

HIGGS

Hydrogen in Gas Grids

A systematic validation approach at various admixture levels into high-pressure grids

D1.5

Data Management Plan

Date	30 June 2020
Grant Number	875091
Author(s)	Teresa Villuendas ¹ , Vanesa Gil ^{1,2} , Armin Bollien ³ , Pablo Benguria ⁴ , Luiz Carlos de Sousa ⁵ , Lola Storch de Gracia ⁶
	<ol style="list-style-type: none"> 1 Fundación para el Desarrollo de las Nuevas Tecnologías del Hidrógeno en Aragón (FHA) 2 Fundación Agencia Aragonesa para la Investigación y el Desarrollo (ARAID) 3 Deutscher Verein des Gas- und Wasserfaches e.V. (DVGW) 4 Fundación TECNALIA Research & Innovation 5 HSR Hochschule Fur Technik Rapperswil (HSR) 6 Redexis
	Author printed in bold is the contact person
Status	Started / Draft / Consolidated / Review / Approved / Submitted / Accepted by the EC / Rework [use bold style for current state]

Dissemination level:

PU Public

RE Restricted to a group specified by the consortium*

PP Restricted to other programme participants*

CO Confidential, only for members of the consortium*

*(including the Commission Services)



This project has received funding from the Fuel Cells and Hydrogen 2 Undertaking under Grant Agreement No. 875091. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and Hydrogen Europe and Hydrogen Research.

Document history

Version	Date	Description
1.1	2020-05-18	First draft of DMP
1.2	2020-06-15	Updated draft
1.3	2020-06-30	Final draft

The contents of this document are provided “AS IS”. It reflects only the authors’ view and the JU is not responsible for any use that may be made of the information it contains.

Table of Contents

Document history	2
Executive Summary	5
1 Objective	6
2 Description of work	7
2.1 Data summary.....	7
2.2 FAIR Data	8
2.2.1 Making data findable, including provisions for metadata	8
2.2.2 Making data openly accessible	10
2.2.3 Making data interoperable.....	10
2.2.4 Increase data re-use (through clarifying licences)	10
2.3 Allocation resources.....	11
2.4 Data security	11
3 Partner-specific data sets	12
4 Ethical aspects	16
5 Conclusions	17
Bibliography and References	18
Acknowledgements	19

List of Figures

Figure 1. WPs output and its relation with the DMP.	7
Figure 2. Naming procedure for HIGGS research data files.	9

List of Tables

Table 1. Template within PDF files associated to each experiment.....	9
Table 2: Data expected to be generated by DVGW.....	12
Table 3: Data expected to be generated by FHA.....	12
Table 4: Data expected to be generated by TECNALIA.	13
Table 5: Data expected to be generated by HSR.	14
Table 6: Data expected to be generated by DVGW.....	15

Executive Summary

As a Project taking part in the Open Research Data Pilot, HIGGS intends to develop a Data Management Plan (DMP), the first draft being the present document and to be updated during the project lifetime. This document also aims to pool results generated in the project that may lead to intellectual property (IP). HIGGS project will perform large amount of experimental and computational data which will cover the gaps of knowledge of the impact that high levels of hydrogen could have on the gas infrastructure, its components and its management.

This DMP will thus contain all forms of knowledge generated by the project. Whenever significant changes arise in the project, such as

- New data sets
- Changes in consortium policies
- New exploitable results
- External factors

a new version of the DMP shall be uploaded taking into account the major developments. In any case, the DMP shall be updated as part of the mid-term and final project reviews of HIGGS.

1 Objective

The main objective of this document is to outline a preliminary strategy for the management of the data generated in the framework of HIGGS project activities. DMP provides a structured form of repository for the consultation of data, measurements, facts and know-how gathered during the project, for the benefit of a more systematic progress in science. Where the knowledge developed in the EU-funded project is not governed by intellectual property for the purpose of commercial exploitation and business development, it is important to valorise the results of project activities by facilitating take-up of key data and information for further elaboration and progress by other projects and players in Europe. Procedures for the management of research data, technological data bases and market analysis data will be addressed.

2 Description of work

This document is developed in the framework of WP1 “Management and Coordination” activities and its policy will be defined in compliance with the “Guidelines on FAIR Data Management in Horizon 2020” and the guidance concerning Open Access and the Open Research Data Pilot at the European Research Council under Horizon 2020.

2.1 Data summary

HIGGS project main objective addresses the potential of hydrogen injection into the transmission high pressure natural gas grid as a way to decarbonise the gas system and gas uses.

As specific technical objective, a R&D platform will be developed in order to validate the technical, legal and regulatory conditions for safe operation of the gas transmission system. To this purpose; specific results about the mapping of legal, regulatory and technical aspects will be obtained as well as unique data from testing campaigns.

Data generated throughout the duration of the project may be classified as following (see Figure 1).

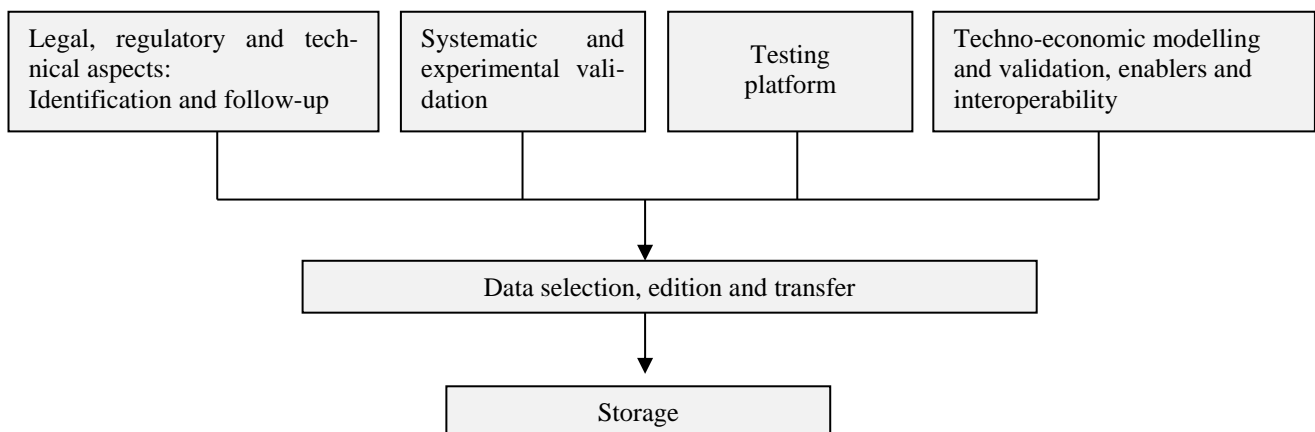


Figure 1. WPs output and its relation with the DMP.

The DMP is developed aiming at typical format of data storage like CSV or Excel file format, and it always should allow to document a wide range of experimental conditions, recording devices, input and output variables, etc.

It is important to highlight that new relevant data may be identified throughout project lifetime as essential to be shared together with the experimental validation results. In that case, this work will be updated during the project lifetime.

2.2 FAIR Data

Good data management is the key conduit leading to knowledge integration and reuse by the community after the data publication process.

In general terms, the EC guidelines suggest research data should be 'FAIR', that is Findable, Accessible, Interoperable and Re-usable.

While open access to research data thereby becomes applicable by default in Horizon 2020, the Commission also recognises that there are good reasons to keep some or even all research data generated in a project closed.

HIGGS data that is expected to be generated will be identified in section 3. Similarly, it will be analysed and justified any data affected by Industrial Property Rights (IPRs), third parties restrictions and other reasons to kept any result closed.

After the publication of any paper or public deliverable within HIGGS project framework, it is proposed to establish a deadline of **2 months** to deposit the data associated with them. This data (metadata and/or description files) will allow others beneficiaries to validate the results already published. There is no expected embargo period.

2.2.1 Making data findable, including provisions for metadata

EC “Guidelines for open access to publications, data, and other research outputs” highlight that quality-controlled research data has to be deposited in a repository as soon as possible ensuring:

- *Persistent and unique identifiers (PIDs)*
- *Long term sustainability*
- *Metadata o Curation and quality assurance*
- *Access (e.g. free and easy access to re-use)*
- *Security o Privacy o Common format*
- *Provenance (e.g. maintains a detailed logfile of changes to datasets and metadata)*

To this purpose, research data is expected to be shared in CSV or Excel format which data naming is proposed in Figure 2.

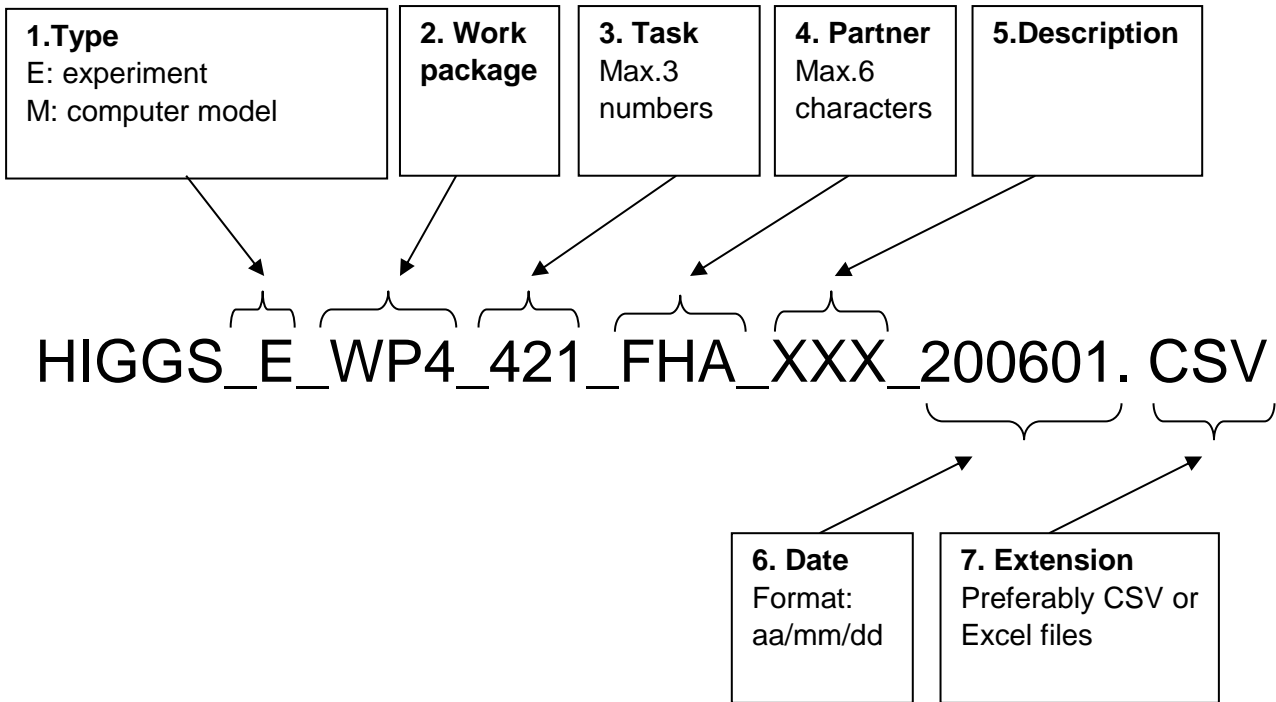


Figure 2. Naming procedure for HIGGS research data files.

Every .csv or .xls file will have an associated document with .pdf extension where summarised information of the experiment will be given. This file will gather the information showed in Table 1.

Table 1. Template within PDF files associated to each experiment

Field	Possible information
Data File name	Naming procedure in Figure 2.
Linked information	Work Package Task / Subtask Paper (DOI) Deliverable / Any public document
Summary	Short description of the experiment Scenario considered Keywords
Author	Main participants The relevant agencies attending the experiment The data provider (the person providing the experiment to database)

Boundary conditions	Experiment date and time Set point conditions (pressure, temperature...)
Experimental procedure	Description of the experiment
Experimental results	Figures that summarised the results if applied
Others	Measurement errors, or any note/event to allow a better understanding of the results.

2.2.2 Making data openly accessible

HIGGS selected final data will be deposited in ZENODO (<https://zenodo.org/>), an open access repository developed by CERN and linked to OpenAIRE (Open Access Infrastructure for Research in Europe). This repository is open for deposition of all type of data and from any stage of the research at no cost, reason why it is highly recommended in EC guidelines for institutions without its own repository.

The total files size limit per record is 50GB so it may be considered as suitable for HIGGS expected data outcomes.

2.2.3 Making data interoperable

HIGGS aims to ensure data interoperability, so every file will be shared in English language (though any language is allowed in the repository) and metadata will be given en SI base units for any measuring data.

Apart from the quality assurance resulted from the data processing, the data file will be as much clear as possible in terms of data labels, graphs, or additional notes when required.

Moreover, the metafile supporting the experimental data also ensure a better knowledge and interoperability of the results.

2.2.4 Increase data re-use (through clarifying licences)

Based on the fact that data will be introduced in the ZENODO repository, the data will remain re-usable as long as the repository operates. **Unless specified otherwise**, Zenodo metadata may be freely reused under the **CC0** waiver. CC0 enables scientists, educators, artists and other creators and owners of copyright- or database-protected content to waive those interests in their works and thereby place them as completely as possible in the public domain, so that others may freely build upon, enhance and reuse the works for any purposes without restriction under copyright or database law.

To ensure this re-usability of **HIGGS data**, the license model should be selected as much open as possible, **considering CC BY license as a possibility**. This license lets others distribute, remix, adapt, and build upon your work, even commercially, as long as they credit you for the original creation. This is the most accommodating of licenses offered. It is the license recommended for maximum dissemination and use of licensed materials.

2.3 Allocation resources

The data management of HIGGS project is expected to be free. HIGGS website may be used as additional repository where the last version of the DMP will be found as well as any data/document specified in section 2.5.

The selection and quality assurance of the data is responsibility of the partners conducting/involved in the experimental campaigns / simulations. To this purpose, they will follow the guidelines established in the last version available of the DMP and will send the files with the processed data to FHA, who will be in charge of the deposition, management and tracking of these files.

Every partner is committed to process their results and send them to FHA before the 2 months since the publication of the associated deliverable/publication/task deadline.

The use of an open repository for data as ZENODO, will allow maintaining the data as long as the repository is open and free of charge.

2.4 Data security

HIGGS data will be deposited in the repository after 2 months of each publication. Until a dataset is not fully finalised and ready for publication, each partner will be responsible of their data security, taking into account:

- If there is no restriction on sensitive data and there is consensus among the dataset contributors, other data exchange systems as email, are also allowed.
- To avoid data lost and once data is processed by each partner, coordinator (FHA) and the partners involved in each task will have a copy of the final files deposited/to be deposited in the server of their company.

Other issue to highlight is that Zenodo allows users to upload files under closed access. Closed access means that zenodo.org users will not be able to access the files you uploaded. The files are however stored unencrypted and may be viewed by Zenodo operational staff under specific conditions. This means that “closed access” on Zenodo is not suitable for secret or confidential data.

3 Partner-specific data sets

All partners involved in research’s activities were asked to provide detailed information about the data generated during the entire project, which is implemented in the tables hereafter.

Table 2: Data expected to be generated by DVGW.

Partner: DVGW		Work Package: WP2	
<i>What is the purpose of the data collection/generation and its relation to the objectives of the project?</i>		Create up-to-date information on present RSC, equipment and infrastructure of the high pressure natural gas grid	
<i>What types and formats of data will the project generate/collect?</i>		Publications, datasets	
<i>What is the origin of the data?</i>		Literature, survey with information from stakeholders	

Data/result identifier and nature (dataset, prototype, app, design, publication, etc.)	Origin (literature, experiments, analysis, modelling, etc.)	Restrictions (IP, patents, third parties, other)	Data/result storage (Zenodo repository, HIGGS website, none)
SoA RSC map incl. testing methods of the EU and its member countries, Infrastructure and equipment of the existing high pressure natural gas grid (Publications)	Literature, survey	tbd	Public deliverables on the homepage, share specific information with other projects/ stakeholders

Table 3: Data expected to be generated by FHA.

Partner: FHA		Work Package: WP3	
<i>What is the purpose of the data collection/generation and its relation to the objectives of the project?</i>		Design, preparation and commissioning of testing facilities. Manufacturing and lab-scale testing of hydrogen selective membranes.	
<i>What types and formats of data will the project generate/collect?</i>		<ul style="list-style-type: none"> - Diagrams and photos (test bench) - Numerical data concerning testing campaigns will be included in WP4. 	
<i>What is the origin of the data?</i>		Experiments and know-how obtained during the project lifetime.	

Data/result identifier and nature (dataset, prototype, app, design, publication, etc.)	Origin (literature, experiments, analysis, modelling, etc.)	Restrictions (IP, patents, third parties, other)	Data/result storage (Zenodo repository, HIGGS website, none)
--	---	--	--

D1.5 Data Management Plan

- Test bench: Block diagram, P&D HAZOP Material information Operating conditions Needs for adaptations, maintenance...	Know-how, literature	IP (D 3.1, D 3.3, D 3.4 and D 3.5)	TBD
Permits/licenses	Know-how, literature	IP (D 3.2)	
- Membrane gas separation module: Material and gas separation performance	Experiments and modelling	IP (D 3.1 and D 3.3)	
- Testing campaigns: Test duration Type of test			Discussed in WP3 but belonging to WP4 (D4.3, D4.4). Data from tests: Zenodo

Table 4: Data expected to be generated by TECNALIA.

Partner: TECNALIA	Work Package: WP4
<i>What is the purpose of the data collection/generation and its relation to the objectives of the project?</i>	Design, preparation and commissioning of testing facilities. Manufacturing and lab-scale testing of hydrogen selective membranes. Experiment design and testing plan Identification of materials and technical elements that potentially could suffer hydrogen embrittlement
<i>What types and formats of data will the project generate/collect?</i>	- Diagrams and photos (test bench) - Numerical data concerning hydrogen embrittlement testing and membranes performance
<i>What is the origin of the data?</i>	Experiments and know-how obtained during the project lifetime.

Data/result identifier and nature (dataset, prototype, app, design, publication, etc.)	Origin (literature, experiments, analysis, modelling, etc.)	Restrictions (IP, patents, third parties, other)	Data/result storage (Zenodo repository, HIGGS website, none)
Membrane gas separation module: Material and gas separation performance	Experiments and modelling	IP (D 3.1 and D 3.3)	Zenodo repository

Characterization results after hydrogen embrittlement experiments	Experiments		Zenodo repository
---	-------------	--	-------------------

Table 5: Data expected to be generated by HSR.

Partner: HSR		Work Package: WP5	
<i>What is the purpose of the data collection/generation and its relation to the objectives of the project?</i>		<p>Evaluation report on existing modelling software containing the results of the assessment and an appraisal its suitability to the project's purpose.</p> <p>Techno-economical Model to study the impact of hydrogen injection into HPGN under different scenarios, operating conditions and network configurations.</p> <p>Modelling results will document the operational and business impacts of the above for further research or operational purposes.</p> <p>All of the above are required to fulfill project objective O6.</p>	
<i>What types and formats of data will the project generate/collect?</i>		<p>a) The techno-economical model itself Format: a programming language yet to be determined).</p> <p>b) The resulting calculated physical simulation data (flowrates, compositions and heating values, pressures etc.) of the gas in the network. Format: excel files (t.b.c.)</p> <p>The resulting calculated economic indicators notably costs and prices to consume. Format: excel files (t.b.c.)</p>	
<i>What is the origin of the data?</i>		Know-how obtained during the project lifetime and simulation platform.	
Data/result identifier and nature <i>(dataset, prototype, app, design, publication, etc.)</i>	Origin <i>(literature, experiments, analysis, modelling, etc.)</i>	Restrictions <i>(IP, patents, third parties, other)</i>	Data/result storage <i>(Zenodo repository, HIGGS website, none)</i>
Evaluation report on existing modelling software (D5.1)	Know-how, literature		HIGGS website (?)

Techno-economical model of HPGN with H2 (D5.2)	Know-how, literature	IP of third party may be involved, depending on chosen solution ¹⁾ , notably if model is extension of existing software D5.2 is confidential to the consortium.	D5.2 is confidential to the consortium.
Physical simulation data	Know-how, literature		

Notes:

1) Part of the work in WP5 is to evaluate existing Software to decide on writing new software either as a standalone or as an extension to existing software. In case an extension is selected, IP of a 3rd party (software proprietor) may need to be considered.

Table 6: Data expected to be generated by DVGW.

Partner: DVGW		Work Package: WP6	
<i>What is the purpose of the data collection/generation and its relation to the objectives of the project?</i>		Establish a compilation of the main findings and assessments during the project incl. a pathway and recommendations summary to enable wider injection of H2 in EU gas networks	
<i>What types and formats of data will the project generate/collect?</i>		Publications, datasets	
<i>What is the origin of the data?</i>		Literature, results of the project	
Data/result identifier and nature <i>(dataset, prototype, app, design, publication, etc.)</i>	Origin <i>(literature, experiments, analysis, modelling, etc.)</i>	Restrictions <i>(IP, patents, third parties, other)</i>	Data/result storage <i>(Zenodo repository, HIGGS website, none)</i>
Summarisation of main findings; Pathway and recommendations (Publications)	Literature, results of the project	Tbd	Public deliverables on the homepage, if applicable sharing information with stakeholders

4 Ethical aspects

Any issues identified.

5 Conclusions

The Data Management Plan (DMP) aims to provide a structured form of repository for the consultation of data, measurements, facts and know-how gathered during the project, for the benefit of a more systematic progress in science.

To this purpose, every partner is responsible to select and process the data generated within the tasks they are involved to create the files that will be shared with the scientific community.

This work will be carried out following the roadmap given in the DMP, which is intended to be a living document where information can be updated as the project advances and additional needs arise.

Bibliography and References

- [1] “ZENODO repository,” [Online]. Available: <https://zenodo.org/>.
- [2] European Commission, “Guidelines on FAIR Data Management in Horizon 2020,” 2016 (Version 3.0).
- [3] European Commission, “Guidelines for open access to publications, data and other research outputs,” 2020.
- [4] European Commission, “Horizon 2020 online manual,” [Online]. Available: https://ec.europa.eu/research/participants/docs/h2020-funding-guide/grants/grant-management/dissemination-of-results_en.htm.
- [5] M. Wilkinson, M. Dumontier, I. Aalbersberg and a. cols, “The FAIR Guiding Principles for scientific data management and stewardship,” *Sci Data*, vol. 3, no. 160018, 2016.
- [6] “Creative Commons,” [Online]. Available: <https://creativecommons.org/>.

Acknowledgements

This project has received funding from the Fuel Cells and Hydrogen 2 Undertaking under Grant Agreement No. 875091. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and Hydrogen Europe and Hydrogen Europe Research.