Data Management Plan (updated version)

Deliverable D7.9

Release Status: Public
Author: Romain F H Viguier, SCCS
Date: 13 February 2025
Filename and Version: PilotSTRATEGY_Data_Management_Plan_D7.9_V05
Project ID Number: 101022664

PilotSTRATEGY (H2020- Topic LC-SC3-NZE-6-2020 - RIA)

Document History

Location

This document is stored in the following location:

Filename	PilotSTRATEGY_Data_Management_Plan_D7.9_V05
Location	

Revision History

This document has been through the following revisions:

Version No.	Revision Date	Filename/Location stored:	Brief Summary of Changes
V01	15 Oct 21	PilotSTRATEGY_Data_Management_Plan_Initial_D7.9_V01	First draft
V02	01 Nov 21	PilotSTRATEGY_Data_Management_Plan_Initial_D7.9_V02	Feedback and comments from partners
V03	24 Apr 24	PilotSTRATEGY_Data_Management_Plan_updated_D7.9_V03	Updates on Metadata
V04	10 Jun 24	PilotSTRATEGY_Data_Management_Plan_updated_D7.9_V04	Added info regarding Repository
V05	13 Feb 25	PilotSTRATEGY_Data_Management_Plan_D7.9_V05	Added summary of datasets table

Authorisation

This document requires the following approvals:

AUTHORISATION	Name	Signature	Date
WP 7 Leader	Romain F H Viguier		13/02/25
Project Coordinator	Isaline Gravaud	IG	07/03/25

Distribution

This document has been distributed to: All partners

The PilotSTRATEGY project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101022664

Disclaimer

The information and views set out in this document are those of the author(s) and do not necessarily reflect the official opinion of the European Union. Neither the European Union institutions and bodies nor any person acting on their behalf may be held responsible for the use which may be made of the information contained therein.

The PilotSTRATEGY project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101022664

Executive summary

The objective of the Data Management Plan (DMP) is to identify the data used and generated in the project and to outline how this data will be made findable, accessible, interoperable, and reusable, in accordance with the principles of FAIR data management. This document adheres to the European Commission's guidelines on FAIR data management in Horizon 2020 (version 3.0, 26th July 2016) and provides information about the collected data, including its purpose, utility, accessibility, and reusability.

This is an updated version of the project's DMP and includes the selection of a repository for long-term data storage, and a summary of datasets. The DMP is being continuously reviewed, updated, and completed throughout the project, with the final version constituting the project's Final DMP (deliverable D7.10).

The PilotSTRATEGY project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101022664

Table of Contents

The PilotSTRATEGY project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101022664

The PilotSTRATEGY project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101022664

1. Introduction

Research data management in project PilotSTRATEGY will follow the Horizon 2020 program's guidelines on FAIR data management and shall be governed by Article 29 (Dissemination of results – open access – visibility of EU funding) of the Grant Agreement.

The dissemination of project PilotSTRATEGY data is subject to a procedure for dissemination approval that is stated in Article 8.4 (Dissemination) of the Consortium Agreement.

The Data Management Plan will describe the research data used and generated by the project and will ensure that research data are findable, accessible, interoperable, and re-usable. The DMP is established in collaboration with all partners in agreement with the dissemination requirement stated in the Grant agreement and with the procedure for dissemination approval stated in the consortium agreement.

The WP7 leader, SCCS/UEDIN is responsible for the coordination of task 7.6 (Data Management Plan) which includes:

- The selection of a repository for project data to provide open access to publications and data generated by the project
- Preparation of rules and instructions for data deposition in the project
- Preparation and annual update of the DMP

The project has constituted five local teams, which are groups of partners from a specific region, in charge of conducting the main work in their area. Local teams are composed of one or more institutions and are coordinated by a Public Organisation that will ensure project outputs are Findable, Accessible, Interoperable, and Reusable (FAIR).

These Public Organizations are:

- BRGM for the Paris-Orléans area
- IGME for the Ebro Basin
- UEVORA for the Lusitanian Basin
- CERTH for the West Macedonia area
- GIG for the Upper Silesia area

2. Data Summary

2.1 Purpose of the data collection/generation

PilotSTRATEGY project is investigating geological CO_2 storage sites in industrial regions of Southern and Eastern Europe for the purpose of large-scale carbon capture and storage (CCS) development.

Carbon Capture and Storage (CCS) has a crucial role in Europe's climate ambitions, particularly for the future viability of industrialised regions. Meeting the challenge will depend on the right amount of geological CO_2 storage becoming available in time. PilotSTRATEGY aims to add CCS to the feasible climate mitigation options available to local industries and communities.

The PilotSTRATEGY project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101022664

Storage pilots will help us develop Europe's CO_2 storage capacity and build confidence in CCS. PilotSTRATEGY will provide the technical basis for the implementation of pilot storage sites in five target regions, with a focus on deep saline aquifers, a largely untapped storage resource.

The project will assess factors such as a storage site's integrity, capacity, hydrodynamics, and monitoring options and undertake research into stakeholder engagement, including an evaluation of factors that affect CCS acceptance.

The project research focussing on the Paris, Lusitanian, and Ebro basins will;

- provide 3D characterisation of storage sites and storage complexes, including sedimentology, structural geology, initial field stress and natural seismicity
- establish an injection strategy and an assessment of CO_2 behaviour in the long term
- provide detailed Measurement, Monitoring and Verification (MMV) strategies as well as plans for preventive and corrective measures
- carry out pre-FEED studies for CO_2 storage resources based on pre-feasibility studies from the STRATEGY CCUS project
- involve regional stakeholders in the pilot's conceptualisation, design and monitoring strategies development
- develop recommendations for stakeholder and public participation in the delivery of geological CO_2 storage pilots

The project research focussing on West Macedonia and Upper Silesia will;

- enhance understanding of CO_2 storage resources by studying new data, reprocessing current data and conducting new dynamic simulation studies
- allow these resources to be part of contingency storage planning for the regions
- involve regional stakeholders, including communities and the wider public, in CCUS public engagement studies

2.2 Types and formats of data generated/collected

Different types and formats of data is being collected and generated in project PilotSTRATEGY.

Data from experimental and modelling work is generated in technical work-packages (WP2-WP5) in the lab and from field experiments.

Qualitative and quantitative data is generated through interviews, questionnaires, and workshops by human science related activities (WP6). This data is anonymised and processed in the project.

All information related to types and formats of generated /collected data is collected using a questionnaire. This questionnaire is transmitted via the WP leaders and filled by the partners. Completed questionnaires are uploaded on the project Sharepoint.

The data questionnaire has been updated to tackle metadata (see **Erreur ! Source du renvoi introuvable.**) and is joint in Annex 1.

The PilotSTRATEGY project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101022664

2.3 Project data

Data collected and generated in the PilotSTRATEGY project is listed in Table 1 below. Collection of project data will be updated during the project's lifetime.

2.4 Data utility

The data collected and generated in the PilotSTRATEGY project will be useful for different groups of stakeholders:

- Research institutes, universities, other R&D organisations actives in CO_2 storage assessment or appraisal will use the data to inform, educate and further research work
- Municipalities, Local authorities, NGO, decision makers will use the data to support decisions to be made regarding the development of CO_2 store in Deep Aquifer.
- CCUS developers, geological CO_2 storage pilot operators, will use the data to develop further CO_2 stores, to plan future CCUS infrastructure.

The PilotSTRATEGY project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101022664

Table 1: summary of datasets

WP/Task/Del	Description of Dataset	Access	Contact	Organisation	Archive
WP2, T2.2.1	3D seismic data onshore (Paris basin), Numerical data	For project partners only	Patrick Robert and Emmanuelle Robins patrick.robert@s3seismic.com; emmanuelle.robins@pole-avenia.com	S³ and PoleAvenia	S³
WP2, T2.2.2	Specifically locations of the baseline seismometer should be confidential to avoid vandalism until recording finishes	Not confidential	Manuel Ron Martin Ron.martin.manuel@repsol.com	REPSOL	Repsol internal repository plus specified folder in BRGM-Pilot Sharepoint
WP2, T2.2.3	Seismicity data, Seismic Catalog and waveforms from IPMA (Instituto Português do Mar e da Atmosfera), SEISAN, SAC, or Obspy (python) program, as well as the instrument responses (stationXML)	Confidential, temporary, restricted access), For project partners only	José Fernando Borges jborges@uevora.pt	Universidade de Évora	University of Évora
WP2, T2.3	Data sets of reservoir complex geomechanical characterisation: 1-P wave propagation velocity 2 – Point Load Strength Index Test	Not confidential	Carlos Alexandre da Silva Ribeiro cribeiro@uevora.pt	Universidade de Évora	Department of Geosciences & Hercules Laboratory

The PilotSTRATEGY project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101022664

	3 - Uniaxial Compressive Strength with the Schmidt Hammer				@ University of Évora; Project sharepoint
WP2, T2.3	DRX data on storage complex samples	Not confidential	Carlos Alexandre da Silva Ribeiro cribeiro@uevora.pt	Universidade de Évora	Department of Geosciences & Hercules Laboratory @ University of Évora; Project sharepoint
WP2, T2.3.1	LAS files, well data provided by DGEG (national authority)	Confidential	Daniel Marques da Silva Daniel.m.silva@galp.com	Galp	DGEG
WP3, T3.1	P10-P50-P90 Reservoir models (ACTNUM. PORO, PERMX, PERMY, PERMZ and NTG).	Confidential For project partners only	Pedro Pereira pmpereira@uevora.pt	Universidade de Évora	Project sharepoint
WP3, T3.1	Grids with related petrophysical properties (three grids for three scenarios of uncertainties)	*Confidential, only for members of the consortium (including the Commission Services)*	Bouquet Sarah; Alina Berenice-Christ, Luca Mattioni Sarah.bouquet@ifpen.fr; alina-berenice.christ@ifpen.fr ; luca.mattioni@ifpen.fr	IFPEN	PilotStrategy sharepoint
WP3, T3.1	Geological models with the results of modelling of petrophysical properties (porosity, permeability) - grids for three scenarios of uncertainties. The sharefolder location is the following path:	Confidential For project partners only	Tomasz Urych, Krzysztof Stańczyk	Central Mining Institute – National	Project sharepoint

The PilotSTRATEGY project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101022664

	Documents/WP3-Numerical/WP3_Deliverables/D3.1/D3.1_Poland_GridFiles The well log data and 2D seismic data used in WP2 and in data in WP3 were provided by the Portuguese Directorate of Energy and Geology.		turych@gig.eu, kstanczyk@gig.eu	Research Institute	
WP3, T3.3	Dynamic model - Data requires Eclipse or Intersect simulators Static Model from Lopin	Not Confidential	Manuel Ron Martin Ron.martin.manuel@repsol.com	REPSOL	Repsol internal repository plus specified folder in RGM-Pilot Sharepoint
WP4, T4.2	Repsol internal workflows – procedures for CCS Well Design could be partly confidential. This will be clearer as the work progresses.	Not Confidential	Manuel Ron Martin Ron.martin.manuel@repsol.com	REPSOL	Repsol internal repository

The PilotSTRATEGY project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101022664

3. FAIR data

3.1 Making data findable, including provisions for metadata

Data should be discoverable, with fully searchable metadata to inform prospective users of the data prepared to recognised data management standards and published in data repositories.

A listing of datasets will be maintained on the PilotSTRATEGY website with links to the data, giving a central source of information describing data associated with the project. This will be coordinated between the data manager and the project dissemination team at the University of Edinburgh. Release of datasets will be made publicly known through the project website and social media accounts.

3.1.1 Metadata

Metadata is high level data that provides a brief summary of something, so that it can be easily catalogued and searched for.

Metadata includes a data title and a data description.

Data title: The data title should employ commonly used terms related to the dataset. It should succinctly convey the content of the dataset.

Data Description: The data description should offer a concise overview of the dataset's content, of 4000 characters maximum (but could be much shorter) aiding in determining its relevance. Most key information should be included in the first one or two sentences as they tend to be used by research engines. It may include references to publicly available papers or reports that elucidate the dataset or provide useful links to webpages.

Partners completing metadata should ensure that this is of high-quality enabling users in future to find a dataset and determine if they wish to use it. Metadata must include a good explanatory title and an accurate concise description (e.g. what, where, when, how, why, who).

What: Provides a concise description of what has been recorded and the format of the data. It aims to immediately inform the reader about the nature of the resource.

Where: Indicate the spatial coverage of the data.

When: The timeframe during which the data was collected, providing clarity on the temporal scope of the dataset.

How: A brief overview of the methodology used in data collection.

Why: The purpose behind data collection and the potential users who may find the data valuable.

Who: The parties responsible for collecting and interpreting the data are identified in this section.

Datasets generated will be INSPIRE compliant, with full metadata conforming to Directive 95/46/EC of the European Parliament and DOI (Digital Object Identifiers) where appropriate. Metadata

The PilotSTRATEGY project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101022664

containing details of the dataset will be captured in a standardised discovery metadata format which complies with ISO standard 19115.

The Research Data Alliance provides a Metadata Standards Directory

(http://rd-alliance.github.io/metadata-directory/)

3.1.2 Digital object identifiers (DOIs)

It is recommended that Digital Object Identifiers (DOIs) are applied to archived datasets where appropriate to enable citation of the information, particularly when data are referenced in a publication. This is a pre-requisite of leading science journals.

The DOI:

- allows data to be cited in the same manner as a scientific journal article
- enables credit to be assigned the dataset creators
- recognises the value of the data
- and the effort that has gone into its creation
- ensures the discoverability, permanence, and stability of the dataset

A DOI can be assigned before the dataset is released so that it can be referenced in the associated publication and the dataset can be released, when notified, at the same time as the publication. Datasets can be cross-linked back to the article.

Data can be archived without a DOI as not all data are appropriate for a DOI. For a dataset to be assigned a DOI, it must be provided to the data repository in good condition, with appropriate metadata and of a suitable level of technical quality. The data depository will be responsible for ensuring the data meets the required level of quality.

A DOI gives assurance to future users that the dataset is:

- Stable
- Complete
- Permanent
- Of good technical quality

The data repository is giving its stamp of approval, saying that the dataset is complete and that all the necessary metadata are available.

3.1.3 Data access statement

Partners must include a statement in their publication(s) describing how to access the data (or a statement explaining why access to underlying data has been restricted). If data are openly available, the name(s) of the data repositories should be provided, as well as any persistent identifiers (e.g. DOI) for the dataset.

The PilotSTRATEGY project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101022664

3.2 Making data openly accessible

Due to the high level of public and industry interest in the potential impacts of CCS, the default position for the project will be for all finalised datasets to be open access. It is a requirement that all open data are accessible for the long-term. This makes the research process more robust by enabling validation of results and maximising the value obtained from publicly funded data. All public (written) deliverables should also be archived.

Data should be archived as open access which:

- underpins a publication
- has long-term interest with potential for re-use (including currently unforeseen uses)
- validates research findings
- is worth keeping

Benefits of open access:

- Accelerations of the research and discovery process
- Avoidance of the duplication of research efforts
- Enhanced opportunities for collaborations
- Broader and faster opportunities for the adoption and commercialisation of research findings

3.3 Categorisation of data access

However, not all data generated by PilotSTRATEGY must be open. The need to balance openness and protection of scientific/commercial information should be considered and certain datasets may need to remain closed according to the principle "as open as possible, as closed as necessary".

Project datasets should therefore be categorised. It should be carefully considered which data can be made public (open access) from the onset, which should be placed under temporary embargo (< 2 years) before open release, and which must remain confidential. WP leaders should discuss with their work package participants/task leaders to determine this.

3.4 Data access committee

There is not a need for a separate data access committee, but this will be an item for discussion on Project Management Board meetings.

Examples of data which could be closed or restricted:

- Confidential information

The PilotSTRATEGY project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101022664

- External industry data
- Commercial sensitivity/interest (e.g. new tools being developed with potential for patenting)
- Data with IPR issues
- Sensitive data containing personal information

Participant consent may also need to be obtained. This should be agreed during early stages of the project.

IPR and innovation for the project should be considered to ensure there are no conflicts between data sharing and these.

If certain datasets cannot be shared (or need to be shared under restrictions) the restrictions associated with the data must be valid/reasonable. Metadata should include a statement specifying any restrictions.

For confidential data, it is recommended that a discovery metadata record is published to signpost that the data exists without necessarily archiving the data. This should contain a brief description, which directs any potential user to the data owner contact if more information is required or to discuss the possibility of data access, which could lead to future collaborations. It should be discussed with the participants if they wish to advertise their (confidential) data in this way or if they prefer to make no information available in some cases.

4. Allocation of resources

Data management costs are covered as part of the grant. WP7 lead will cover data manager role and identify repository to ensure long-term data management complies with current best practice to allow continuous data availability.

Beyond the end of the project, data archived in the repositories will be preserved and maintained for the long-term using the data repository resources. The costs associated with this will not be substantial.

5. Data security

Sensitive data should be encrypted when transferring.

6. Ethical aspects

All data must conform to the EU General Data Protection Regulation (GDPR) when personal information is involved. Data can be anonymised if needed.

There may be sensitivities with some data which means it is unsuitable for sharing. This should be decided by the data generator in discussion with the WP lead and the project Coordinator.

7. PilotSTRATEGY data repository

The project has selected Zenodo as a repository for project data, to provide open access to publications and to increase the visibility of data generated by the project.

The PilotSTRATEGY project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101022664

7.1 PilotSTRATEGY community

The project has created a community on Zenodo. The link to the Zenodo community is https://zenodo.org/communities/pilotstrategy

7.2 Step per step guide on Zenodo

To help you get started, here's a step-by-step guide on how to upload to Zenodo: https://help.zenodo.org/docs/deposit/create-new-upload/

7.3 Metadata for Zenodo

Each data set and project deliverable uploaded to Zenodo must be accompanied by metadata, including the title, description, and keywords. Please refer to **Part 3 of the current document** on making your data FAIR, which covers the use of metadata.

8. Resources for DMP development

The Research Data Alliance provides a Metadata Standards Directory that can be searched for discipline-specific standards and associated tools.

The EUDAT B2SHARE tool includes a built-in license wizard that facilitates the selection of an adequate license for research data.

Useful listings of repositories include:

Registry of Research Data Repositories

Some repositories like Zenodo, an OpenAIRE and CERN collaboration, allow researchers to deposit both publications and data, while providing tools to link them.

Other useful tools include DMP online and platforms for making individual scientific observations available such as ScienceMatters

The PilotSTRATEGY project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101022664

ANNEX 1 - DATA QUESTIONNAIRE

To refine further the PilotSTRATEGY data management plan, it would assist if you could complete this questionnaire and upload it to the project SharePoint, in the following file:

Documents/WP7-Communication/Data_Management_Plan/

This will help identify and categorise project data and any potential problem areas.

1	Work package				
2	Task				
3	Contacts				
	Name				
	Organisation				
	e-mail				
4	Project timescale of data collection/production, from		to		
	Project timescale of data collection/production, from		to		
5	Where will you deposit data for long-term archiving?				
6	When do you expect to deposit data?				
7	Specific datasets				
8	All data will not necessarily be appropriate for long-term preservation with a data repository. Which of the above datasets may not be appropriate to deposit?				
9	Will you be providing any software?				
	If yes, please provide more details				
	Are there likely to be any licensing requirement?				
10	Are you using any existing data from project partners?				
	If yes, please provide some details				
11	Are you using any external third-party data?				
	If yes, please provide some details				

The PilotSTRATEGY project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101022664

12	Does data require an embargo before open access release?	
	If yes, please provide some details	
13	Is any data confidential or restricted?	
	If yes, please provide some details	
14	Are there any intellectual property rights or commercial sensitivity issues that will restrict access to data?	
	If yes, please provide some details	

15 Making data findable, including the provision for Metadata.
Metadata is high level data that provides a brief summary of something, so that it can be easily catalogued and searched for. **Metadata** includes a data title, a data description and a set of Keywords.

Data title:
Please propose a data title. The data title should employ commonly used terms related to the dataset. It should succinctly convey the content of the dataset.

Data Description:
The data description should offer a concise overview of the dataset's content, of 4000 characters maximum (but could be much shorter) aiding in determining its relevance. Most key information should be included in the first one or two sentences as they tend to be used by research engines. It may include references to publicly available papers or reports that elucidate the dataset or provide useful links to webpages.
Partners completing metadata should ensure that this is of high-quality enabling users in future to find a dataset and determine if they wish to use it. Metadata must include a good explanatory title and an accurate concise description (e.g. what, where, when, how, why, who).
***What**: Provides a concise description of what has been recorded and the format of the data. It aims to immediately inform the reader about the nature of the resource.*
***Where**: Indicate the spatial coverage of the data.*
***When**: The timeframe during which the data was collected, providing clarity on the temporal scope of the dataset.*
***How**: A brief overview of the methodology used in data collection.*
***Why**: The purpose behind data collection and the potential users who may find the data valuable.*
***Who**: The parties responsible for collecting and interpreting the data are identified in this section.*

Key words.
Please add a set of keywords to make your data more easily findable.

The PilotSTRATEGY project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101022664

Datasets generated will be INSPIRE compliant, with full metadata conforming to Directive 95/46/EC of the European Parliament and DOI (Digital Object Identifiers) where appropriate. Metadata containing details of the dataset will be captured in a standardised discovery metadata format which complies with ISO standard 19115.
The Research Data Alliance provides a Metadata Standards Directory (http://rd-alliance.github.io/metadata-directory/)

The PilotSTRATEGY project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101022664