D1.2: Data Management Plan

WP1: Project Management, Communication, Dissemination and Exploitation

Version 1.0

Authors:

Giuseppe Carcasi; Angela Nocita; Camilo Patiño Puerta; Marjorie de Belen

WP leader: RFI

Delivery date: 2023-07-31

Dissemination level: Public

Type: Report

Deliverable Information

D1.2 Data Management Plan			
Work Package	WP1: Project Management, Communication, Dissemination and Exploitation		
Lead Partner	Rete Ferroviaria Italiana (RFI)		
Author(s)	Giuseppe Carcasi; Angela Nocita; Camilo Patiño Puerta; Marjorie de Belen		
Due Date	31 July 2023		
Version no.	1.0	**Status**	Final
Dissemination Level	Public		

Project Number	101121851	**Project Acronym**	MaDe4Rail
Project Title	Maglev-Derived Systems for Rail		
Start date	1 July 2023		
Project Duration	12 months		
Call Identifier	HORIZON-ER-JU-2022-02		
Topic	HORIZON-ER-JU-2022-FA7-02: Technological development of Maglev-derived Systems		
Type of Action	HORIZON JU Research and Innovation Actions		
Granting Authority	Europe's Rail Joint Undertaking		

Revision History

Revision	Date	Contributor(s)	Description
V0.1	2023-07-21	Camilo Patiño Puerta (RFI); Marjorie de Belen (RFI)	First draft
V0.2	2023-07-22	Stefano Ricci (DITS)	Draft reviewed by WS1 leader
V0.3	2023-07-26	Florian Brinkmann (DLR)	Draft reviewed by WS2 leader
V1.0	2023-07-28	Giuseppe Carcasi(RFI); Angela Nocita (RFI)	Final revision, document for submission

Disclaimer

This document has been produced in the context of the MaDe4Rail project. The MaDe4Rail project is part of the Horizon Europe Framework Programme (HORIZON) for research and development and is as such funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or Europe's Rail Joint Undertaking (ERJU). Neither the European Union nor the granting authority has liability with respect to this document and they cannot be held responsible for it. The user thereof uses the information at its sole risk and liability.

Table of Contents

List of Tables

1 Executive Summary

This document constitutes the deliverable D1.2 of the MaDe4Rail project and describes the Data Management Plan (DMP) for the project. The deliverable aims to ensure the highest quality of project outputs through the definition and standardization of data types, formats, data storage and security, maintenance and alignment practices under the FAIR data principles, in order to meet the Open Science requirements of the Horizon Europe programme.

This deliverable serves also as a handbook for all project partners, offering structured data management processes, methods, and tools to be implemented for data storage and administration in order to deliver quality assured results and outputs and maintain the integrity of the project.

The MaDe4Rail data management plan is not a fixed document but will be updated as the project activities are carried out, aligning with technical developments, generated data (Project Data) and used data (Provided Data).

Keywords: Data Management; Data Processing; Data Storage and Security

2 Abbreviations and Acronyms

Abbreviation/Acronym	Description
CA	Consortium Agreement
CDEP	Communication, Dissemination and Exploitation Plan
DITS	Development and Innovation in Transport Systems – La Sapienza
DLR	Deutsches Zentrum für Luft- und Raumfahrt
DMP	Data Management Plan
EC	European Commission
ERJU	Europe's Rail Joint Undertaking
EU	European Union
FA	Flagship Area
FAIR	Findability, Accessibility, Interoperability, and Reusability
FSI	Ferrovie dello Stato Italiane
GA	Grant Agreement
GDPR	General Data Protection Regulation
IP	Intellectual Property
IPR	Intellectual Property Rights
ITF	Italferr
MaDe4Rail	Maglev-Derived Systems for Rail
Maglev	Magnetic levitation
MAWP	Multi-Annual Work Programme
MDS	Maglev-Derived Systems
NDA	Non-Disclosure Agreement

ORDP	Open Research Data Pilot
RFI	Rete Ferroviaria Italiana
TRV	Trafikverket
TU Delft	Technische Universiteit Delft
Uni.Eiffel	Université Gustave Eiffel
UPM	Universidad Politécnica de Madrid
WP	Work Package
WS	Work Stream

3 Introduction and Objective

MaDe4Rail: Maglev-derived Systems for Rail is one of the projects of Europe's Rail Flagship Area (FA) 7 that is part of the Innovation Pillar and is described in the ERJU Multi-Annual Work Programme (MAWP). The project aims to explore non-traditional and emerging Maglev-Derived Systems (MDS), which can be defined as innovative, fast track-bound transportation systems for rail application that use maglev-based technologies, such as linear motors with magnetic/pneumatic levitation, as their foundation. They can be stand-alone systems with their own dedicated infrastructure or can be, in principle, integrated within the existing railway infrastructure.

The project will identify technical enablers and technologies for MDS defining a common system architecture, analyse risks and needs for standardisation on safety and security, elaborate the design concept and prototype for a MDS vehicle and evaluate the technical and economic feasibility and effectiveness of introducing MDS in Europe. The MaDe4Rail project is expected to have significant impacts such as contributing to the development of MDS, promoting more sustainable passenger and freight transport and reinforcing the role of railway as the backbone of a multimodal, sustainable, efficient and advanced mobility network by possibly, upgrading the existing railway lines/facilities through the adoption of maglev-derived technologies.

The main objective of the Data Management Plan (DMP) is to provide a structure on how data used (Provided Data) or generated (Project Data) will be managed through the duration of the MaDe4Rail project and beyond. The DMP addresses the different data types in the project; and how data will be formatted, stored and shared following the FAIR Data Principles and in connection to the Data Protection considerations defined in the Article 15 of the Grant Agreement (GA). The DMP can continuously be updated during the whole duration of the project as new data and datasets will be used or created within project execution.

The GA, in Annex 1, Section 1.2.5: *Open Science Practices and Research Data Management,* states that: "*information will be generated throughout the project that will be managed by MaDe4Rail according to the principles of Findability, Accessibility, Interoperability, and Reusability (FAIR) to ensure its results are open to use to the research and development communities and external stakeholders*". Thus, this DMP provides methods for the MaDe4Rail project's internal data management in alignment with FAIR principles, while maintaining security of proprietary confidential data (background information and Intellectual Property Rights) and ethical aspects.

4 Data Summary

4.1 Data Types

There are three (3) types of data foreseen for the development of the MaDe4Rail project: "Project Data", "Provided Data" and "Personal Data".

4.1.1 Project Data

Project Data refers to results in accordance with the provisions of Article 16.2 of the GA. In this context, Project Data means any figure produced directly by hardware and/or algorithm other than personal data according to Article 4 no. 1 of the European General Data Protection Regulation (GDPR), including, but not limited to, all output of measurements and sensors, raw data, refined data, analyses data, geodata, evaluation data and electronical data and/or data in writing produced by the project's activities.

Project Data types that will be generated in the MaDe4Rail project may include, among others, the following:

- Technical reports (periodic technical reports and internal technical reports) and technical specifications
- Presentations for meetings or for dissemination of results during conferences, events, fairs, webinars, workshops
- Videos for dissemination of results
- Test reports
- Technical drawings and designs.

4.1.2 Provided Data

Provided Data means a subset of Project Data which are obtained by a partner outside of the project or by external sources/parties and introduced into the Action during the project implementation.

Provided Data types that will be used in the MaDe4Rail project and their origins may include, among others, the following:

- Publicly available deliverable reports of relevant projects for MaDe4Rail;
- Publicly available technical reports, presentations specification documents from relevant projects for MaDe4Rail;
- Confidential or Sensitive data that is brought to the project by one of the partners as "Background", and is identified in Attachment 1 of the CA, wherein the restrictions and/or conditions for implementation or exploitation are also specified. Such Background Data should be treated in accordance with Section 9.1 of the CA;
  Confidential or Sensitive data that is brought to the project by external sources/parties, and will be governed by a specific Non-Disclosure Agreement (NDA), where the restrictions and/or conditions for implementation or exploitation will be specified, among the external party and the members of the Consortium involved in the WP wherein the data will be utilized,

4.1.3 Personal Data

Personal Data is defined by the GDPR as *"any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in*

particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person".

Personal Data that will be used in the MaDe4Rail project may include, among others, the following:

- Contact data from representatives of the MaDe4Rail project partners (name, last name, organization, position, phone number, address, email address)
- Contact data from eventual stakeholders interviewed for the MaDe4Rail project (name, last name, organization, position, phone number, address, email address)

All Personal Data within the MaDe4Rail project will be treated with the highest ethical standards and the applicable legislation on ethical principles (see Section 8).

4.2 Data Formats and Conventions

In order to comply with Findability and Accessibility principles, standard computer file formats and naming conventions are defined for the MaDe4Rail project.

4.2.1 Data Formats

For the Project Data, Table 1 presents the defined standard computer file formats to ensure document compatibility for each "Data Type" under "Data Formats".

Data Type	Data Formats
Technical reports and technical specifications	Microsoft Office Word (.docx), PDF (.pdf), Microsoft Office Excel (.xlsx), or compatible open-source files
Presentations	Microsoft Office PowerPoint (.pptx), PDF (.pdf) or compatible open-source files
Videos	.mp4, .avi, .mov
Images and logos from partners	.jpeg, .jpg, .png
Test reports	Microsoft Office Word (.docx), PDF (.pdf), Microsoft Office Excel (.xlsx) or compatible open-source files
Technical drawings	.dwg, .ifc, .bcf, .dng, .rvt, Standard file formats (including .pdf)
Datasets coming from algorithms or raw data on test cases in standard file formats associated to the instrument the data is generated on	Formats that are generally accessible (e.g. .xml, .txt, .csv)
Geographical information	.kml, .prj, .shp, .gpx, .dbf, .qgz

Table 1: Data type and data formats for the MaDe4Rail project

4.2.2 File Naming Conventions and Formatting

According to the GA, all deliverables foreseen for the MaDe4Rail project are of the type "Document/Report". A template for all deliverables (of the Document/Report type) has been produced and is available on the shared folder. The template already provides the following general structure to be maintained and followed:

- Cover Page
- Disclaimer
- Table of Contents
- List of Tables
- List of Figures
- Executive Summary
- Abbreviations and Acronyms
- Introduction and Objective
- Core Chapters
- Conclusion
- References
- Annexes (if any)

For the documents and other types of data files, the following naming and formatting rules should be used:

- The document naming procedure should be: YYYY-MM-DD_MaDe4Rail_Name_of_the_Document_ Partner'sAcronym_Version_Responsible'sInitials.FileExtension
- The date should follow ISO format 8601: YYYY-MM-DD
- The name of the document should be self-explanatory and concise in order to keep the file findable
- The versioning system will follow a numbering rule, where drafts will be 0.X (e.g., V0.4) and final versions will be X.0 (e.g., V1.0) following revisions
- Preferred fonts are Open Sans, Calibri
- Font size:
    - Documents and reports: 11
    - Presentations: minimum 14.

4.3 Data Management

All data with Consortium access level generated in the MaDe4Rail project will be accessible during the project to all partners through a specific shared folder for online collaboration, as specified in D1.1: Quality Plan. After an analysis of possibilities, considering security, flexibility, easiness and cost efficiency, the Consortium has opted for the commercial solution Microsoft SharePoint Online. Data storage and data sharing among project partners in the shared folder will be managed and maintained by the Project Coordinator.

4.3.1 Shared Folder Structure

The Shared Folder features the following channels:

- **General**: contains all the MaDe4Rail data, documents and information that is not related to any WP in particular. The General channel contains the following folders:
    - Administration: contains project administrative documents such as, guidelines and project communication material for partner reuse (e.g., logos, press releases, visuals and images)

- o Contracts: contains official MaDe4Rail project contracts like the Grant Agreement (GA) and Consortium Agreement (CA), eventual amendments, and eventual NDAs signed between members the Consortium and external parties.
- o Directory: contains the address list and mailing list of representatives from each organization for every Work Package and also eventual address lists and mailing lists for external parties and stakeholders.
- o Templates: contains the templates for official documents and presentations for the MaDe4Rail project.
- **External Access**: this area will be used for sharing data and files with the European Commission (EC) Project Officer. It will be managed by the Project Coordinator (RFI) and will contain a folder for every WS and every WP.
- **Work Package 1**: the WP1, corresponding to Project Management, Coordination, Dissemination and Exploitation has a separate channel as its activities are transversal to the other WPs. It will contain the WP1 deliverables (including older versions), financial documents and all other related documents to the WP1 activities. The Project Coordinator (RFI) will be in charge of managing the WP1 channel. It will also include a folder for Dissemination, that will contain the agenda, presentations, minutes and other documents related to Made4Rail events and meetings. It will also comprise all the documents and files for dissemination and communication to external parties (e.g., papers, journal articles, presentation, etc.).
- **Work Stream 1**: There will be a channel for WS1 and a specific separated channel for each WP contained in WS1 (WP2, WP3, WP4 and WP5). The WS leader (DITS) will be in charge of managing the WS1 channel, and each WP leader will be in charge of managing the respective WP channel and is free to set it up as better suited to the WP needs.
- **Work Stream 2**: There will be a channel for WS2 and a specific separated channel for each WP contained in WS2 (WP6, WP7 and WP8). The WS leader (DLR) will be in charge of managing the WS2 channel, and each WP leader will be in charge of managing the respective WP channel and is free to set it up as better suited to the WP needs.

4.3.2 Access and Protection

The shared folder is accessible to registered MaDe4Rail partners. Access to the shared folder will be managed by the Project Coordinator. The representative of each organization of the MaDe4Rail project will have access to the General, External Access and WP1 channels. Additionally, the representative of each organization that participates in a specific WP, will have access to the WP specific channel and to all information inside of them, as well as the possibility to upload and download files. Additional to the representatives of each organization that participates in the WP, the Project Coordinator and the WS Leader (of the WS where the WP is included) will have access to every WP channel.

Aside from the partners, access will also be provided to the Project Officer for the External Access channel only. Security measures are in place to protect the shared folder for unauthorized access and disclose of confidential information. Nevertheless, it should be noted that sensitive data should not be shared through this shared folder and should be securely maintained on internal severs of project partners.

As a general rule, all the project partners are obliged to adhere to the Intellectual Property Rights (IPR), Background and Results, Access Rights and Rights of Use as defined in Article 15 of the GA, and to the Section 9 and Section 10 of the CA. With regards to data sharing, especially for communication and dissemination

purposes, all data results that are not deemed as sensitive or confidential will be shared and available to the general public.

5 FAIR Data Principles in the MaDe4Rail Project

As stated in the GA (Annex 1, Section 1.2.5), the MaDe4Rail project is committed to adopt the FAIR (Findability, Accessibility, Interoperability, and Reusability) principles for the management of the data collected (Provided Data) or generated (Project Data) throughout the project, to ensure all outcomes are available to use for research, innovation and development purposes and for a proper communication and dissemination of the project's own results.

5.1 Making Data Findable

The Project Data produced will follow a specific naming convention and clear version numbers that will make it findable both for humans and computers (see Section 0). A clear set of keywords should also be provided to all documents produced within the MaDe4Rail project in order to optimise findability and thus, possibilities of re-use.

5.2 Making Data Openly Accessible

Project Data produced for deliverables with public dissemination level will be open to the general public and will follow Open Research Data Pilot (ORDP) principles. Project Data produced for deliverables with public dissemination level could be used for scientific publications and will consequently be made available through publication on specific open access repositories. All scientific publications produced will fulfil the requirements of the Gold[1] or Green[2] Open Access Publishing policies. To this end, the Open Research Europe platform will be one of the considered when publishing scientific articles foreseen in the CDEP.

5.3 Making Data Interoperable

Data will be produced or collected in standard file formats associated to the instrument the data is generated on, and can be read out by anyone who is in possession of the dedicated software. To facilitate exchange and re-use of data, data sets will be exported and stored in formats that are more generally accessible (see Section 4.2.1).

To ensure the interoperability of the data, standard vocabularies will be used when applicable, including vocabularies from commonly accepted literature in the specific field. Abbreviations and acronyms will always be written in full and clarified the first time they are used, and they will be specified in a specific chapter of the document (see Section 0).

5.4 Increase Data Re-Use

All the data produced for deliverables with public dissemination levels is expected to have long-term value. The data with long term value will be described in detail and the categories of communities that it could be

---

[1] Gold open access refers to open access publishing. More specifically, according to the European Commission, it may refer to an article that is "immediately published in open access mode. In this model, the payment of publication costs is shifted away from subscribing readers."

[2] Green open access refers to a self-archiving open access. The European Commission specifies that "the author, or a representative, archives (deposits) the published article or the final peer-reviewed manuscript in an online repository before, at the same time as, or after publication. Some publishers request that open access be granted only after an embargo period has elapsed."

useful for will be identified. All data with public dissemination level will be made available for re-use. This will contribute to the Open Access of the data. However, the CDEP will define a specific publication policy and publication notification procedures to be implemented in order to ensure protection and exploitation of IPR.

6   Allocation of Resources

Although there are no specific efforts defined for data management, most of the activities foreseen in order to carry out the DMP fall inside the tasks of WP1. Table 2 presents the staff efforts in Person-Month for each partner in WP1, which includes the effort for data management. Nonetheless, all the WPs must ensure that the DMP is followed and fulfilled within each Task.

Partner	Person-Month Effort
RFI	1.68
FSI	0.44
ITF	0.60
DITS	1.80
DLR	0.80
IRONLEV	0.40
MER MEC	1.20
NEVOMO	0.90
PRORAIL	0.50
TU DELFT	0.50
TACV LAB	0.10
TRV	0.60
KTH	0.00
UNI.EIFFEL	0.20
UPM	1.00
GESTE	0.40
Total	**11.12**

Table 2: Partner's staff efforts for Work Package 1.

The Project Coordinator (RFI) is the main responsible for defining the DMP and managing the shared folder in general, and the General, External Access and WP1 channels content in the specifics. WS Leaders (DITS and DLR), WP Leaders and Task Leaders, will also ensure the compliance of the DMP at the respective levels for the actions of which they are responsible, and will manage the contents of the specific channels respectively.

7 Data Security

All data, datasets documents and files must be kept at the software drives of each responsible partners' organization. Each partner is obliged to maintain the data regarding backups and secure storage (with limited and protected access to team members only), guaranteeing that only users relevant to MaDe4Rail project implementation have access to the Project Data and Provided Data during the duration of the project. Access to data and datasets will be set up by the respective IT departments of each organization.

The CA, signed by all partners, defines confidential information as all information in whatever form or mode of communication, which is disclosed by one partner to another in connection with the project during its implementation and which has been explicitly marked as "confidential". The CA defines the treatment of confidential information and thus confidential data, agreed between partners of the Consortium.

Each partner has filled up the Attachment 1 of the CA containing the background data/information that it would share to the consortium to perform the Action. Specific restrictions for implementation and for exploitation are stated for each type of data. More details could be found in Section 8, Section 9 and Section 10 of the CA.

8 Data Ethics

All partners of the MaDe4Rail project must ensure that the data management will be carried out in line with the highest ethical standards and the applicable EU, international and national level legislation on ethical principles. All MaDe4Rail partners will comply with the provisions set out in Article 14 "Ethics and Values" of the GA, and with Article 19 "Ethics" of the Regulation (EU) 2021/695 of the European Parliament and of the Council of 28 April 2021 establishing Horizon Europe.

All Project Data and Provided Data from MaDe4Rail research activities will be managed according to Article 15 "Data Protection" of the GA. Also, all personal data that is processed within the project, must comply with the GDPR. Project results or confidential information that is disclosed between the beneficiaries shall not include personal data. Personal data that is processed within the MaDe4Rail project will not be disclosed to any other external party. In case any additional issues related to data ethics surface within the project, partners must inform the Project Coordinator.

9 Conclusion

The deliverable D1.2: Data Management Plan, summarizes the activities to carry out in order to ensure the highest quality of project outputs through the management of Project Data and Provided Data. The document defines data types, data formats, naming conventions, document formatting and data management, as well as the FAIR principles adopted by the MaDe4Rail project, data ethics and data security aspects foreseen to achieve the project's objectives. The Project Coordinator is responsible for the data management in general, while WS, WP and Task Leaders have the responsibility at their corresponding level.

The deliverable D1.2 is meant to guide all partners of the MaDe4Rail project in the handling, storing and sharing of data and datasets in order to deliver the highest quality outputs for the project. The DMP will be followed and applied during the entire duration of the project towards successful delivery of deliverables and planned results, and will be updated, if necessary, for revisions and for the creation of new versions.

10 References

1. Europe's Rail Joint Undertaking Governance and Process Handbook, September 2022, Version 2.5

2. MaDe4Rail Grant Agreement (GA Nr. 101121851), available to all Consortium partners

3. MaDe4Rail Consortium Agreement, available to all Consortium partners

4. Regulation (EU) 2021/695 of the European Parliament and of the Council of 28 April 2021 establishing Horizon Europe, April 2021

5. Europe's Rail Joint Undertaking Multi-Annual Work Programme, March 2022, Version 2.0

6. European General Data Protection Regulation GDPR, 2016

7. Open Data, Software and Code Guidelines | Open Research Europe (europa.eu)

8. Data Management, Horizon 2020 Online Manual | Open Research Europe (europa.eu)