This project is supported by the Clean Hydrogen Partnership and its members.

BalticSeaH2

Cross-border Hydrogen Valley around the Baltic Sea

D1.1 Data Management Plan 1st version

WP1 – Project Coordination and Management

Deliverable Lead	Merie Joseph Kannampuzha & Oleg Todorov (CLIC Innovation)
Deliverable Due Date	30.9.2023
Actual Submission Date	29.9.2023
Version	Final D1.1

BalticSeaH2 Project GA ID: 101112047

Co-funded by the European Union

Document information

Grant Agreement number:	**101112047**			
Project acronym:	BalticSeaH2			
Project full title:	Cross-border Hydrogen Valley around the Baltic Sea			
Start date:	1st June 2023			
Duration:	60 months			
Deliverable title:	Data Management Plan 1st version			
Work Package:	WP1			
Date of delivery:	*Contractual:*	30.9.2023	*Actual:*	29.9.2023
Type:	DMP			
Dissemination level:	PU - Public			
Lead beneficiary:	1 CLIC			
Responsible author:	Merie Joseph Kannampuzha & Oleg Todorov			
Contributions from:	CM, WP1 team, all partners			
Target audience:	[x] Public [ ] Restricted to other programme participants (including the Commission Services) [ ] Restricted to a group specified by the consortium (including the Commission Services) [ ] Confidential, only for members of the consortium (including the Commission Services)			

Document history

Version	Issue date	Stage	Description	Contributor
First shared draft	21.9.2023	Draft for comments	Version sent for inputs and comments to CM	Oleg Todorov & Merie Kannampuzha
Final	29.9.2023	Comments integrated	Final version with inputs from WP leaders and partners added	Merie Kannampuzha

Co-funded by the European Union

Contents

Acronyms and abbreviations

Acronym	Full name
AAL	Aalto Korkeakoulusäätiö Sr
ABB	ABB Oy
BIA	Baltic Innovation Agency OÜ
BOR	Borealis Polymers Ltd
BPSE	Ballard Power Systems Europe
BSR	Baltic Sea Region
CDG	Communication and Dissemination Group
CDP	Communication and Dissemination Plan
CKAN	The Comprehensive Knowledge Archive Network
CLIC	CLIC Innovation Ltd
CON	Convion Ltd
CYB	Cybernetica AS
D	Deliverable report
DMP	Data Management Plan
DMS	Data Management System
DPO	Data Protection Officer
EAB	External Advisory Board
ECD	Energy Cluster Denmark
EF	Energiforsk AB
EHC	Estonian Hydrogen Cluster

Co-funded by the European Union

ELC	Elcogen Ltd
ELO	Elomatic Ltd
EOSC	European Open Science Cloud
EVN	Energy Valley Norway
FAIR	Findable, Accessible, Interoperable, Reusable
FLX	Flexens Ltd
GA	Grant Agreement
GDPR	General Data Protection Regulation
GG	Gasgrid Finland Ltd
GNH2	Green North Energy Ltd
GSTC	Green and Smart Technology Cluster Latvia
H&Ko	OÜ Hendrikson & Ko
HLN	Helen Ltd
IWEN	IWEN Energy Institute GmbH
KBFI	National Institute of Chemical Physics and Biophysics
KPI	Key Performance Indicator
LHY	Lhyfe Labs
NTE	Neste Oyj
OEP	Open Energy Platform
OpenAIRE	Open Access Infrastructure of Research in Europe
P2X	P2X Solutions Ltd
PHC	Polish Hydrogen Cluster
PID	Personal Identifier
PtX	Power-to-X
PU	PowerUP Fuel Cells OÜ
RBP	Freeport of Riga Authority
RH	Rønne Havn A/S
RISE	RISE Research Institutes of Sweden AB
SRS	Skyborn Renewables Sweden AB
SS	Solarstone
TLN	City of Tallinn
UU	Uppsala University
VEA	Lithuanian Hydrogen Energy Association
VEN	Vantaa Energy Ltd
VKL	Viking Line Abp
VTT	VTT Technical Research Centre of Finland Ltd
VV	MTÜ Viru Vesinik
WLA	Wärtsilä Finland Ltd
WP	Work package
YRA	Yara Finland Ltd

Co-funded by the European Union

Executive summary

This document presents the data management plan for BalticSeaH2 project. BalticSeaH2 Data Management Plan (DMP) is prepared by CLIC for delivery by M4 (D1.1) based on the Horizon Europe Data Management Plan Template. This deliverable is part of WP1 Project management and coordination. CLIC is the responsible partner for the DMP. All WP leaders and project partners will contribute to implementing the data management plan described here. This plan is intended for the internal use of the BalticSeaH2 Consortium as well as for the purpose of evaluation of the data management activities of the project by external stakeholders.

BalticSeaH2 generates data from the assessment tasks, the feasibility studies, the demo cases of hydrogen production, transmission and end use as well as from the project management activities, stakeholder engagement, valley replication and communication and dissemination activities. During this project, data will be collected using multiple methods such as measurements, stakeholder workshops, research outputs as well as questionnaires. The data will be collected and analysed using relevant tools and software for each WP. The DMP defines the Consortium's strategy for effective and ethical data archiving and sharing, Personal Data Management and general data management guidelines. The DMP will be updated at least once per reporting period. The DMP will be GDPR compliant and operational under the FAIR principles.

1. Introduction

This document describes the Data Management Plan (DMP) for BalticSeaH2 project. The DMP defines the Consortium's strategy for effective and ethical data archiving and sharing, Personal Data Management and general data management guidelines. The DMP will be updated at least once per reporting period.

BalticSeaH2 generates data from the assessment tasks, demo cases as well as management, valley replication and communication and dissemination activities. During this project, data will be collected using multiple methods such as measurements, stakeholder workshops, research outputs as well as questionnaires. The data will be collected and analysed using relevant tools and software for each WP. The data Frelated activities of BalticSeaH2 will be driven by the FAIR data principles. All data related activities in BalticSeaH2 will be GDPR compliant. Ethics section of this document details on the personal data management.

The Consortium has appropriate technical and organizational measures in place to carry out data management and protection during the project. A centralized secure storage is provided in Microsoft Teams folder of BalticSeaH2 for all internal project data, reports and documents. All project Consortium members will have access to this data. BalticSeaH2 valley will be one of the cross-border Hydrogen valleys in Europe. Hence the Consortium is strongly committed to share learned experiences and best practices to enable replication of Hydrogen valleys in the rest of Europe by open science and access. All public deliverables and reports of the project will be published promptly and placed in a public repository. When applicable, public repositories like European Open Science Cloud (EOSC), Open Access Infrastructure of Research in Europe (OpenAIRE), Zenodo, CKAN (The Comprehensive Knowledge Archive Network, OpenRemote (A powerful open source IoT-platform) or Open Energy Platform (OEP) will be utilized to share public outputs from the project.

Co-funded by the European Union

2. BalticSeaH2 Data summary

BalticSeaH2 generates data from the assessment tasks, demo cases as well as management, valley replication, and communication and dissemination activities. Data are collected throughout the project to build the H2 market pool concept. In BalticSeaH2, the following data will be generated:

1. Data from the project management tasks (meetings).
2. Data from stakeholder engagement activities, desktop studies, assessment reports, feasibility studies and other research outputs.
3. Data from surveys and validation studies.
4. Data from the demo cases for hydrogen production, transportation, and utilization.
5. Data from modelling, optimization and software for hydrogen commercialization, production and storage.
6. Data from communication and dissemination activities of the project.

Main types of data in BalticSeaH2 include qualitative data, quantitative data, personal data, digital image and video data as well as social media posts and various reports which are deliverables from the project. In addition, raw data will be generated from measurements and experiments. The main formats of data will be measurement data from project demos in MS Excel format, CSV format, documents, and reports in the formats of MS Word, MS Power Point, Adobe Acrobat PDF as well as image and video formats like png, jpeg, mp4. To support project management and carry out various events related to BalticSeaH2, personal data might be collected. The Consortium members will try to minimize the amount of personal data collected.

The datasets generated in BalticSeaH2 will be processed and analysed using relevant software. The data analysed in connection with the project work and used to develop the H2Pool market concept will be generated by the various project activities, surveys, questionnaires, workshops and the demo cases. However, the possibility of reutilizing any existing relevant open data will also be examined carefully. Potential reutilization of any open data will be ensured by careful documentation of datasets as well as description and publication of data collection methods, protocols, workflows, and models.

A summary of data generated and collected in BalticSeaH2 as well as how the data will be handled is presented below in the Table 1. This table reflects how it is foreseen at the beginning (M4) of the project. Any changes arising during the five-year period of the project will be updated in each reporting period.

Co-funded by the European Union

Table 1 BalticSeaH2 project - data generated, project activities and used software foreseen at the start of the project.

Type of data	Nature of data	Format of the data	Open access	BalticSeaH2 activities and used software foreseen at the start of the project
Raw data	Data from monitoring	TXT, CSV or compatible format	NO	Data from measurements & experiments.
Quantitative data	Data from the project demo cases and stakeholder workshops	Excel, CSV or compatible format	NO, shared among project partners when applicable	WP1: Quantitative data to measure the technical, environmental, economic, societal, regulatory, and scientific KPIs. WP2: Quantitative survey data collected for diagnosis of country profile overview; Quantitative data collected from stakeholders for development of BalticSeaH2 visions. WP3: Quantitative data collected to carry out assessment of optimal factors for cost efficient hydrogen production and spatial- temporal analysis of freshwater availability; Quantitative data from comparison of various electrolyser technologies; Data from the evaluation study on large scale hydrogen storage and electrochemical hydrogen compressor; Numerical details from the Hydrogen production use cases including financial feasibility, environmental impact, and performance metrics. WP4: Quantitative data from the Hydrogen consumption use cases in industry, marine, road transport and energy sectors as well as built environment to carry out economic feasibility analysis, environmental analysis as well as performance and technical analysis. Wp5: Use case data external to the project for renewable energy assessment for hydrogen production; Quantitative data from sector coupling use cases, PtX applications and grid balancing; Quantitative data for modelling sector integration framework. WP6: Quantitative data from project Consortium members for Hydrogen valley market model, data from digital twins; Online H2Pool marketplace framework KPI metrics. WP7: Market analysis data on hydrogen opportunities, Impact assessment data on hydrogen economy development.

				WP8: Communication metrics from different channels where shared.
Qualitative data	Interview, surveys, questionnaires, results from workshops	MS Excel, Word, Power Point	YES (project partners) when public	WP1: Qualitative data to measure the technical, environmental, economic, societal, regulatory, and scientific KPIs. WP2: Qualitative survey data collected for diagnosis of country profile overview; Interview data collected from stakeholders for development of BalticSeaH2 visions, Use case data from other EU Hydrogen valleys and relevant hydrogen initiatives. WP3: Data collected for study of reutilization of existing natural gas pipelines with the scenarios of building new hydrogen pipelines; Data collected for analysis of interconnection of valleys in BSR. WP4: Qualitative data from the Hydrogen consumption use cases in industry, marine, road transport and energy sectors as well as built environment to carry out economic feasibility analysis, environmental analysis as well as performance and technical analysis. Wp5: Qualitative data from sector coupling use cases and PtX applications and grid balancing. WP6: Qualitative data from project Consortium members for Hydrogen valley market model. WP7: Qualitative data from assessment of potential impact creation and replication. WP8: Qualitative data from communication specific questionnaires or interviews and events.
Personal data	From the staff working with BalticSeaH2 project, Consortium members	Name, Affiliation, Email, Phone number (optional), Roles in the project / organization, Project WP responsibilities,	NO	The collected personal information will be treated as confidential and will be handled in accordance with data protection and regulation (GDPR). WP1: Personal and contact details will be collected for the project management and coordination from the staff working with the project.

		Shipping and mailing addresses of each partner organization		
Personal data	From participants of EAB, project workshops, strategic weeks, surveys, questionnaires, and interview	Name, Affiliation, Email, Preferred choice of food and information on dietary restrictions for physical meeting	NO	The collected personal information will be treated as confidential and will be handled in accordance with data protection and regulation (GDPR). WP1: Personal and contact details of external advisory board members. WP2: Contact details of stakeholders from the quadruple helix model groups and community members to engage them in workshops and the development of BalticSeaH2 visions. WP3: Not foreseen. WP4: Not foreseen. WP5: Not foreseen. WP6: Not foreseen. WP7: Contact details collected as part of external study visits WP8: Data relevant for the dissemination, and communication activities, such as mailing and participant lists, with the participants' consent.
Digital image and video data	Photographic images, videos	jpeg, mp4 and other imagery and video formats	YES	WP1: Videos and digital images from management events will be collected for communication and dissemination activities. WP2: Video and images from the on-site strategic week, stakeholder workshops, webinars, public information days and private sector seminars and networking events; Use case videos, animations for public information days. WP3: Videos and images to showcase the production use cases of Hydrogen. WP4: Videos and images to showcase the consumption use cases of Hydrogen.

				WP7: Videos and images during Foresight visioning workshops and knowledge sharing internal and external study visits. WP8: Visual identity, templates and logo, Podcasts, WP specific videos and infographics, Videos and digital images taken during policy briefs, roundtable discussions, regulatory nodes, matchmaking activities with funders, social awareness and acceptance activities and other physical and online events for communication and dissemination. The consent of participants will be asked in order to use it for project website, social media posts, press releases, communication materials and e-newsletters as detailed in the CDP.
Social media	Social media posts and blog posts	Text, images, or other social media format	YES	WP2: Use case social media and blog posts for public information days. WP3: Social media posts to showcase the production use cases of Hydrogen. WP4: Social media posts to showcase the consumption use cases of Hydrogen. WP8: Social media posts in relevant social media accounts (e.g., LinkedIn, X (Twitter)) and blog posts generated for communication and dissemination of project activities.
Documentation	Project deliverables, assessment, feasibility study reports, research outputs. CAD/CAM documentation (converted to PDF)	Word, PDF, PPT	YES, when public	WP1: Public deliverables D1.1, D1.2, D1.3 (M48 version of the safety plan), D1.4 (M60 version of the evaluation handbook). WP2: Public deliverables D2.1, D2.4 (M60 version of the BalticSeaH2 valley's vision document), D 2.5, D 2.6 (M60 version of the public awareness report). WP3: Public deliverables D 3.1 (M36 version of the assessment report), D 3.2, D 3.3, D 3.4, D 3.5 (M55 version of the H2 quality standards report), D 3.6. WP4: Public deliverables D4.1, D 4.2, D4.3, D 4.4, D4.5, D4.6. WP5: Public deliverables D5.1, D5.2, D5.3, D5.4. WP6: Public deliverables D 6.1, D 6.2, D 6.3.

				WP7: Public deliverables D 7.1, D 7.2, D 7.3, D 7.4, D 7.5, D 7.6. WP8: Public deliverables D 8.2, press releases, position papers, factsheets, presentations, online brochures, WP specific leaflets and posters and other relevant documents. Data for deliverables will be reviewed by Consortium before publishing according to Consortium agreement (publications 30 days and presentations 10 days prior publishing)

Co-funded by the European Union

3. FAIR data

EU Commission recommends that project research data should follow the FAIR data principles, meaning that data is Findable, Accessible, Interoperable and Reusable. The data management plan (DMP) of BalticSeaH2 will be influenced by the FAIR data principles, which basically describe how research outputs should be organised so they can be more easily accessed, understood, exchanged and reused. The project DMP additionally highlights what type of data will be generated, if and how it will be made accessible for verification and reuse, how to ensure data quality and reusability, how will metadata be generated, and how the data will be curated and preserved.

Making data findable, including provisions for metadata

Persistent identifiers (PIDs) are essential e.g., for research output, because they unambiguously identify the data and facilitate data citation. An example of a PID is a Digital Object Identifier (DOI), normally used in scientific publications. When depositing project data in a repository, a repository that assigns a persistent identifier (for example Zenodo) should be considered.

To be able to discern and easily identify data sets, each data set will be assigned with an unique name. This name can also be used as the identifier of the data sets. All data files produced, including emails, will include the term “BalticSeaH2”, followed by file name which briefly describes its content, followed by a version number (or the term “FINAL”), and followed by the short name of the organisation which prepared the document (if relevant). Each data set that will be collected, processed, or generated within the project will be accompanied by a brief description.

Another important feature is the metadata describing project research data, since it supports findability, citation and reuse. Rich metadata provides important context for the interpretation of your data and makes it easier for machines to conduct automated analysis. The following general metadata structure is proposed for project documents and data sets:

General overview

- Unique name / ID
- Work package
- Resource type
- Title
- Version
- Date of creation
- Creator
- Contributors
- File location

Content description

- Description and relation to the project objectives
- Abstract / Keywords
- Geographical location
- Code list used
- Metadata schema used

Co-funded by the European Union

Technical description

- Software used to create the file
- File format
- Necessary software
- File size

Sharing and preservation

- Use / Users
- Rights / Licence
- Dissemination
- Access
- Restrictions
- Repository for open data
- Permanent identifier (e.g., DOI, URN)

Accessibility of data from the project

During the project data will be shared and used by BalticSeaH2 Consortium members. Any decision regarding sharing of datasets to third parties will be made by the project Management Committee (MC). The project coordinator and project participants will take all the appropriate measures to ensure availability and usability of data for third parties for research and study purposes.

Data and research outputs for public use will be disseminated and shared as early and as openly as possible. Both gold and green approaches will be used for publications, depending on the circumstances. Agreement on the ownership and management of the results will be defined in WP1 management and WP8 exploitation and business planning activities. In line with Horizon rules, knowledge will be owned by the partner(s) that generate(s) it.

Project Consortium can propose the creation of Data taskforce and Data access Committee, in case the amount of generated data becomes relevant, and their accessibility could potentially create bottlenecks in normal Project management, operation and dissemination.

If data reuse requests need to be approved after the project duration, the request will be considered by the data owner and the Management Committee. The permission for data reuse is granted only if there are no conflicting confidentiality and IPR issues.

Data which are publicly available will be made available in repositories. When applicable, public repositories like European Open Science Cloud (EOSC), Open Access Infrastructure of Research in Europe (OpenAIRE), Zenodo, CKAN (The Open-Source Data Management System), OpenRemote (A powerful open source IoT-platform) or Open Energy Platform (OEP) will be utilized to share public outputs from the project.

Making data from BalticSeaH2 project interoperable

Data which can be integrated with other data, applications and workflows are interoperable. That means not creating data with proprietary software and making it available in open formats. The utilization of community agreed schemas, controlled vocabularies, keywords, thesauri or ontologies where possible, is recommendable.

Co-funded by the European Union

Interoperability of data/research outputs

An early-stage identification of standards (as defined in the Project Plan):

1) Microsoft Word 2010 for text-based documents (or any other compatible version). doc, .docx, .xls, .xlsx, .ppt, .pptx, .msg. Also, especially where larger datasets need to be dealt with, .csv and .txt file formats will be used. All finished and approved documents will also be made available as .pdf documents.
2) Illustrations and graphic design will be made available as .jpg files. Computer Aided Design with formats: .dwg that will be also converted to .pdf documents. Video files in formats: .mkv, .avi, .mpeg, .mov recorded and received through camcorders or dedicated advertisement movies completed in different video editor software. Other CAD and CAM standards that are currently undefined.

Increasing data reusability

The digital data generated and processed by the project team will be used in different ways, depending on the policies and exchange requirements attached to it. For both public and private data, the goal is to preserve the data and make it available to project participants throughout the project and beyond. Database compliance is forced to ensure that security policies are correctly implemented in databases that validate vulnerability, wrong and corrupted data.

Increase data reuse (through clarifying licences)

After the project completion, the ownership of datasets will belong to the grant beneficiaries that generated them. Creative Commons license CC-BY-SA or CC-BY or similar public copyright license will be used for any opened datasets, unless there are compelling reasons to select more restricted type of public license. Creative commons licenses will by default include also a disclaimer of liability for the reuse of open data.

4. Allocation of resources

All Consortium members will ensure that the data shared are accordant with FAIR principles, quality controlled and as open as possible in applicable contexts.

The costs for data curation and storage during the project is included in the budget for project management. Each of the Consortium members will be responsible for covering the costs of generating, curating and storing data corresponding to their demo cases from their allocated project budget. Following the finalization of the project, each Consortium member will be responsible for the costs of longer-term storage and ensuring access of data which is open.

5. Data security

The task forces responsible for T1.2 data management and T1.3 quality assurance and risk management activities will take overall responsibility for data management and quality assurance. CLIC as the coordinator will lead these tasks, ensuring that the DMP is carried out, with crucial input from other partners throughout the project.

Co-funded by the European Union

For the duration of the project, datasets will be stored on the responsible partner's storage system. Every partner together with their IT teams is responsible to ensure that the data are stored safely and securely and in full compliance with EU data protection laws. At the end of the project, each Consortium member will be responsible to ensure for long term storage, curation and sharing of data sets which are open.

All data structures used by the project will include a secure protection of sensitive data. Third party solutions will be carefully vetted to ensure the safety and security of data. Transport layer security and transparent data encryption will be used. Database compliance is forced to ensure that security policies are correctly implemented in databases that validate vulnerability, wrong and corrupted data.

6. Ethics

BalticSeaH2 is compliant with EU, international and national ethics requirements and legislation and ensures that the provisions on ethics regulation and rules are implemented into all project activities. BalticSeaH2 project will follow the guidelines of Regulation (EU) 2016/679 (General Data Protection Regulation) for handling personal data to ensure privacy and rights of participants.

For project management purposes, the contact details of individuals working in the project are collected. In addition, contact details of individuals who are participating in the project external advisory board will also be collected. These data are stored in the Microsoft Teams excel file in the project Teams folder maintained by CLIC and not shared to the wider public. The contact details of organizers of public workshops are shared with the prior consent in the communication materials to register for the events.

Personal data may also be collected when meetings or workshops are arranged related to the stakeholder engagement activities. Informed consent procedures will be followed and specific safeguards implemented if the groups to be involved in these activities could be considered so called vulnerable groups. In case, ethics approvals, opinions or statements are required in these more specific cases, the action cannot be started before the received decision. The personal data are not shared with the Consortium but kept confidential by the partner organizing the meeting or an event.

BalticSeaH2 will produce videos which may include the staff affiliated to institutions which are project Consortium members who will introduce and explain the role of their affiliated institutions in the project. In such cases, the Consortium members will be presented in the movie with their names, affiliations, and positions in the project. Prior consent will be sought from the respective person before publishing the videos.

Research integrity and ethical principles related to data collection and use are covered in detail in the ethics section 14.1 of the Grant Agreement. Ethical aspects will be considered by all Consortium participants and monitored by the task 1.6 leader VTT in collaboration with CLIC.

7. References

EC, 2016: Horizon Europe Data Management Plan Template v1.0 – 13.10.2016

Co-funded by the European Union

EC, 2016: General Data Protection Regulation (Regulation (EU) 2016/679 of the European Parliament and of the Council)

8. Funding statement

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 Clean Hydrogen Joint Undertaking. Neither the European Union nor the granting authority can be held responsible for them.