



# DELIVERABLE 1.3

## ROADMAP FOR

### DATA MANAGEMENT PLAN

#### DEVELOPMENT

Within the framework of **PNCADAI – National Program for Open Science and Open Research Data**

Part of **Measure RE-C05-i08 – More Digital Science**  
under the **PRR – Recovery and Resilience Plan**

(**Contract Date** – 13/03/2025) to (**Completion Date** – 31/12/2025)

**Date:** 31/12/2025

**Prepared by:** João Aguiar Castro (INESC TEC), Inês Sousa (INESC TEC),  
Maria Paola Tomasino (CIIMAR), António Muñoz (BIOPOLIS), and Nuno  
Fonseca (BIOPOLIS)

By:



Supported by:



# INDEX

Executive Summary .....	3
Scope & Audience .....	3
Key Principles .....	3
How to Read the Roadmap .....	5
THE ROADMAP .....	7
WAYPOINT 1: Diagnosing the data .....	8
WAYPOINT 2: Describing how data are created and processed .....	9
WAYPOINT 3: Structuring and describing data .....	10
WAYPOINT 4 : Storage, access, and active data management.....	11
WAYPOINT 5: Legal, ethical, and contractual constraints .....	12
WAYPOINT 6: Sharing data beyond the project.....	13
WAYPOINT 7: Preserving data beyond the project .....	14
WAYPOINT 8: Roles and Responsibilities.....	15
Recommendations.....	16

## Executive Summary

This document presents a lightweight Roadmap to support the development of Data Management Plans (DMPs) across diverse research contexts. It helps those involved in preparing, reviewing, or maintaining DMPs decide how much detail is appropriate, rather than prescribing a single model or level of specification.

The Roadmap is based on the principle of proportionality: the level of effort and detail in a DMP should reflect the characteristics of the data and their management needs, not project size or funding requirements alone.

Guidance is organised around a set of Waypoints that highlight key decision moments commonly encountered when working with DMPs. These Waypoints are not steps or requirements, but prompts to support judgement and improve the practical usefulness of DMPs over time.

Rather than translating FAIR into prescriptive requirements, the Roadmap focuses on decisions that shape whether data can, in practice, become findable, accessible, interoperable, and reusable.

## Scope & Audience

This Roadmap applies to the development, review, and ongoing use of DMPs across a wide range of institutional, disciplinary, and project contexts.

It is intended for:

- Researchers with coordination or leadership responsibilities
- Research support staff and project managers
- Data stewards and others involved in DMP support and review

The document is designed for selective reading and practical use. Different users may engage with different Waypoints depending on their role, the stage of the project, and the characteristics of the data involved. No specific tools, templates, repositories, or technical solutions are prescribed.

## Key Principles

This Roadmap is grounded in a small set of guiding principles that inform decision-making throughout the development and use of a DMP. These principles do not replace formal requirements, but support their practical application across different stages of DMP development.

- **Proportionality** - DMP detail scales with data workflows complexity and diversity, not project size.
- **Operational Clarity** - prefer concrete, verifiable decisions over generic statements.
- **Data context sensitivity** - data handling decisions depend on data type, infrastructure and partnership dependencies, and practical constraints.
- **Reuse-driven thinking** - the DMP is intended to remain useful and reusable throughout the project lifecycle.
- **Distributed accountability** - roles, responsibilities, and review points are explicit and shared across the project.
- **Iteration and learning** - the DMP is maintained, reviewed, and improved as experience accumulates during the project lifecycle.

### The DMP is a living document

Periodic review is not a sign of weak planning, but of maturity. This Roadmap encourages progressive refinement of the DMP as data, methods, and constraints evolve.

## How to Read the Roadmap

The Roadmap is organised around a sequence of Waypoints. Each Waypoint represents a moment where specific decisions about data management typically need to be made in the context of a DMP.

Waypoints do not prescribe actions, phases, or implementation steps. They identify decision points where particular questions about data management become relevant, and where the level of detail in the DMP needs to be considered. In the context of a DMP, these decision points arise when the document is written, reviewed, or updated, and when choices need to be made about how explicitly data management decisions should be described.

Waypoints are not a workflow, a checklist, or a maturity model. They support proportional and context-sensitive descriptions of data within a DMP.

### What Waypoints are for

Each Waypoint focuses on a limited set of decisions that shape how data will be managed. Its purpose is to:

- clarify which decisions are relevant at that point;
- support judgement about whether minimal description is sufficient or greater detail is required;
- make data-related complexity visible without moving into technical implementation.

## Layers of detail

Within each Waypoint, guidance is organised into three layers of detail. These layers do not represent quality, maturity, or compliance. They indicate how much detail the DMP needs in order to adequately reflect the complexity of the data and their workflows.

### What is needed in most cases

Captures the essential information typically sufficient to understand the data context and support early data management decisions, particularly where data workflows are simple and stable.

### What strengthens clarity and reuse

Adds contextual detail that improves transparency, interpretation, coordination, or reuse, even when data are not inherently complex.

### What may need further explanation

Highlights dependencies, constraints, or risks that may require clearer or more detailed description to support later decisions or avoid misunderstandings.

# THE ROADMAP

<b>WAYPOINT</b> <b>1</b> Diagnosing the data	Establishing a clear picture of the data landscape, and identifying where explicit description will be needed
<b>WAYPOINT</b> <b>2</b> Describing how data are created and processed	Making explicit how datasets are generated, transformed, and combined over time
<b>WAYPOINT</b> <b>3</b> Structuring and describing data	Defining how datasets are generated, transformed, and combined over time
<b>WAYPOINT</b> <b>4</b> Storage, access, and active data management	Deciding how data are stored and accessed while they are actively being used in the project
<b>WAYPOINT</b> <b>5</b> Legal, ethical, and contractual constraints	Identifying constraints that shape how data can be accessed, used, shared, or reused
<b>WAYPOINT</b> <b>6</b> Sharing data beyond the project	Deciding whether, how, and under what conditions data will be shared with audiences beyond the project team
<b>WAYPOINT</b> <b>7</b> Preserving data beyond the project	Establishing which data is retained for the long term and how continued availability is ensured
<b>WAYPOINT</b> <b>8</b> Roles and responsibilities	Clarifying who is responsible for maintaining, reviewing, and updating data management decisions over time

### Purpose

To build a shared understanding of the data landscape before decisions are made about documentation, storage, sharing, or compliance.

### Key decisions

- What data will be produced, collected, or reused
- How many distinct datasets are involved
- How datasets relate to each other (independent, sequential, derived)
- Whether any data require special handling (e.g. *personal, confidential, restricted*)

### What is needed in most cases

- Identify the main datasets using short, plain-language descriptions
- Indicate the type of data involved (e.g. *measurements, images, text, code*)
- State whether data are newly produced, reused, or combined

**The DMP provides a clear initial picture of the project data, allowing readers to understand the scope of the data landscape at a glance.**

### What strengthens clarity and reuse

- Group datasets into clear families (e.g. *raw data, processed data*)
- Describe how datasets relate to each other (e.g. *inputs, transformations, outputs*)
- Flag early where different datasets may require different handling or access conditions

**The DMP helps readers understand how the project's data are organised and connected, supporting interpretation and reuse beyond the immediate project team.**

### What may need further explanation

- Explain how datasets are connected or derived, where this is not obvious
- Which data introduce additional complexity (e.g. *heterogeneous formats, sensitive or restricted data*)
- Why certain datasets require different handling within the project

**The DMP makes data dependencies and sources of complexity explicit, preparing later informed decisions**

# WAYPOINT 2

## Describing how data are created and processed

### Purpose

To describe how the datasets identified in the DMP come into being and how they change, in order to surface dependencies or constraints that matter for data management decisions.

### Key decisions

- How each dataset is generated, collected, or acquired
- Whether data creation follows a standard or repeatable process
- Whether datasets are processed, transformed, or derived from other datasets
- Whether creation or processing steps introduced dependencies relevant to later decisions

### What is needed in most cases

- Describe how data are generated or collected (e.g. *observation, experiment*)
- State whether data creation is standardised or ad hoc
- Indicate whether data are processed, without detailing steps

**The DMP briefly explains how project data come into being and whether they change over time, moving beyond description of datasets to their creation and evolution.**

### What strengthens clarity and reuse

- Describe the main processing or transformation steps applied to the data
- Indicate which datasets are derived from others within the project
- Clarify whether processing is manual, automated, or software-assisted

**The DMP clarifies how datasets evolve from one another over time, helping readers understand data provenance and interpret derived data correctly.**

### What may need further explanation

- How dependencies between datasets arise when data are created or processed in stages
- Which processing steps affect data quality or interpretation
- Where data creation or processing relies on specific tools or external services, if relevant

**The DMP highlights where data creation and processing introduce dependencies or constraints that matter for later data management decisions.**

# WAYPOINT 3

## Structuring and describing data

### Purpose

To determine how datasets will be structured, named, and described so that they can be understood, located, and interpreted during and after the project.

### Key decisions

- How datasets are organised (folders, collections, logical groupings)
- Whether consistent naming and versioning are needed
- What descriptive information (metadata) is required to make the data intelligible
- Whether different datasets require different levels or standards of description

### What is needed in most cases

- Describe how data files are organised (*e.g. folders, logical groupings*)
- Indicate how file naming and versioning are handled
- State what basic descriptive information is recorded to help others understand the data

**The DMP explains how data can be identified and understood during and after the project.**

### What strengthens clarity and reuse

- Describe naming, versioning, or file structures in detail
- Identify which descriptive elements are particularly important to enhance reuse
- Indicate whether community or domain metadata standards are used

**The DMP enables others to navigate, interpret, and reuse the data with minimal additional guidance.**

### What may need further explanation

- Describe why different data require different structures or descriptions
- Explain how data structure choices may influence interpretation or reuse, where relevant
- Flag risks related to inconsistent naming, versioning, or description

**The DMP makes clear where data structure or description choices may affect how the data are interpreted or reused.**

# WAYPOINT 4

## Storage, access, and active data management

### Purpose

To clarify where data are stored during the project, who can access them, and how data are protected while in active use.

### Key decisions

- Where the data are stored at different stages of the project
- Who has access to which datasets
- How data integrity, security, and backup are ensured
- Whether storage or access constraints affect other data management decisions

### What is needed in most cases

- Identify the primary storage location(s) for active data
- Indicate who has access to the data
- State whether routine backup or institutional storage is used

**The DMP shows where data are kept during the project and who can access them.**

### What strengthens clarity and reuse

- Distinguish storage for different stages of data (e.g. *raw, working, processed*)
- Clarify access levels for different roles or partners
- Indicate how long data remain in active storage

**The DMP explains how storage and access arrangements differ across data and roles during the project.**

### What may need further explanation

- Why different storage locations or services are used during the project
- How access restrictions are handled where data sensitivity or collaboration requires it
- Which storage constraints affect data handling (e.g. *capacity, performance, security*)

**The DMP highlights where storage and access arrangements introduce constraints or dependencies that affect ongoing data work.**

### Purpose

To identify legal, ethical, and contractual constraints that affect data management decisions, and to ensure their implications are understood early.

### Key decisions

- Whether any datasets are subject to legal, ethical, or contractual constraints
- Which datasets are affected, and in what way
- Whether constraints apply throughout the project or change over time
- How constraints influence storage, access, sharing, or preservation decisions

### What is needed in most cases

- State whether any data are subject to legal, ethical, or contractual constraints
- Indicate the general nature of those constraints (e.g. *personal data, confidentiality, third-party restrictions*)
- Identify which data are affected, at a high level

**The DMP flags whether constraints apply to the project's data and what kind of limitations they introduce.**

### What strengthens clarity and reuse

- Clarify how constraints affect data handling or sharing
- Indicate whether mitigation measures are planned (e.g. *anonymisation, aggregation, access controls*)
- Clarify responsibilities for compliance within the project

**The DMP explains how identified constraints shape data management choices and clarifies how compliance is addressed within the project.**

### What may need further explanation

- How constraints overlap or evolve over time
- Where constraints introduce dependencies or risks affecting later sharing, reuse, or preservation

**The DMP highlights where legal, ethical, or contractual constraints require careful coordination or may affect decisions later in the data lifecycle.**

### Purpose

To determine whether data will be made accessible beyond the project team, through which repositories this may occur, and under what broad access and reuse conditions.

### Key decisions

- Which data will be shared beyond the project team
- When data are expected to become accessible
- Through which repository or repositories data will be shared

### What is needed in most cases

- Identify which data will be shared beyond the project team
- When data may become accessible and under what general conditions
- Indicate the expected repository type(s), naming one or more candidate repositories (*e.g. institutional, general or disciplinary*)

**The DMP makes clear which project data are intended for external access, when this may occur, and through which routes access is provided.**

### What strengthens clarity and reuse

- Name the repository or repositories intended for data sharing
- Indicate the general access model (*e.g. open, embargoed, controlled*)
- Indicate expected reuse conditions, where known (*e.g. licence type*)

**The DMP shows how and where data can be accessed beyond the project, and under what general conditions they may be reused.**

### What may need further explanation

- Why different data are shared through different routes or under different conditions
- How timing or approvals affect when data become accessible
- Which factors may limit or delay reuse, where this is not obvious

**The DMP highlights where data sharing is conditional or differentiated, helping readers understand how access and reuse may vary across data.**

### Purpose

To establish which data should be retained beyond the project and how their long-term availability will be supported.

### Key decisions

- Which data should be retained beyond the end of the project
- For how long retained data are expected to remain available
- Through which repository, service, or preservation route long-term availability will be supported

### What is needed in most cases

- Identify which data are intended for long-term retention beyond the project
- Indicate the expected retention period, at a high level
- State the repository or service expected to support long-term availability

**The DMP identifies which project data are retained beyond the project and outline their expected long-term availability.**

### What strengthens clarity and reuse

- Clarify whether different data are retained for different durations
- Name the repositories supporting long-term availability
- Indicate any general conditions affecting continued availability

**The DMP clarifies retention durations, names the repositories supporting long-term availability, and notes any general conditions affecting access over time.**

### What may need further explanation

- Why certain data are retained while others are not, where this may not be self-evident
- How retention periods were determined, particularly when they differ across data
- What “long-term availability” means in practice in the given context (*e.g. expected duration, review points*)

**The DMP explains retention choices, how retention periods were defined, and what long-term availability means in the project context.**

# WAYPOINT 8

## Roles and Responsibilities

### Purpose

To clarify who is responsible for maintaining, reviewing, and updating data management decisions throughout the project and beyond.

### Key decisions

- Who is responsible for key data management decisions and actions
- At which points responsibilities apply or may change over time
- How responsibility is distributed across roles, partners, or organisations

### What is needed in most cases

- Identify the main role(s) responsible for data management within the project
- State who is responsible for maintaining and updating the DMP
- Indicate where responsibility lies for decisions related to sharing and long-term availability

**The DMP clearly identifies who is accountable for data management decisions within the project.**

### What strengthens clarity and reuse

- Clarify responsibilities for different data-related decisions, where they are distributed across roles
- Distinguish responsibility from execution, where tasks are delegated
- Indicate when responsibilities change over the project lifecycle

**The DMP provides a transparent picture of how responsibility for data management decisions is distributed and maintained over time.**

### What may need further explanation

- Why specific roles are responsible for certain decisions, particularly in collaborative or multi-partner projects
- How responsibility is handled when roles overlap or change over time
- Who acts as the point of contact for data-related questions beyond the project

**The DMP helps readers understand the rationale behind responsibility assignments and how continuity is ensured when roles change or overlap.**

## Recommendations

### DO

- Use the DMP as a working document that support real decisions
- Adjust the level of detail to the data and workflows actually in place
- Consult guidance and reference documents to identify FAIR practices
- Write descriptions that are clear to someone outside the project
- Use the DMP as a shared reference within the project team
- Review the DMP on a regular basis, and whenever changes are required
- Provide information that supports future understanding, reuse, or review

### DON'T

- Copy generic text from templates or previous DMPs without adaptation
- Describe ideal practices that do not reflect how data are managed in practice
- Add technical detail that does not support concrete practices
- Treat the DMP as a document completed once at the start of the project
- Assume that more detail automatically means a better DMP
- Put the DMP on the shelf

### Keep the DMP relevant

A useful DMP is one that continues to support informed decisions over time, adapting to changes in data, workflows, and project needs evolves.



**FAIRway**

