Coordination and Support Action SET4H2

Data Management Plan

D1.5

WP1 / T1.3
August 2024

Authors: Anne Dreysel, Paula Menne, Horst Krämer, DLR

Funded by the European Union

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 CINEA. Neither the European Union nor CINEA can be held responsible for them.

Technical References

Project Acronym	SET4H2
Project Number	101160662
Project Title	Support to the SET Plan IWG on hydrogen
Project Duration	1st May 2024 – 30th April 2026 (24 months)

Deliverable No.	D1.5
Title of Deliverable	Data management Plan
Subtitle of Deliverable	
Dissemination level*	PU*
Work Package	WP 1– Organisation and Management
Task	T1.3 Monitoring, data management, risk management and quality assurance
Lead Beneficiary	DLR
Contributing Beneficiary	

* PU = Public (Fully open); SEN = Sensitive (Limited under the conditions of the Grant Agreement); EUCl = EU Classified (RESTREINT-UE/EU-RESTRICTED, CONFIDENTIEL-UE/EU-CONFIDENTIAL, SECRET-UE/EU-SECRET under Decision 2015/444.

V	Date	Beneficiary	Author(s)
1.0	06/08/2024	DLR	Paula Menne, Anne Dreysel, anne.dreysel[at]dlr.de
1.0	07/08/2024	EUREC	Anna Spoden
1.0	23/08/2024	DLR	Horst Krämer
Final	28/08/2024	DLR	Anne Dreysel

Executive Summary

The Coordination and Support Action (CSA) SET4H2 is focused on the support and uptake of research and innovation on hydrogen technology through the Implementation Working Group (IWG) on hydrogen, as part of the European Strategic Energy Technology (SET) Plan. The goal is to continue the cooperation on hydrogen within Europe and to implement the Strategic Research and Innovation Agenda published in 2022 as result of the ERA pilot initiative Agenda Process on Green Hydrogen. The CSA plays a pivotal role in maximizing and broadening the uptake of learnings from national and regional hydrogen R&I programmes, hydrogen activities and recent developments and facilitate their mutual coordination on a European level.

Additionally, it focuses on creating greater synergies between hydrogen and other renewable energy technologies – particularly those under the SET Plan but in its greater scope also with the European Research Area (e.g. ERA Action 11) and the Net Zero Industry Act. SET4H2 will implement an integrated, systematic and interdisciplinary approach to address research needs in the hydrogen sector and enhance cooperation across Member States.

The cooperation and synergies among Member States and SET-Plan countries will help to improve Europe's position in the global market and sustain the key European industries' competitiveness. These are enhanced by developing positive impacts on the next integrated SET Plan, Horizon Europe work programme, national/regional hydrogen strategies as well as National Energy and Climate Plans (NECP).

The Data Management Plan (DMP) describes the strategies and measures for handling research data during and after the CSA SET4H2 project. Its goal is to approach and determine the technical, organizational and legal aspects of research data management as early as possible in order to practise good data management and sharing, follow uniform standards and avoid data loss.

The deliverable D1.5 Data Management Plan (DMP) is part of the work package WP1 Organisation and Management, and a direct output from task T1.3 Monitoring, data management, risk management and quality assurance.

Table of Contents

List of Tables

List of Abbreviations and CSA SET4H2 Partners

List of Abbreviations

Abbreviation	Long form
CETPartnership	Clean Energy Transition Partnership
CHJU	Clean Hydrogen Partnership
CSA	Coordination and Support Action
DoI	Declaration of Intent
DMP	Data Management Plan
EERA	European Energy Research Alliance
GDPR	General Data Protection Regulation
IWG	Implementation Working Group
NECP	National Energy and Climate Plans
SET Plan	European Strategic Energy Technology Plan
SRIA	Strategic Research and Innovation Agenda
TRI	Transition Initiatives
TWG	Temporary Implementation Working Group

CSA SET4H2 consortium partners

Nr.	Partner	Acronym	Country
1	Deutsches Zentrum für Luft- und Raumfahrt e.V.	DLR	Germany
2	Ministero dell'Università e della Ricerca	MUR	Italy
2.1	Alma Mater Studiorum – Università di Bologna	UniBO	Italy
3	Direção-Geral de Energia e Geologia (Directorate General for Energy and Geology)	DGEG	Portugal
4	Balgarska Asotsiatsia za Vodorod, Gorivni Kletki I Sahranenie na Energia (Bulgarian Hydrogen, Fuel Cell and Energy Storage Association)	BGH2A	Bulgaria
5	Österreichische Energieagentur - Austrian Energy Agency	AEA	Austria
6	Association of European Renewable Energy Research Centers	EUREC	Belgium
7	Hydrogen Europe Research	HER	Belgium

1. Data Summary

SET4H2 will produce insight reports, fact sheets, policy briefs and summaries of events to facilitate communication of findings from the SET Plan to the European R&I community, policymakers, industry and society. By providing accessible and openly available information, SET4H2 ensures that lessons learned are widely understood and implemented, driving institutional changes and fostering exchanges and synergies between all stakeholders.

The main data source for SET4H2 is the website for the IWG on hydrogen. The website serves as a central hub for information on hydrogen policies on different levels, technological developments, activities, and achievements. It facilitates communication among stakeholders, offering a repository of key documents.

The project will re-use both quantitative and qualitative data from hydrogen-related European initiatives, new findings published on eCorda and CORDIS, as well as all data from national and international initiatives and policymakers:

- Quantitative Data: This includes statistics and metrics related to various hydrogen initiatives on national and international level. These data will be sourced from existing databases and reports, providing numerical insights into the performance and impact of policies and programs in relation to the implementation of the Strategic Research and Innovation Agenda of the ERA pilot initiative Agenda Process on Green Hydrogen.
- Qualitative Data: This includes survey responses, country reports, and various documents that offer detailed descriptions and analyses of hydrogen-related activities and ERA Actions.

The key types of documents the CSA SET4H2 will produce are:

- Insight Reports: Comprehensive reports analysing specific aspects of hydrogen policies, national and international R&I and funding strategies related to hydrogen, and latest technology developments in hydrogen as well as related energy technologies, providing in-depth insights and recommendations.
- Fact Sheets: Concise documents summarizing key statistics, findings, and information about hydrogen initiatives and policies, making data easily digestible for stakeholders.
- Event Summaries: Summaries of meetings, workshops, and conferences, capturing key discussions, outcomes, and action points.
- Policy Papers: documents aimed at policymakers, summarizing key issues, research findings, and policy recommendations related to hydrogen.

Additionally, the project will collect personal contact data of stakeholders involved in or interested in the hydrogen policies and R&I to build a community of stakeholders around the IWG on hydrogen via Mailchimp. Contact data will include:

- Name (First and Last Name)
- Organisation
- Country
- E-Mail Address
- Stakeholder Group relevant to the project

The expected size of the data cannot be exactly estimated but will include data from multiple sources and types, implying a potential dataset of about 10 Gigabytes. Personal data will not to be shared outside the project as it would violate the General Data Protection Regulation (GDPR). Data Processing Agreements with all relevant data processors are put in place in compliance with GDPR provisions.

Type of Data	Users	Authentification/Verification	Data Privacy
Quantitative Data	Researchers, policymakers, civil society organizations	Access restricted to authorised users with ExtraNet Accounts; stored on DLR's internal SharePoint.	Data anonymised and managed per GDPR regulations; not shared outside project without consent.
Qualitative Data	Researchers, policymakers, civil society organizations	Access restricted to authorised users with ExtraNet Accounts; stored on DLR's internal SharePoint.	Data anonymised and managed per GDPR regulations; not shared outside project without consent.
Personal Contact Data	Project stakeholders, event registrants	Access restricted to project team; managed via Mailchimp and SharePoint.	Collected based on informed consent; includes name, organization, country, email address.
Insight Reports	General public, SET Plan community	Publicly available on project website and Open Research Europe platform.	No personal data included.
Fact Sheets	General public, SET Plan community	Publicly available on project website and Open Research Europe platform.	No personal data included.
Event Summaries	General public, SET Plan community	Publicly available on project website and Open Research Europe platform.	No personal data included.
Policy Briefs	Policymakers, researchers	Publicly available on project website and Open Research Europe platform.	No personal data included.

Table 1: Data Overview

2. FAIR Data

2.1 Making Data Findable, Including, Provisions for Metadata

Consistent file naming conventions will be applied to all data files. For example, file names will include the type of data (e.g., survey responses, interview transcripts), date of collection (YYYYMMDD format), and version number (e.g., "20240101_SurveyResponses_v1"). In addition, each document will entail detailed document information. An example:

Name	Data management Plan (DMP)
Deliverable Number	D1.5
Dissemination Level	PU-Public
Work Package	WP1: Organisation and Management
Task	T1.3 Monitoring, data management, risk management and quality assurance
Lead Beneficiary	DLR
Contact Person	Anne Dreysel, anne.dreysel(at)dlr.de
Version	v.1
Version Date	31. August 2024

Table 2: Example for information on deliverable

Metadata, if needed, will include relevant keywords related to SET Plan or hydrogen topics.

A version control system will be implemented to document and maintain different versions of data. This will ensure that any changes or updates to the data are tracked, and previous versions can be retrieved if necessary.

The project will use the Dublin Core metadata standard for any metadata information required.

2.2 Making Data Openly Accessible

Sensitive data will be stored on DLR's internal SharePoint, with access restricted to authorised users with an ExtraNet Account. Only employees from the project partners assigned to the CSA SET4H2 and active members of the IWG on hydrogen will receive ExtraNet Accounts.

For public dissemination, selected documents will be made available on the project's website under "Reports and Studies" as soon as their quality is assured and they do not contain sensitive information not meant for publication.

The collected data will also be registered with OPENAIRE. The policy related outputs (policy briefs, insight reports, fact sheets, etc.) will be published on Open Research Europe platform which provides

open access. We will follow the routes described in the EU guidelines on open access to scientific publications and research data.

The project will not produce new research data for a trusted repository. Instead, it will focus on re-using and sharing existing data through appropriate channels, such as the project website, ensuring that valuable data is widely and openly accessible.

Data or documents will also not be subject to an embargo to give time to publish or to seek protection of intellectual property. There is no need to ascertain the identity of persons accessing the data or to establish a data access committee due to the nature of this project.

2.3 Making Data Interoperable

Data will be formatted in widely-used standards such as CSV for quantitative data, ensuring compatibility with various data analysis tools. Qualitative data will be formatted in PDF or Word documents. Metadata will adhere to Dublin Core standards and use of standardised vocabularies.

2.4 Increase Data Re-use

Data will be shared under licenses that allow for re-use, such as Creative Commons licenses (e.g., CC BY 4.0), where applicable.

Comprehensive documentation will accompany datasets, including details about data collection methods, processing steps, and any limitations or considerations.

Data will be retained and maintained according to the project's duration and beyond, as needed. The project's website will remain online and be maintained for two years after the end of the project.

The quality assurance process for documents and data involves several structured steps to ensure thorough evaluation and adherence to set guidelines. The project coordinator conducts the final quality review, and the review process includes initial drafting by the task leader, partner feedback, revisions, and final coordinator review, ensuring compliance with editing standards, language clarity, and correct layout.

Disagreements during the review process require the author to respond to reviewer comments in writing, providing explanations if changes cannot be incorporated. If necessary, the project coordinator will mediate persistent disagreements by appointing an external reviewer. Formal assessment checks ensure compliance with corporate design and clarity, while content assessments evaluate alignment with objectives, the accuracy of the executive summary, logical content presentation, adherence to academic standards, appropriateness for the target group, correct dissemination level, and the exclusion of sensitive data in public deliverables.

3. Allocation of Resources

Allocations for data management are included in the overall project budget. Specific budget items will cover the development and maintenance of the SharePoint repository, the project website, and any required licenses or software.

Data management tasks will be undertaken by the project team members, including data collection, processing, and dissemination. Team members receive training to ensure adherence to best practices for data management and security.

4. Data security

Data will be stored on DLR's internal SharePoint, with access restricted to authorised personnel with an ExtraNet Account. Regular backups will be performed to ensure data integrity and availability.

Personal data will be managed in compliance with GDPR regulations. Access controls, encryption, and anonymisation techniques will be employed to protect data subjects' privacy, ensuring that personal data is securely handled.

Documents intended for public release will be published on the project's WordPress website, adhering to the latest security standards to protect against unauthorised access and breaches.

All accounts for the WordPress back-end, Mailchimp access, and SharePoint are password protected. Passwords will be changed regularly to enhance security, ensuring that accounts remain secure.

4.1 Additional Security Measures for DLR servers, including SharePoint

Unauthorised individuals are prevented from accessing data processing facilities where personal data is processed. Server hardware is physically located in a separate, rented room in a DLR data center in Cologne-Porz.

All accesses to server systems are logged. Access rights are assigned based on tasks and responsibilities, either at the system level (e.g., administrators) or programme level (e.g., software developers). Users typically receive access rights to specified files, data sets, or program modules as per application requirements.

Systems are ensured to be restorable in case of failure through comprehensive backup and recovery mechanisms. Daily backups of all relevant data on web and database servers are maintained. This ensures that data can be recovered in case of system failures.

A documented backup/recovery concept for the DMZ includes physical offsite backup storage and disaster recovery plans. Central IT of DLR provides escalation and emergency plans, applied within DLR. All data is backed up and restored as per a dedicated backup/restore concept tested for functionality. This ensures that data can be recovered in case of disasters.

Automatic backup and archiving of all employee email communication are conducted via DLR's central mail servers. This ensures that email communication is securely stored and can be retrieved if needed.

Data collected for different purposes is processed separately, ensuring compliance with privacy regulations.

4.2 Mailchimp Security Measures

Mailchimp implements an extensive range of security and data protection measures to safeguard user data. These include physical security at data centers, encryption of data in transit using TLS (TLS 1.2 or higher), and encryption at rest with AES-256. DDOS mitigation is present at all data centers. They conduct regular security audits and vulnerability assessments to identify and mitigate risks. Additionally, Mailchimp employs access controls to limit who can access sensitive data, segregates user accounts to prevent corruption and overlap and utilizes intrusion detection systems to monitor for unusual activity. The company also follows industry best practices for incident response and has enhanced measures following recent breaches. Finally, Mailchimp regularly conducts external and internal security tests, using different vendors, to test for vulnerabilities inside the application.

4.3 WordPress Security Measures

WordPress implements SSL certificates for securing data transferred between users and the website. The website is being backed-up regularly and a Web Application Firewall (WAF) protects the site by filtering and monitoring incoming traffic, blocking malicious requests before they reach the server. All WordPress sites are encrypted to ensure privacy and security. A security tool, Jetpack Scan, checks every WordPress.com site daily for dangerous plugins, themes, malware, and other vulnerabilities. If there are any issues, a security team will quickly resolve the issue by updating or reverting files, as needed.

User Access Control is properly managed. Assigning the least amount of privilege necessary and regularly reviewing user access can prevent unauthorised modifications and reduce the risk of security breaches.

The project will also ensure to keep WordPress core, themes, and plugins up to date to ensure security.

5. Ethical Aspects

Personal data collection will be based on informed consent, ensuring that data subjects are aware of the purposes and methods of data processing. Consent forms and information sheets will be provided to participants, detailing how their data will be used and protected.

Data will be anonymised where possible, particularly when published or shared beyond the immediate project team. This will involve removing or masking identifying information to protect data subjects' privacy.

6. Data Sharing and reuse

Data will be shared through the project website to maximize reach and impact. Regular updates will ensure that the information remains current and useful.

The project will engage with the SET Plan and hydrogen community through newsletters, social media, and events, encouraging the sharing and re-use of data and insights.

7. Data Preservation

Data necessary for the project's objectives will be preserved for the project duration and beyond, as necessary, to support ongoing and future research and policy-making efforts. This ensures that data remains available for future use.

The preservation strategy will include regular backups, migration of data to newer formats as needed, and ensuring that all data remains accessible and usable over time.

8. Review and Updates

The Data Management Plan will be reviewed and updated throughout the project lifecycle to ensure it remains aligned with project needs and best practices.

Scheduled reviews every six months will ensure the DMP evolves with the project and incorporates any new requirements or insights.


SET4H2