WaterWise Hub

Deliverable 1.2

Data Management Plan (DMP)

Funded by
the European Union

GRANT AGREEMENT NUMBER	101184151
FULL TITLE / ACRONYM	THE WATERWISE HUB: AN EXCELLENCE HUB ON WATER IN THE CIRCULAR ECONOMY HUB
START DATE	01 – 01 – 2025
END DATE	31 – 12 – 2028
PROJECT URL	www.waterwisehub.eu
WORK PACKAGE No and title	WP1 – Coordination and Project Management
DELIVERABLE TITLE	D1.2 – Data Management Plan (DMP)
ACTUAL DATE OF DELIVERY	31 March 2025
DISSEMINATION LEVEL	Public
LEAD BENEFICIARY	NTUA
RESPONSIBLE AUTHORS	Alexandros Kritikos, George Bariamis
CONTRIBUTIONS FROM	Alexandros Kritikos, George Bariamis George Karavokiros

Document History

Version	Description	Author	Organisation
V 0.1	First Draft	Alexandros Kritikos, Georgios Karavokiros	NTUA, TETHYS
V 0.2	Review	Georgios Bariamis	NTUA
V 0.3	Final Draft	Alexandros Kritikos, Georgios Karavokiros	NTUA, TETHYS
V 1.0	Final Version	Georgios Bariamis	NTUA

Quality Control

Author	Organisation	Role	Date
Georgios Karavokiros Georgios Bariamis	TETHYS, NTUA	Data Manager Deliverable Leader	21-3-2025
Prof. Christos Makropoulos	NTUA	Scientific Project Coordinator	24-3-2025
Georgios Bariamis	NTUA	Project Coordinator	28-3-2025

Contents

Abbreviations

AES	Advanced Encryption Standard
CC-BY	Creative Commons Attribution License
CC0	Creative Commons Zero (public domain dedication)
CERN	European Organization for Nuclear Research
CSDM	Case Study Data Manager
CSV	Comma-Separated Values
DMP	Data Management Plan
DOI	Digital Object Identifier
EOSC	European Open Science Cloud
EO	Expected Outcome
FAIR	Findable, Accessible, Interoperable, and Reusable
GIS	Geographic Information System
GDPR	General Data Protection Regulation
GPL	General Public License
HDF5	Hierarchical Data Format version 5
IP	Intellectual Property
IPR	Intellectual Property Rights
JSON	JavaScript Object Notation
MIT	Massachusetts Institute of Technology License
NetCDF	Network Common Data Form
NGO	Non-Governmental Organization
NTUA	National Technical University of Athens
OAI-PMH	Open Archives Initiative Protocol for Metadata Harvesting
RBAC	Role-Based Access Control
R&I	Research and Innovation
SHP	Shapefile format (in GIS)
SSL/TLS	Secure Sockets Layer / Transport Layer Security
TDM	Technology Data Manager
VPN	Virtual Private Network
WE	Water Europe
XML	Extensible Markup Language

Executive Summary

The WaterWise Hub is part of Europe's initiative on Excellence Hubs, focusing on sustainable water management in the Circular Economy. The goal of the project is to improve sustainability of water use and resource efficiency by bringing together research, innovation and business development while proliferating the knowledge in this new paradigm of water management. By utilizing the quadruple helix (academia, industry, government and society) it aims to propel innovation and research for achieving a fast and significant impact on its domain.

The current Data Management Plan outlines the policy and practices for handling the data utilized within the project, throughout its lifetime. At its core, the DMP is established upon the FAIR principles of data management, aligning with European Union's Open Science and Data Protection mandates, including GDPR compliance.

The WaterWise Hub will collect, generate and re-use data from various sources (research, models, stakeholders, regulatory bodies etc.) and of various types (text, spreadsheet, software-parsed datasets etc.). All the data will be stored in secure, trusted, open repositories (OpenAIRE, Zenodo community) as well as proprietary platforms that provide additional functionalities as well (Water Europe Marketplace). Abiding to accessibility, all datasets will be accompanied by Persistent Identifiers (PIDs) and preserved for a minimum of 10 years beyond the project's completion. Regarding GDPR, any personal data will be anonymized and aggregated to ensure compliance.

Taking into consideration the sensitive data of the project (e.g. related to Intellectual Property Rights), data access will be allowed through Role-Based Access Control (RBAC) to assure that only the appropriate authorized users can access (view, edit or manage). Furthermore, regular security audits and backups will be conducted, serving as an additional assurance on the matter.

Through this comprehensive Data Management Plan, the WaterWise Hub plans to maximize its impact on research, policy development and business innovation in sustainable water management while the refined and potent database created from the project is expected to contribute in the long-term towards water resource sustainability and innovation across Europe.

The DMP is intended to be a dynamic document and will be updated as needed throughout the project.

1. Introduction

1.1 Purpose of the Data Management Plan

In accordance with the standards of Horizon Europe funded programmes, a Data Management Plan (DMP) is to be established for the proper management of all data relevant to the WaterWise Hub project. The DMP will serve as a comprehensive framework for handling the produced and utilized data, according to EU's Open Science and Data Protection mandates. The current document and plan will ensure that all the data-related activities are complying with the FAIR principles (Findable, Accessible, Interoperable and Reusable), facilitating the responsible data collection, processing and sharing inside the project Consortium, as well with the general research community.

In detail, the DMP provides a comprehensive methodology and guidelines to managing data throughout the implementation of the project, thoroughly addressing issues such as data generation, storage, protection, accessibility, and long-term preservation. Clarifications are provided on technical data-handling matters such as metadata standards, open-access repositories, and security protocols in order to enhance data reusability and interoperability. Additionally, contextual matters to data production, storage and utilization are addressed, such as adherence to the EU General Data Protection Regulation (GDPR) and ethical standards in general.

By following this DMP, the WaterWise Hub will be able to streamline the management of data through the project and maximize the impact of its research outputs, by facilitating dissemination of knowledge and thus further supporting research and innovation in sustainable water management. As a live document, it will be regularly updated through the project's gradual implementation in order to include potential new types of data that will be produced and identified in later stages.

1.2 Overview of the WaterWise Hub Project

The WaterWise Hub is an Excellence Hub dedicated to the advancement of water management in the era of the Circular Economy. The general goal of the Hub is to improve water sustainability and resource efficiency through catalyzing Research and Innovation, Business Development and Knowledge Dissemination. In this ambitious effort, the combination of disciplines and domains is a critical element of the Hub's design. For this reason, the selection of stakeholders follows the Quadruple Helix principle, drawing people from academia, industry, government and civil sectors.

The function and reach of the Hub are established at several levels. The core of the Hub is composed of the Greek and the Romanian components, with multiple champions from the Widening European countries, altogether supported by the insights, experience and

advancements of partners in the UK and the Netherlands. Through this approach, a scheme of cultivation and diffusion of knowledge is established, fostering wide and fast impact on various pressing water challenges (pollution, scarcity, resilience to climate change etc.)

Indicatively to the actions of the WaterWise Hub, some key components are the development of an R&I Strategy, an Action and Investment Plan, a FundFinder Portal, a Strategic-Tactical & an Operational Decision-making toolbox, Tech Challenges, an Innovation Accelerator, a Technology Transfer Office and more. Putting emphasis on the dissemination of knowledge, specialized components are to be developed here as well, such as Life-Long Learning Training Courses, a Knowledge Hub, a Matchmaking Platform and a Mentoring Scheme.

Overall, the WaterWise Hub is designed to be a self-sustainable and ever-growing establishment that through its multi-faceted action will keep propelling water management towards the Circular Economy.

1.3 Data Management in the Project Lifecycle

The sheer diversity, scale and complexity of activities of the WaterWise Hub are meant to create proportionately diverse, sizable and complex types of data. Therefore, effective data management is necessary for the success of the project. The datasets that will be created require a structured approach in order to ensure their quality, security and accessibility. The DMP is created to standardize this effort, and it will be updated through the implementation of this project in order to ensure full coverage over any new datasets that come up.

At the initial stage of the project some fundamental principles will be established regarding expected types of data, collection methods, storage protocols and metadata specifications. This will set the ground for aligning with the FAIR principles, required by the EU, right from the outset. Building on these foundations, in the progression of the project, an ongoing assessment will follow to ensure that all data is documented according to the pre-defined conventions, are securely stored and readily accessible to the permitted stakeholders. It is during this phase that new types of data may come up and therefore the DMP will be updated accordingly.

Upon completion of the project, all datasets, findings and outputs will go through a final review before being made fully and openly accessible to the public (for the applicable datasets) through trusted repositories such as Zenodo / OpenAIRE. For datasets that are characterized as confidential, the corresponding policies will be applied (e.g. GDPR, EU Open Science etc.) and they will be archived accordingly.

This structured approach ensures a fluent and complete capture of all data, creating harmonized and standardized datasets, accessible and easy-to-navigate to any interested party. In this manner, the long-term impact of the project is maximized and the WaterWise Hub will continue its operation on a robust information infrastructure.

The DMP is intended to be a dynamic document and will be updated as needed throughout the project.

1.4 Target Audience

Given the importance of its core subject (water) and its diverse outreach, the data of the WaterWise Hub has a wide spectrum of target audience, both internal and external to the project. Therefore, the DMP is designed to serve the full variety of interested parties, ensuring that Open Science principles are effectively implemented. The core identified categories of such parties are:

- **Project Partners:** Being the party that will mostly create the datasets, the DMP has a direct importance here in order to provide the relevant guidelines for the creation of these datasets according to the specifications. Also, it the navigation and utilization of the datasets is of critical importance for the execution of the required tasks and the successful completion of the project.
- **Granting Authority:** In accordance with the Grant Agreement, the Consortium is required to follow all agreed terms regarding data handling. The DMP provides the Granting Authority with a detailed blueprint that assures how these terms will be fulfilled.
- **Researchers and Academia:** The DMP sets clear specifications and guidelines on the types, structure, format, metadata etc. of the project's data. Therefore, to the non-participants of the project it can serve as a manual of how the created datasets can be accessed and navigated.
- **Policymakers and regulatory bodies**: Evidence-based promotion of new policies is an important aspiration of the WaterWise Hub. In this effort, convenient and effective access on the side of policymakers and official regulatory bodies allows for a strong corroboration of the policy changes proposed from the Hub's findings.
- **Industry and Innovation Stakeholders:** innovation and business generation are another goal of Hub, both within and outside it. By providing easy access to the project's findings towards innovators and entrepreneurs, they are allowed in refined market insights that can help them identify opportunities and invest in them.
- **General Public / Civil Society:** Raising awareness (and thus public support) on the importance of water and the existing challenges is part of the propelling force for the transition to the Circular Economy. Through access to the datasets, and the facts that emerge from them, the public can be more effectively sensitized and mobilized on the matter, encouraging also participation in the decision-making process.

Providing easy and convenient access to the project's datasets to this diverse set of stakeholders can benefit both them and the project. Given the diversity of their backgrounds, skills and

intentions, the DMP proves to be an essential document that will balance accessibility, GDPR, ethical considerations and IP constraints.

2. Data Summary

Given the multidisciplinary nature of the project and the diversity of activities performed, a large volume and variety of data types is expected to be collected, generated and re-used for multiple purposes and from multiple stakeholders, both internal and external to the project. Such data can be included in the official deliverables of the project but may also play an intermediary or administrative role and therefore their contribution may not be "directly" visible on the project's outcomes.

2.1 Project data and purpose of use

2.1.1 Data to be collected

Individual pieces of data will be needed to be collected as to provide a starting ground for many of the project's activities. In brief description, their nature and purpose are listed in the Table 1 below:

The project will generate various types of data such as:

- **Survey data** from citizen engagement activities
- **Modelling data** (e.g. from hydrological/hydraulic and behavioural analyses)
- **Geospatial data** (e.g. GIS layers, remote sensing and earth observation data)
- **Documentation and metadata** (codebooks, protocols)

Table 1: Data to be collected

Data to be Collected	Purpose (Task – Deliverable)
Stakeholder engagement and mapping data	Identifies key stakeholders, fosters engagement, and informs policy recommendations. (T1.3, T2.1)
Policy and regulatory framework data	Supports legal compliance, regulatory alignment, and policy development. (T2.2, D2.1, D3.1)
Market data on investment trends and funding opportunities	Provides funding insights for research, innovation, and startup support. (T3.1, T3.2, D3.1, D3.2-3-4)
Pilot project implementation and performance/KPI data	Tracks pilot project progress (Open Calls program), assesses sustainability, and validates innovation models. (T4.3, T4.4, D4.4, D4.5)
Meteorological, hydrological, environmental and water governance data	Analyzes water resource management, climate resilience, and policy impacts. (T4.1, D4.1, D2.1, D4.1, D4.2)
Market data on technology development and innovation tracking data	Monitors technological challenges, startup acceleration, and market adoption. (T5.1, T5.2, D5.1, D5.2)
Training modules content and participant feedback	Evaluates skills development, training effectiveness, and knowledge transfer. (T6.2, D6.2)
Mentorship program data and engagement statistics	Assesses mentoring success, mentee growth, and impact of knowledge-sharing. (T7.1, T7.2, D7.1, D7.2)
Dissemination and communication outreach data	Tracks outreach performance, knowledge dissemination,

Data to be Collected	Purpose (Task – Deliverable)
	and stakeholder engagement. (T8.1, D8.1, D8.3)
Website and social media engagement metrics	Measures digital reach, content performance, and engagement trends. (T8.2, D8.2)
Workshops, conferences, and networking event data	Records event participation, assesses networking impact, and captures collaboration insights. (T8.3, D8.4)
Market trends, business development, and commercialization data	Supports business case evaluation, market entry strategy, and post-project sustainability. (T8.5, D8.5)

2.1.2 Data to be generated

Through the activities of the project, new data will be generated as results or intermediary steps. Indicatively, some are the following in Table 2 below:

Table 2: Data to be generated

Data to be Generated	Purpose (Task – Deliverable)
Stakeholder engagement	Measures stakeholder influence, policy alignment, and impact assessments. (T1.3, T2.1)
Policy suggestions / white papers / reports	Promotion of policy change (D2.1)
Investment analysis and funding strategy reports	Evaluates investment trends, guides financial planning, and supports research funding strategies. (T3.1, T3.2, D3.1, D3.2-3-4)
Pilot project performance assessments	Validates pilot project outcomes, assesses sustainability, and benchmarks innovation success. (T4.3, D4.4)
Meteorological, hydrological and environmental modeling outputs	Supports climate resilience modeling, water governance policy formulation, and environmental assessments. (T4.1, T4.2, D2.1, D4.1, D4.2)
Technology challenge case studies and innovation reports	Analyzes technology innovation, assesses startup acceleration, and evaluates market readiness. (T5.1, T5.2, D5.1, D5.2)
Training impact assessments and upskilling evaluations	Evaluates training program effectiveness, tracks knowledge transfer, and measures learning outcomes. (T6.2, T6.3, D6.2)
Mentorship program evaluations and mentee progression records	Tracks mentorship effectiveness, documents mentee progress, and assesses impact on professional development. (T7.1, T7.2, D7.1, D7.2)
Communication and dissemination effectiveness reports	Measures dissemination impact, outreach success, and public engagement in research activities. (T8.1, D8.1, D8.3-4)
Website traffic and social media engagement reports	Assesses digital engagement, monitors web traffic trends, and evaluates online content performance. (T8.2, D8.2)
Event participation summaries and networking impact assessments	Captures event participation, networking effectiveness, and collaboration results. (T8.3, D8.3-4)
Business model feasibility studies and commercialization strategies	Supports business case development, market entry strategy, and long-term project sustainability. (T8.5, D8.5)

2.1.3 Data to be re-used

Building extensively on past findings and keeping in compliance with any contextual requirements (e.g. laws and regulations), existing datasets will be utilised as well in the project, as indicated in Table 3 below:

Table 3: Data to be re-used

Data to be Re-used	Purpose (Task – Deliverable)
Stakeholder engagement case studies and past collaboration records	Improves stakeholder engagement strategies and strengthens policy impact assessments. (T1.3, T2.1, D2.1)
Existing policy frameworks, legislative documents, and regulatory compliance reports	Guides policy mapping, regulatory alignment, and compliance with environmental laws. (T2.2, D2.1, D3.1)
Historical investment trends and financial benchmarking data	Provides insights into research investment trends, funding mechanisms, and economic forecasting. (T3.1, T3.2, D3.2-3-4)
Water resource management models and sustainability frameworks	Supports climate resilience modeling, sustainable water governance, and hydrological research. (T4.1, T4.2, D2.1, D4.1, D4.2)
Technology challenge methodologies and startup acceleration best practices	Benchmarks technology acceleration processes, startup incubation models, and innovation funding strategies. (T5.1, T5.2, D5.1, D5.2)
Educational and training best practices from previous EU projects	Enhance training design, course content, and knowledge-sharing effectiveness. (T6.2, D6.2)
Mentorship program structures and knowledge transfer strategies	Improves mentorship program effectiveness, participant engagement, and career progression strategies. (T7.1, T7.2, D7.1, D7.2)
Communication and dissemination methodologies from past projects	Ensures efficient dissemination of project findings, outreach efforts, and public engagement. (T8.1, D8.1, D8.3-4)
Social media engagement analytics and digital outreach benchmarks	Optimizes digital engagement strategies, content visibility, and audience reach. (T8.2, D8.2)
Networking best practices and stakeholder engagement strategies	Strengthens stakeholder networking, collaborative research, and project dissemination impact. (T8.3, D8.3-4)
Business sustainability reports and commercialization strategies	Provides validated business models, market entry strategies, and commercialization approaches. (T8.5, D8.5)

2.2 Data Types and Formats

The various data will be obtained, maintained and processed in the corresponding format, according to the type of the data. For most of the conventional information representations (text, spreadsheets, charts, presentations etc.) the Office Open XML standard will be followed, for the purpose of maximizing accessibility. In detail, the expected types and format are presented below in Table 4:

Table 4 Data types and formats

Type of Data	Supporting format
Text Documents (Reports, Policies, Agreements)	.pdf, .docx, .txt, .odt
Spreadsheets (Financial Data, Performance Tracking)	.xlsx, .csv, .ods
Presentations (Training Materials, Awareness Campaigns)	.pptx, .odp, .pdf
Audio Recordings (Interviews, Stakeholder Consultations)	.mp3, .wav, .ogg
Video Files (Workshops, Webinars, Dissemination Content)	.mp4, .avi, .mov, .mkv
Geospatial Data Files (Hydrological and Environmental Models)	.shp, .geojson, .kml, .nc (netCDF), .tiff
Structured Datasets (Metadata, Research Data, Open Science Repositories)	.json, .xml, .csv, .rdf
Legal and Compliance Documents (IPR, Licensing Agreements, Regulatory Data)	.pdf, .docx, .txt
Website Analytics and Social Media Engagement Data	.csv, .json, .xml, .xlsx
Event and Participation Logs (Networking, Conferences, Stakeholder Engagement)	.csv, .xlsx, .pdf, .json

2.3 Data Origin and Expected Size

Various sources will be utilized for obtaining the necessary data, raw pieces or established datasets.

- **Project documents/records:** A source regarding the organization and monitoring of the project's progress comes from the official documentation of the project, e.g. the Grant Agreement.
- **Regulatory – Policy frameworks:** Official standards, whether mandatory or best-practices will be utilized in multiple tasks of the project, whether that is to conduct the project, propose new policies, guide third parties on their operations, provide upskilling to professionals etc.
- **Scientific research and Open Data repositories:** A variety of scientific data will be necessary to develop outputs of the project (e.g. policy suggestions, strategic decisions, decision-making models, curriculums etc.). The same data will also be required to justify and explain the results of the project.
- **Financial and Investment Data Sources:** As a large part of the project has to do with business creation, the identification and organization of funding sources is an important section of data collection. Investment databases, funding programs, economic trend reports etc. will be utilized for this purpose.
- **Innovation and Technology Development Records:** fostering innovation is another core aspect of the project, thus monitoring and mapping this landscape will be important. Tech trends, IP records, latest developments and more will be of such value.

- **Stakeholder Engagement:** the stakeholder network of WaterWise Hub is going to be its greatest source of general knowledge and expert insights. Through direct inquiries or group work (e.g. Thinktank) invaluable pieces of information will be obtained for various parts of the project.
- **Social media and Public Events:** With a strong focus on communication, dissemination and extroversion in general, the interaction of the project with the public will be a significant source of data (through web/app analytics, event participation reports, networking activities etc.)

Most data expected to be collected, generated and utilized for the needs of the project are going to be of a text/numerical nature. Therefore, the total estimated size is to be at the GB scale. Given that a variety of third parties will be engaging in the project activities and that more opportunities for impact (e.g. areas of policy suggestion) might be identified along the project implementation, this number is naturally expected to vary significantly.

2.4 Data Utility

WaterWise Hub's work on data collection and generation will create a repository of refined knowledge valuable to a multitude of stakeholders and groups outside of the project. The first such group is policy makers and regulators. Indicatively, government agencies, EU bodies and local water authorities. Here the data are valuable in the sense that it can assist the process of updating water management regulations, circular economy practices and environmental sustainability policy overall. The refined and consolidated data can help such stakeholders to make highly-informed decisions.

Another group is academic and research institutions (universities, research centers, environmental science departments etc.). As a pool of refined and pre-processed data, it holds potent material for scientific studies, comparative analyses and innovation in the field of water sustainability. Some of the most pertinent activities are development of new models and scientific publications.

In the broader sense of creation and dissemination of knowledge, the data can be useful to educational and training institutions (a field which the project also emphasizes on). The project's collection of consolidated knowledge in the water sector and the development of new can be a foundation in the development of curricula, training material and lifelong learning initiatives for vocational schools and professional training centers.

Significant benefit is also expected for the industry and business sector. The data, highlighting the challenges and opportunities of the water sector, can inform companies such as water technology, sustainable innovation and utility services to develop new products and business strategies. The highly robust data-driven insights generated can facilitate the innovation process, allow for optimized resource use and help increase competitiveness overall.

Narrowing further to innovation, the data hold significant value for startups and entrepreneurs (proof, that a significant part of the project is dedicated on this). In a similar manner that it can assist businesses in general, here the data can assist in developing products, business models, market-entry strategies and altogether the commercialization of novel solutions.

Coming to societal organisations, such as NGOs, the data can be valuable for supporting fact-based motivation towards their target groups and fields of action, e.g. raising public awareness and advocating for water sustainability and environmental protection. In practical terms, the data can be used to build campaigns, lobby for policy change and provide workshops and education to the public on sustainable water management.

Lastly, the data are also open to simple citizens and communities. While not designed for utilization from the "layman", a common person with interest in the matter of water sustainability (and thus some familiarization with the basic terminology and notions) can become much better informed, apply effective practices, motivate other people in their community and take accurately focused initiatives.

3. FAIR Data

WaterWise Hub is committed to following the FAIR principles of data handling, in order to maximize its impact beyond the project's context as well. Here are the relevant details on how this will be achieved

3.1 Making Data Findable

All datasets created within the project will be assigned with Persistent Identifiers (PIDs) to ensure their long-term traceability and accessibility, for example, Digital Object Identifiers (DOIs). All datasets will be stored in well-established and trusted repositories such as Zenodo, where the PIDs will be assigned and will be compliant with the EOSC standards.

Additionally, the project has already established both OpenAire profile and Zenodo Community to ensure that all produced data will be easily discoverable and accessible. This approach aligns with the project's commitment to supporting the FAIR data principles, enhancing the visibility of research outputs.

As planned, large part of the datasets will be stored in the WaterWise Knowledge HUB, built around the Water Europe Marketplace (WE Marketplace). Data within the WaterWise Knowledge HUB are made findable through multiple means: a) full-text search capability, b) the application of appropriate filters according to data categories such as technologies, projects, products, case studies, events, organizations, and stakeholder profiles and c) georeferenced case studies accessible via an interactive map. Each item within the platform has its own unique URL to ensure ease of reference and access. The use of established data catalogs such as WE Marketplace allows systematic cataloguing and registration of datasets, significantly enhancing data discoverability and interoperability across related platforms and communities.

To enhance their discoverability and reachability and depending on the data category, rich metadata will accompany every dataset, providing comprehensive and descriptive information about the contained data. Any standardized metadata frameworks that are applicable to the project's datasets will be followed, as to facilitate further this goal. Such are Dublin Core, DataCite, INSPIRE, and ISO 19115 (for environmental and hydrological data) as well as the OpenAIRE Guidelines for Literature Repository Managers - Data Archives for (for research outputs and publications).

For data lacking existing standards, custom metadata profiles will be developed using best practices recommended by Horizon Europe and international data management protocols. This approach guarantees sufficient descriptive and structural information for effective interpretation and reuse.

To further facilitate the findability of data, optimal search keywords will be identified and included in the metadata. Drawing from relevant key thematic areas of the project (water management, circular economy, sustainability etc.). The inclusion of such standardized,

indicative and popular keywords will allow for easier and faster discovery of the data through general or specialized search engines.

Lastly, in this effort, the metadata will be structured accordingly, to support automated harvesting and indexing by platforms such as data aggregators and metadata repositories. The records will be designed to be machine-readable in accordance with protocols such as OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting). This way, the WaterWise Hub data will be easily included in large-scale databases like OpenAIRE, EOSC etc.

3.2 Making Data Accessible

As mentioned previously, for the storage of data, a reliable and trusted open-access repository such as Zenodo will be utilized. This repository was deemed the optimal choice based on its reliability, long-term preservation capabilities, and compliance with FAIR principles, also allowing for the assignment of persistent identifiers (e.g. DOIs) to the datasets. Any necessary customizations will be made to ensure that the data will be uploaded, curated and preserved according to the relevant standards thus facilitating their discoverability and accessibility.

Committed to Open Science, the WaterWise Hub will keep its data "as open as possible, as closed as necessary". Therefore, while following its mandate for transparency and reusability, certain datasets will be excluded. These include:

- **Personal/Sensitive Data & Privacy**: If data violates privacy or data protection (e.g. GDPR), they will not be openly shared and will be kept confidential or anonymized as needed to comply with legal and ethical standards.
- **Intellectual Property & Commercial Exploitation**: If immediate open access undermines IPR or planned commercialization, data will be temporarily or permanently closed. As Horizon Europe explicitly mentions commercial exploitation concerns are a valid exception, e.g. data that are commercially valuable may be kept closed if openness would jeopardize a patent or reveal trade secrets. In some cases, WaterWise HUB may impose an embargo period on the data, releasing it only after a patent is filed or a publication is out.
- **Confidentiality Agreements & Partner Interests**: If a project partner has legitimate interests or confidentiality obligations (e.g. third-party agreements, national security, or other "protected legitimate interests"), data will be restricted.

The datasets designated for open-access will be made accessible through free and standardized protocols such as HTTP and OAI-PMH. Our repository in Zenodo will provide file downloads over HTTP and expose metadata via OAI-PMH for harvesting. For the restricted-access datasets, controlled authorization mechanisms will be utilized.

For the case of the metadata, all of them will be openly available and licensed under a public domain dedication, CC0 (in compliance with Horizon Europe's Open Data mandate). The contained information will guide users to access the corresponding data, with additional contact details and further security protocol instructions for the restricted datasets. Even in the case of

data deleted, removed, made restricted or expired, the metadata will be maintained as to allow for fulfilling the needs of findability and citation records.

Other important information contained in the metadata, regarding accessibility, will include references to specific software tools that are necessary to access and operate the data (name, version, configuration details etc.). The use of open-source software will be prioritized and promoted accordingly in the metadata, as to remove any possible obstacles on the accessibility and reproducibility of the results for the user.

According to the Horizon Europe guidelines, the project is required to keep data available after its completion for a minimum of 5 years and possibly much longer. In fact, Zenodo's policy notes that items are retained for the lifetime of the repository (the foreseeable future of CERN, with an experimental program at least ~20 years) and that in case of closure, content will be migrated elsewhere. Any data made inaccessible due to obsoletion, or lack of maintenance will be designated as such, while their metadata will remain available for the purposes of traceability and citation capability, as explained previously. For the datasets that require specific software to be operated, the relevant references, documentation and open-source code will be preserved and updated in the metadata, where applicable.

3.3 Making Data Interoperable

In order to facilitate the integration and harmonization of data, the project will adhere to community-endorsed interoperability best practices that follow broadly accepted standards and methodologies. Both the data's and the metadata's structure will be based on common vocabularies and formats that are widely recognized and accepted by the scientific, technical and business communities.

The main metadata standard will be that of DataCite, since it is endorsed by European Open Science Cloud (EOSC) and OpenAIRE. In specific for environmental and hydrological data the ISO 19115 (Geographic Information - Metadata), Dublin Core (general metadata descriptions) and INSPIRE (spatial datasets) will be utilized. Implementing these will allow for the optimal cross-dataset compatibility, through supporting machine-readability and automated processing.

Coming to the formats of the datasets themselves, the most common structured, non-proprietary, and widely accepted ones will be chosen, to the maximum extend possible. These include CSV, JSON, XML, NetCDF and HDF5 since these are the universally recognized as the most interoperable, long-term accessible and easy-to-integrate in data processing pipelines. For any data with geospatial components, formats of GeoTIFF and SHP will be preferred, compatible with geographic information systems (GIS). Lastly, for any text-predominant data the widespread formats of TXT, PDF, and DOCX will be utilized.

A significant part of the project's data will be data reused from other sources, such as case study data and applied technologies. On this matter, qualified references to these external datasets will be incorporated into the metadata records to create the required contextual connections. Such sources can be datasets from past projects, relevant EU repositories and public databases. Through this inter-linkage not only data usability is enhanced but also the creation of integrated datasets is cultivated, allowing in the future for broader analytical opportunities.

Considering here as well the mandate for long-term usability of the data, all the datasets will be accompanied by extensive metadata that include essential information such as contextual information, data lineage, and versioning while metadata schemas and data dictionaries will be maintained alongside the datasets in order to preserve their interpretability.

3.4 Increasing Data Reuse

In alignment with Horizon Europe's Open Science principles, the project is fully committed to promoting reusability, transparency, and accessibility through the adoption of open-source software and the sharing of all relevant research outputs beyond traditional publications and datasets. Recognizing that software, algorithms, and workflows are critical components of reproducible science, the project will make available, wherever possible, the tools and instruments developed during its implementation.

Metadata will accompany each dataset, clearly documenting the software required for reuse or validation. This includes precise information on the software name, version, configuration settings, dependencies, and whether the tool relies on commercial or proprietary platforms. Such transparency ensures that third parties can easily access, understand, and reuse the data, even years after the project has concluded.

This practice is not only compliant with the Horizon Europe Model Grant Agreement—which mandates the inclusion of all tools and instruments needed for data reuse—but also reflects best practices as described in the Annotated Grant Agreement. By reducing reliance on proprietary solutions and prioritizing open-source alternatives, the project eliminates potential access barriers and enhances reproducibility.

This commitment will be regularly reviewed to ensure that all outputs meet the highest standards of open science. The project's proactive approach strengthens the long-term impact and sustainability of the research, facilitating broader adoption and fostering innovation within the scientific community.

In accordance with Horizon Europe's open data policy, the data produced and gathered through the WaterWise Hub project will be placed in the public domain for free and wide-reaching reuse. If no other restrictions apply, all data sets will be made available under the Creative Commons Attribution (CC-BY) license which permits redistribution and adaptation of the data as long as

users give proper credit to the data creators. This license allows the data creators to be acknowledged while facilitating wider dissemination and use of the data.

3.5 Other Research Outputs

For the needs and purposes of the project, non-data research outputs created or reused are expected as well. These outputs could be digital (e.g., software tools, models, workflows, simulation algorithms, analytical protocols) or physical (e.g., materials, samples, prototypes). These outputs too shall be governed by the FAIR principles (Findable, Accessible, Interoperable, Reusable) as much as possible, ensuring that they remain useful once the project is over.

Regarding digital research products, these will be necessary for simulating hydrologic conditions, water quality observation, and environmental effects analysis. All the generated digital tools and models will be deposited via trusted open-access repositories (e.g., Zenodo or OpenAIRE) to ensure their long-term availability and traceability.

For interoperability and reuse assurance, the software and models will be accompanied by extensive documentation, including:

- User Manuals and Guides: To enable effective use and tool adoption.
- Source Code and Versioning: Located on sites such as GitHub or GitLab under open-source licenses (e.g., GPL or MIT).
- Metadata and Provenance Information: Including descriptions of how the tools were created, how they are maintained, and how they should be utilized.

Where practicable, programs and digital tools will be constructed using community-vetted coding practices and employ standard data formats (e.g., JSON, XML) to enable greater compatibility with other systems.

4. Allocation of Resources

4.1 Costs and Funding

Given that many data collected or generated will be text-based, no specialized forms repositories/platforms/services are expected to be needed, and thus no significant costs that will require a specific part of the project's budget for this purpose.

The IT provisions of the Project Coordinator (NTUA) can accommodate entirely the needs for initial storage of data, both on local servers as well through its official information sharing provider, Microsoft Teams. As the project progresses, any necessary assessments for further utilization of software will be performed and incorporated in the budget as described by the restrictions and provisions of the Grant Agreement regarding direct and indirect expenses.

Mature datasets, accompanied by the metadata described previously, will be transferred to trusted platforms such as Zenodo and OpenAIRE, which are free of charge. Specially designed according to the EU's principles of Open Science, these repositories fulfill all necessary requirements regarding accessibility and security and therefore facilitate the effort of minimizing costs in pursuit of making data FAIR.

A special notion must be made on all kinds of sensitive data, mostly the ones that relate to the generation of intellectual property. Since IPRs are expected to be a significant part of one of the WaterWise Hub's activity sectors, special care might be required here for the storage of this data. In case a service/platform/medium is identified as necessary, then the corresponding adjustments in budget allocation will be made.

4.2 Responsibilities

The responsibilities of proper data handling, according to the DMP, follow uniformly the entire chain of data flow. Task assignees are to clearly mark any collected or generated data, considering the type of data set it belongs to. The **Work Package Leaders** are the ones responsible for safeguarding appropriate data collection, processing and documentation, providing guidance to the relevant stakeholders. They are to ensure that any data generated and collected follows the agreed standards and protocols outlined in the DMP (e.g. metadata creation and quality assurance).

Once the datasets are past their initial curation, consolidated in form and format, they are transferred to the storage space of the coordinator where the **Data Manager** and the corresponding team will verify and coordinate any remaining edits/additions for the full compliance of the data with the DMP. Also, the level of restrictions will be considered based on the nature of the data, after consultation with the legal and ethical advisors of the project.

As with Risk Management and Quality Assurance, it is the responsibility of all partners to facilitate to the maximum possible extent at any point the accurate data reporting and timely

submission of outputs. Any issues observed must be escalated and reported to the responsible individuals / bodies of the Consortium.

4.3 Long-term Preservation

The aim duration for the preservation of data is at least 10 years after project completion. Therefore, the most suitable platforms for hosting the finalized versions of datasets are Zenodo, OpenAIRE, and the European Open Science Cloud (EOSC), which guarantee long-term preservation and accessibility. To ensure long-term findability a Persistent Identifier (PID, e.g. a Digital Object Identifier (DOI), will be assigned. Metadata will be maintained alongside the data to enable accurate interpretation and reuse, following DataCite standards.

In parallel with Risk Management, backups and redundancy measures will be implemented additionally for critical datasets, as to mitigate risks of data loss. The information service provider of the Coordinator, Microsoft Teams, is a provenly reliable service provider and therefore this minimizes any concerns about data availability for the initial phase of data storage.

5. Data Security

5.1 Storage and Backup

The project will implement protocols for secure data storage and regular backups to safeguard against loss or unauthorized access. All primary research datasets will be stored on trusted platforms with robust security measures.

Final or public datasets will be deposited WaterWise Hub Zenodo Community, that provides long-term preservation and assigns persistent identifiers (DOIs) for findability. Additionally, the WaterWise Hub's knowledge platform is built upon the Water Europe Marketplace, which will serve as a central repository for storing data in the following data categories: technologies of the circular economy, products serving the circular economy, case studies and best practices and stakeholder profiles. By using OpenAIRE, Zenodo and the Water Europe Marketplace, the project will ensure redundancy: datasets will exist in multiple secure locations, reducing the risk of a single point of failure and providing reliable backup copies.

All storage solutions employed will be GDPR-compliant and will have strong built-in security. Data stored on cloud services during the project (such as the Coordinator's Microsoft Teams/SharePoint environment for working datasets) will be protected by enterprise-grade access controls and encryption.

Regular backup procedures will be put in place: critical datasets will be automatically backed up on a defined schedule (e.g., daily incremental backups with periodic full backups) to secure off-site storage. These backup files will be encrypted, ensuring that even if backup media are compromised, the data remain unintelligible to unauthorized parties. Backup integrity will be periodically tested by performing test restorations, and logs of backup operations will be monitored for any failures or irregularities. All project partners will be instructed in the approved storage and backup procedures, and a designated Data Manager will oversee compliance with these protocols. Overall, these measures will ensure that project data is securely stored with adequate redundancy, in full alignment with data protection obligations (including GDPR requirements for confidentiality and integrity).

5.2 Access Control

Access to datasets will be strictly controlled through strong authentication and role-based access control (RBAC) mechanisms, particularly within the Water Europe Marketplace platform hosting the WaterWise Hub data. All users will be required to authenticate with secure credentials before gaining access to non-public data or sensitive functions. Data transmission channels will be encrypted using SSL/TLS protocols (for example, the Marketplace and Zenodo are accessed via HTTPS), preventing eavesdropping or man-in-the-middle attacks during transfers. Internally, any exchange of sensitive data between partners (e.g. via SFTP or the project's collaboration

portal) will also use encrypted connections. In addition, data at rest on servers (for both the Marketplace platform and any internal databases) will be encrypted (following standards such as AES-256) to protect confidentiality in the event of a breach. These encryption measures ensure that whether data is in transit or stored, it remains protected from unauthorized access. Server access will be restricted exclusively through a Virtual Private Network (VPN), ensuring a secure and encrypted connection for all interactions with server resources, thereby significantly reducing the risk of unauthorized access.

Within the Marketplace, a robust RBAC scheme will define what each user can see or do, enforcing the principle of least privilege. The key roles include:

- **Unregistered Users**: Users who have not logged in (the general public). They can only browse and view information designated as publicly available. Unregistered users have read-only access to open data (e.g. public summaries of case studies or technologies), but cannot download sensitive files or view any non-public details.
- **Registered Users**: Users with an approved account on the platform. In addition to all public information, registered users can access additional resources such as detailed dataset metadata or certain data downloads that the project shares with the broader community. They may also interact with platform features (for example, saving items to a personal workspace or contacting data contributors) within permitted bounds. However, they cannot modify any project datasets and cannot access confidential or restricted data without explicit permission.
- **Event Managers**: A specialized role granted to members who organize or manage events within the WaterWise Hub. Event Managers can create and edit event entries on the Marketplace (e.g. posting event details, agendas, and related materials). They have rights to manage event-related data (such as participant lists or outcomes of events), but their access to other project datasets is limited to what a normal registered user can see, unless they hold another role as well.
- **Case Study Data Managers (CSDMs)**: These are project team members entrusted with managing data for specific case studies. CSDMs can upload, edit, and curate case study datasets and their metadata on the Marketplace. They control access levels for their case study data – for instance, marking certain sensitive case study datasets as private or only accessible to the consortium. CSDMs ensure that case study information is kept up-to-date and that any privacy-sensitive data in their datasets is appropriately anonymized or restricted (in line with GDPR and the project's ethics requirements) before broader sharing.
- **Technology Data Managers (TDMs)**: Similar to CSDMs, TDMs are responsible for datasets related to technologies or products showcased in the Marketplace. They can create and manage entries for technological solutions, upload and update technical datasets or documentation, and maintain the accuracy of these records. TDMs have the authority to set permission levels for the technology-related data, for example, restricting proprietary technical details to project-internal access while making high-level information public.

- **Marketplace Administrator**: The Administrator holds the highest level of privileges on the platform. A Marketplace Administrator can manage user accounts (approve new registrations, assign roles, or revoke access) and oversees all content on the Marketplace. Administrators have full read-write access to all datasets, including those marked as restricted or confidential. They are responsible for ensuring the platform's security configurations are maintained (such as enforcing password policies and access rules) and for approving any special access requests. The Administrator role serves as the gatekeeper for sensitive data: for example, if a registered user requests access to a restricted dataset, the Administrator (often in consultation with the relevant CSDM/TDM) must authorize that access.

This role-based model will ensure controlled authorization: each dataset on the Marketplace can be tagged with an access level tied to one or more of the above roles. In practice, any attempt to access a dataset will trigger a check of the user's role against the dataset's permissions. If the user lacks the required role, access will be denied, and the user may submit a request for authorization. Such access requests will be handled through formal approval procedures — for example, a user might contact the dataset's manager or an Administrator to obtain access, and a record will be kept of any permissions granted. All user login attempts will be logged by the system, creating an audit trail for security monitoring.

Authentication on the platform will use secure methods: user passwords will be stored as salted hashes rather than plain text, and user sessions will be protected via secure cookies and timeouts. The platform will enforce strong password policies (e.g. requiring a mix of characters and periodic password updates), especially for accounts with elevated privileges like CSDMs, TDMs, and Administrators. Only authorized personnel (those with the appropriate roles and permissions) will be able to add or modify data on the Marketplace, which prevents any unauthorized changes to the datasets. These authentication and authorization measures, combined with end-to-end encryption, will protect sensitive and personal data from unauthorized access. In accordance with GDPR, personal data will be accessible only to those who are explicitly permitted and require it for defined purposes.

5.3 Ongoing Monitoring and Security Audits

Data security will not be treated as a one-time setup but rather as an ongoing responsibility throughout the project's lifetime. The WaterWise Hub consortium will perform continuous monitoring of its data management systems and platforms to promptly identify and address potential security issues. This includes actively monitoring user access logs and permission changes on the Water Europe Marketplace and watching for any suspicious activities (such as repeated failed login attempts or unusual download patterns). The Marketplace's administrators have tools to alert them to potential security incidents, and any irregularities will be investigated immediately with corrective action taken as needed.

In addition to day-to-day monitoring, the project will conduct **regular security audits** of its data security measures. At defined intervals (for example, annually, as well as at the project's mid-term and final review), the Data Management team together with the project's IT security advisors or Data Protection Officer will systematically review the effectiveness of all security controls.

6. Ethics

6.1 Ethical and Legal Considerations

Adherence to any ethical and legal limitations is of critical importance for the implementation of the WaterWise Hub. Involving a multitude of stakeholders and third parties (organizations as well as individuals) the data collection and processing may include personal or sensitive data, for which handling is strictly dictated by GDPR and national data protection laws and regulations. Therefore, proper care will be taken for the handling and storing of such data and any data containing personal information will be appropriately anonymized or aggregated before sharing, with access being restricted when necessary to protect individual privacy.

Especially on the matter of Intellectual Property Rights related data, any datasets, models, software, or digital outputs that are collected and generated will be evaluated for their relevance to IP issues, in order to ensure that ownership and usage rights are clearly defined and protected. Overall, the project will follow the principle of "as open as possible, as closed as necessary", ergo using a Creative Commons (CC-BY) licensing for openly shared data, while applying a data-handling policy in alignment with EU IPR regulations for the protected or commercially sensitive.

6.2 Informed Consent

For any data-collection activities from third-party stakeholders and individuals, appropriate informed consent will be obtained, particularly when handling personal or sensitive data. Consent forms will be clearly outlined to include data sharing and long-term storage provisions, defining data collection purposes, data use, storage duration, and potential reuse. Participants will be well-advised regarding their rights to withdraw consent at any time. If data is being collected indirectly or from secondary sources, the project will verify that due consent was granted prior to reuse of data. Evidence of consent processes will be maintained and securely stored as part of the data management plan.

7. Other Issues

7.1 Compliance with Other Data Management Procedures

As mentioned above, all the mandatory regulation and best-practice policies will be implemented for the needs of the project. That includes first and foremost the national laws for data management and privacy protection in the locality of the countries where data will be collected, processed and stored. Further than that, for personal information, the GDPR applies to all the data-processing within the WaterWise Hub's operations, as any other applicable EU frameworks. Finally, the Horizon Europe Open Science guidelines are followed facilitating open access to the project's findings and FAIR data sharing.

WaterWise Hub

Funded by the European Union