose

Relation to the objectives of the project

Types/Formats

Re-use of any existing data

Origin

Size

Utility for others

2. FAIR Data

2.1. Making dataset findable, including provisions for metadata

Metadata provision

Metadata standards

Unique identifier
Naming conventions
Search keywords
Version control
2.2 Making dataset openly Accessible
Classification
Sharing and access regimes
Needed method/software
Repository
Access authorisation
2.3. Making dataset interoperable
Data/metadata vocabularies and other I/O standards -
Mapping to common ontologies -
2.4. Increase re-use of dataset (through clarifying licences)
Licence
Re-use availability schedule
Re-use by third parties
Quality assurance
Availability period

Table 7. Description of REMARKABLE's dataset

5.4. Project Publications

Scientific Publications

Along with the dissemination of project deliverables and datasets, we are considering as part of the DMP, further dissemination of project Scientific Publications.

Each publication will be reported using a table with the structure presented in Table 8.

1. Publication summary
Purpose
Relation to the objectives of the project
Types/Formats
Re-use of any existing data
Origin
Size

Utility for others	
2. FAIR Data	
2.1. Making publication findable, including provisions for metadata	
Metadata provision	
Metadata standards	
Unique identifier	
Naming conventions	
Search keywords	
Version control	
2.2 Making publication openly Accessible	
Classification	
Sharing and access regimes	
Needed method/software	
Repository	
Access authorisation	
2.3. Making publication interoperable	
Data/metadata vocabularies and other I/O standards	-
Mapping to common ontologies	-
2.4. Increase re-use of publication (through clarifying licences)	
Licence	
Re-use availability schedule	
Re-use by third parties	
Quality assurance	
Availability period	

Table 8. Description of REMARKABLE's publication

Other Publications

Other publications encompass any published material created by consortium members during the project's lifetime that falls outside the academic research field. This includes press releases, presentations, software documentation, or multimedia productions for dissemination purposes, such as videos. Such material will be accessible on the REMARKABLE website, either in its original form, as a link to relevant social media or data repository platforms (e.g., YouTube or Zenodo), or embedded within multimedia frames (e.g., embedded radio interviews). Comprehensive metadata should accompany these materials on their respective sources.

When producing such content, special attention should be given to providing references to the various sources used. If a source is not publicly available, consent should be sought. Publications from third parties referencing REMARKABLE (e.g., blog posts, special articles, or extensive TV reports on the REMARKABLE approach) or otherwise related to the project require approval from the project's Coordinator and Dissemination Leader. These publications will be archived in a dedicated section of the website for reference.

6. Summary and Conclusions

This deliverable presents REMARKABLE's Data Management Plan (DMP). It outlines key principles, regulations, and methodologies aligned with the DMP, including adherence to EU guidelines on open access to scientific publications and research data, as well as the FAIR Data Management principles. This document defines an adapted FAIR Data template to present expected data, categorized into four main groups: public deliverables, software components, research datasets, and publications.

Dedicated sections detail relevant storage solutions selected for their appropriateness to the project's needs. Given the evolving landscape of EU data protection regulation, particular attention is paid to identifying GDPR roles among consortium partners to clarify data protection responsibilities. Foreseen acquisition and processing of personal data for REMARKABLE's requirements are also outlined to the extent possible.

The methodologies outlined herein will primarily serve as a guiding framework for handling project-derived data. It's important to note that this document is dynamic; this version of the DMP is preliminary and subject to potential amendments and additions throughout the project lifecycle. As this is the only planned release of the deliverable, it will be made available to the EU commission at the end of the project with all changes it was subject to.

For data stemming from publications, storage will primarily utilize OpenAIRE/Zenodo/ResearchGate indexed repositories, with a preference for the green model. Other publications, such as deliverables, will be housed within ULHT's hosting services, while public deliverables will be made available through the project website.