



IDEATION

Data Management Plan V1

Deliverable 1.4
WP1

Author: Angeles Tejado
Date: 29 November 2024



Funded by
the European Union



GRANT AGREEMENT NUMBER	101136799		
ACRONYM / FULL TITLE	IDEATION: InlanD watErs in the digitAI TwIn OceaN		
START DATE	01/06/2024	DURATION	24 months
END DATE	31/05/2026		
PROJECT URL	www.ideation-euproject.eu		
DELIVERABLE	D1.4		
WORK PACKAGE	1		
CONTRACTUAL DATE OF DELIVERY	30 November 2024		
ACTUAL DATE OF DELIVERY			
NATURE	DMP	DISSEMINATION LEVEL	PU
LEAD BENEFICIARY	CETAQUA		
RESPONSIBLE AUTHOR	FIWARE		
CONTRIBUTIONS FROM	ALL		
ABSTRACT	The initial version of the Data Management Plan (DMP) for IDEATION follows the H2020 FAIR data management guidelines. It outlines procedures for monitoring generated data, ensuring privacy and confidentiality, and applying ethical standards for data use, storage, and sharing across the project. Technical standards are defined to determine data eligible for the open data initiative.		



Disclaimer

Any dissemination of results reflects only the author's view and the European Commission is not responsible for any use that may be made of the information it contains.

Copyright message

© **IDEATION Consortium, 2024**

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both. Reproduction is authorised provided the source is acknowledged.

TABLE OF CONTENTS

ACRONYMS	5
EXECUTIVE SUMMARY	7
1 Introduction	8
2 Brief description of IDEATION	9
3 Data Summary	11
3.1 Description of the data	12
3.1.1 Data generated and re-used	12
3.1.2 Data collection purposes.....	12
3.1.3 Origin of the data.....	13
3.1.4 Target Groups (TG)	13
3.2 Data generated in IDEATION	14
3.2.1 Stakeholders map	14
3.2.2 OpenKIWAS	14
3.2.3 Reference Architecture.....	19
3.2.4 Publications.....	19
4 FAIR Data	21
4.1 Making Data Findable – Including Provisions for Metadata	21
4.2 Making Openly Data Accessible	25
4.3 Making Data Interoperable	27
4.4 Enabling Data Reuse	28
5 Other Research Outputs	29
6 Allocation of Resources	30
7 Data Security	32
7.1 Data Publishing and Storage	32
7.2 Quality control of the data	32
8 Ethical Aspects	34
9 CONCLUSIONS	35
ANNEX I - Dataset Template	36
ANNEX II - Zenodo Template	37
ANNEX III - Template for General Datasets	39

ACRONYMS

AI	Artificial Intelligence
APIs	Application Programming Interface
CA	Consortium Agreement
CEF	Connecting Europe Facility
CET	Cetaqua
CC0	Creative Commons Zero
CMS	Content management system
DEP	Digital Europe Programme
DM	Data Manager
DMP	Data Management Plan
DoA	Description of Actions
DOI	Digital Object Identifier
DTO	Digital Twin Ocean
DTs	Digital Twins
EC	European Commission
EMBRC	European Marine Biological Resource Centre
EMSO	European Multidisciplinary Seafloor and water column Observatory
ENG	Engineering
ERDF	European Regional Development Fund
FAIR	Findability, Accessibility, Interoperability, and Reuse
GA	Grand Agreement
GDPR	General Data Protection Regulation
IPR	Intellectual Property Rights
MSF	Multi-Stakeholder Forum
NGO	Non-governmental organization



IDEATION

www.ideation-project.eu

OpenKIWAS	Open Knowledge Inventory for Inland Water Systems
ORE	Open Research Europe
SO	Specific Objective
TG	Target Group
WISE	Water Information System of Europe
WoLLs	Water Oriented Living Labs
WP	Work Package
WPL	Work Package Leader



EXECUTIVE SUMMARY

This document introduces the first version of the IDEATION project Data Management Plan (DMP) at M6, which describes IDEATION structured approach to open, standardised, and responsible data management that supports a collaborative data ecosystem. IDEATION DMP is designed to ensure accessibility, interoperability, and reusability of the data collected, generated and/or used during the project's lifecycle.

The DMP puts strong emphasis on Open Science practices: to guarantee online, open access to scientific information and publications, and FAIR principles: to make IDEATION data findable, accessible, interoperable, and reusable for future users such as researchers, policymakers, and other stakeholders.

To support interoperability, IDEATION follows community standards for seamless data and system integration that will ensure easy use and application in different contexts beyond the project. Rigorous planning for metadata standards, data curation and long-term preservation underscores IDEATION's commitment to high-quality data management, supporting future data integrity and accessibility within the inland water ecosystem.

The IDEATION DMP is a living document. Relevant updates will occur after attaining milestones related to, for instance, the generation of new data or to reflect changes related to the original planning, changes in data/output access provisions or curation policies, changes in consortium practices or changes in consortium composition. Updated regularly at mid-term (M15) and upon project completion (M23), the DMP will serve as a dynamic document, evolving to reflect changes in data practices and project requirements.

1 Introduction

The purpose of deliverable *D1.4 Data Management Plan V1* is to outline the monitoring processes for data privacy, security and confidentiality throughout the project lifecycle, ensuring adherence to ethical standards in data use, storage, and sharing. The document corresponds to the initial version of the DMP to be delivered at M6.

Following FAIR and Open Science¹ principles, this DMP includes information on the following: data description, data generated in IDEATION, FAIR data, other research outputs, allocation resources, data security, and ethical aspects.

The first part of the document (Chapter 2, 3, and 4) provides an overview of the IDEATION project mission and data management strategies, explaining how the project addresses data sharing, open research and FAIR data principles. Initial types and formats of data that will be collected, generated and/or used during the project's lifecycle are also summarised.

The second part of the document (Chapter 4, 5, 6, and 7) focuses on the practical aspects of data management, covering data accessibility, repository selection and security measures. It describes the documentation needed for data validation and re-use, licensing for public domain availability and long-term usability by third parties. Additionally, it addresses resource allocation, roles within the consortium and compliance with ethical standards and data security protocols.

¹ https://rea.ec.europa.eu/open-science_en

https://research-and-innovation.ec.europa.eu/strategy/strategy-2020-2024/our-digital-future/open-science_en



2 Brief description of IDEATION

In order to make this deliverable self-consistent, this section briefly recalls the aspects of the project related to the DMP.

IDEATION project follows a two phases implementation:

- Phase I (M1 to M12), aimed at collecting all the necessary information for preparing the roadmap and the set of recommendations and the design of the architecture
- Phase II (M13 to M24), aimed at refining results from Phase I, designing the final reference architecture, and defining the IDEATION roadmap

This first version of the Data Management Plan (delivered in M6) describes data management practices for Phase I, including data collection, storage, sharing, and security protocols relevant to early-stage activities, such as stakeholder engagement, the creation of the Open Knowledge Inventory for Inland Water Systems (OpenKIWAS), and the collection of initial data models and interoperability guidelines for the final design of the reference architecture. Specifically:

- Collection of **requirements and needs** from the stakeholders that will be matched to the existing solutions and technologies selected and stored for the OpenKWAS (T5.1, starting at M9). During this activity, features and functionality of the solutions will be carefully compared to the requirements. Availability, accessibility (open versus closed systems), FAIRness, cost, ease of use, and support options of each solution will be considered to create a matching map. It will allow visualising the relationship between the requirements and the solutions and will be used in following activities that involve the stakeholders. Requirements will be reported in *D2.3: MSF results and functional and non-functional requirements (KER1) V1* at M9.
- **Data governance** considerations are and will be embedded into each step of the OpenKIWAS development that involves sourcing and categorising data, conducting gap analyses, prioritising policies, and establishing the IDEATION schema/ontology to align with FAIR principles and ensure structured data organisation, interoperability, and long-term reusability. The IDEATION schema/ontology will serve as a specialised framework for the categorization and organisation of important data extracted through the analysis of policies and standards related to inland water. The ontology will encompass the various aspects of inland water, facilitate the prioritisation of policies based on their importance and scope and link data received from different sources for storage, analysis and comparison.
- **Online accessibility and the cataloguing of resources** for wider dissemination and reuse, will be facilitated by an Open Source Content Management System (CMS) customizations to manage and share OpenKIWAS data, which will be then made available online through the IDEATION website. The OpenKIWAS will be made available in its first version

at M9 in the deliverable *D3.1: OpenKIWAS (KER2) V1*. The final version will be ready at M21 in the deliverable *D3.2: OpenKIWAS (KER2) V2*.

- The **reference architecture design** will incorporate this structured data, making it scalable, interoperable, and ready for Phase II advancements. Guidelines on data structures, APIs, standards, and formats will be developed to support data exchange and integration across IDEATION's ecosystem, ensuring the architecture aligns with FAIR principles and the broader landscape of interoperability standards (e.g., WaterML², FIWARE³, IDSA⁴, GAIA-X⁵). The reference architecture will be available in its first version at M12 (deliverable *D4.1: Reference architecture and Interoperability Guidelines (KER3) V1*) and the final version will be ready at M21 and reported in the deliverable *D4.2: Reference architecture and Interoperability Guidelines (KER3) V2*.

Future iterations of the DMP will expand to address the needs and outputs of Phase II, incorporating insights from the project's initial findings and refining practices to support the IDEATION ecosystem as it advances. Accordingly, data from the list of recommendations and the IDEATION roadmap (activities that will start a M21 in WP5) will be also included in the DMP.

² <https://www.ogc.org/standard/waterml/>

³ <https://fiware.org>

⁴ <https://internationaldataspaces.org/>

⁵ <https://gaia-x.eu/>

3 Data Summary

Following the Horizon Europe template of the DMP, this section summarises the the following points:

- Data generated and re-used
- Purpose of the data generation or re-use and its relation to the objectives of the project
- Origin and/or provenance of the data
- Expected size of the data
- Data formats
- Target groups that might find useful the data ('data utility'), outside your project

Accordingly, the template in Table 1 has been defined for the definition of the datasets (the table is also reported in Annex III). The DMP will be composed of the datasets that will be generated during the project lifecycle and reported according to the given template. Each dataset will be provided with an individual DMP using a unified form that simplifies the collection of the most relevant information (Annex I).

Table 1 - Template for the dataset

Partner	Partner name
Dataset	Reference Dataset number / name
Reference	Work Package WPx
Owner/Organization	Referring person who takes responsibility and his/her organization
Type of Data	Raw (Datasets, images, presentations, publications, software, etc...) Processed, Simulations, Metadata, Protocols, Methodologies...
Description and purpose	Data description and purpose of the data collection / generation and its relation to the objectives of the project.
Data Origin	Primary or secondary data, simulated, open data
Data Size	Preliminary estimation of the expected size of the dataset
Data Format	E.g., Text (.docx, PDF, etc...), Tabular (.xls, etc...), Image (.jpg, .TIFF, etc...), Video (.mov, .mp4, etc...), Scripts, codebases, or models
Tools	What methods or software tools are needed to access the DATA?
Utility & Re-use	Clarify to whom the data might be useful
F-Specific Metadata	Keyword(s) that categorize data to make it linked/searchable
F-Naming convention	What naming conventions do you follow? E.g., Convention internal to the organization; Project convention, etc

F-Standards	Reference to existing standards in topic area governing data collection, aggregation, storage and sharing. Adaptation of dataset to community standards to maximize interoperability with other researchers. Potential license restrictions. Discoverability. Need for aggregation and anonymization.
A- Availability	Public or Sensitive (restricted under conditions set out in the GA)
A- Accessibility	Source & Location, Access Restrictions (Note if Authentication needed, Embargo period, etc)
R-Archiving and preservation	Procedure for long-term preservation, length of preservation, an estimation of costs and how this will be covered. Data recovery methods and secure storage
R- Software tools	What methods or software tools are needed to access the data?
Ethical aspects	Are there any ethical or legal issues that can have an impact on data sharing? Is informed consent for data sharing and long-term preservation included in questionnaires dealing with personal data?
Data Security	Specific aspect related to sensitive or personal information

It is worth mentioning that since this document is delivered at M6, the corresponding tables for the datasets are not available, yet. In the following sections, an overview of the data that will be generated is given.

Annex III presents the template for general datasets, more information is given in Section 3.3 for each dataset. The data process has just started, the table with the collected data will be presented at M15 in the second version of this deliverable.

3.1 Description of the data

3.1.1 Data generated and re-used

The data acquisition and curation of IDEATION focuses on sensor data, location data, weather data, satellite data (e.g., from Copernicus), meteorological data, data input from water utilities and policy/decision makers, which will be acquired, collected, pre-processed for consistency and included in the project repository.

3.1.2 Data collection purposes

Data will be collected and processed during the project for the following purposes and in relation to the project objectives. In particular, focusing on the Phase I of the project, the following specific-objectives (SO) will be considered:

- **SO1** - To engage with relevant stakeholders involving all the above mentioned target groups in MSFs for the co-creation of the project solutions and the definition of viable use cases - **Stakeholders Map**
- **SO2** - To comprehensively create an inventory of the regulatory, research, and data landscape of inland water management, analysing policies, cataloguing scientific and



technological advancements, and creating inventories of open data and meteorological services for potential digital twin applications – **OpenKIWAS**

- **SO3** - To define the architectural concept and design, as a reference architecture, for the integration of inland waters in the DTO - **Reference architecture**
- **SO6** - To promote IDEATION at a wide level through suitable networking, clustering, and dissemination activities all along the project duration - **Publications**

3.1.3 Origin of the data

The data will originate from a variety of sources, including:

- **Survey Data:** collected through questionnaires and/or surveys, as well as those used during the MSF to gather information and data from the stakeholders. The survey used during the MSF will be reported in *D2.3: MSF results and functional and non-functional requirements (KER1) V1* at M9.
- **EU and international policies and standards** for inland waters, as well as project outputs related to inland water management.
- **Existing research data**, obtained from previously published studies, datasets, and scientific databases, such as public databases or other data provided by project partners from their internal databases and records, subject to the confidentiality conditions set in the Consortium Agreement or specific agreements signed to reinforce sharing internally the data.
- **Research and technological tools** from the scientific literature related to inland water management
- **Open data:** freely accessible, usable and shareable data made available to anyone.
- **Data from platforms or apps** (e.g., social media posts, crowd-sourcing, citizen science).

3.1.4 Target Groups (TG)

The data generated or re-used in this project may be useful to a broad audience, including the target groups introduced in the Document of Actions (DoA):

- **TG1_ARC:** Academia and research community: Hydrogeology, hydroinformatics and Digital Twin experts, data and computer scientists, AI researchers, data space experts, environmental scientists, biologists, socio-economists;
- **TG2_GOV:** Water regulators, river basin authorities, coastal authorities, government (local, regional, national and EU level), international organisation climate/global change & policy makers;

- **TG3_WUT:** Water Utilities, drinking water suppliers/wastewater managers and operators;
- **TG4_IND:** Industrial: water domain-related SMEs and large companies, IT solution integrators, data users/modelling experts;
- **TG5_PROJ:** National, European and International projects related to MISSION-OCEAN, IDEATION and the DTO;
- **TG6_CIV:** Citizens/Civil society/student community and media professionals (journalists);
- **TG7_NGO:** Foundations/NGO/environmental organisations.

3.2 Data generated in IDEATION

3.2.1 Stakeholders map

Existing stakeholders associated with the Water-oriented Living Labs (WoLLs) and the overall reference scenarios are currently gathered in a register that includes:

- Personal Information: Last Name, First Name, Gender, Email Address.
- Professional information: organisation, role in IDEATION project, involvement in WoLLs, expertise, involvement in the DTO other projects & networks.
- Expertise with particular reference to digital twins and Digital Twin Ocean (DTO).

The stakeholder map has been described in the deliverable *D2.1: Stakeholder registers for all reference scenarios V1* submitted at M3. A sketch of the register is given in Figure 1.



This is the Stakeholder Register for the IDEATION project.	
The aim is to identify relevant stakeholders (SH) to be involved during the project and in the various Reference Scenarios (RS)	
Please read this guide carefully	
The SH register contains several tabs	
Guide	Guidance on how to use the SH register (current tab)
Overview	A brief overview of all RS. This tab automatically collects information about the identified SHs from all RS tabs.
All RS	A list of SHs that are relevant to all RS. Use this tab to enter SHs that might be relevant for more than one RS or the IDEATION project as a whole.
RS# [Country]	A list of SHs for each RS. Use these tabs to enter SHs that might be relevant for a particular RS.

Figure 1 - Stakeholder register

3.2.2 OpenKIWAS

From M1, IDEATION started a comprehensive review to make an inventory of current solutions in the field of inland waters (rivers, lakes, reservoirs, wetlands, snow, and ice),

digital twin for water, and application of the DTO. The IDEATION inventory, i.e., the Open Knowledge Inventory for Inland Water Systems (OpenKIWAS), will include :

1. Policy, regulatory framework and project outputs: includes compiling and analysing EU and international policies, as well as project outputs related to inland water management.

- Collection and analysis of existing and upcoming EU policies, as well as policies from international organisations dealing with inland water. Especially those focusing on monitoring, forecasting, projecting, or simulating all aspects of the inland water cycle, including its physical state, chemical composition, geological features, and biological elements, both in their static (e.g., groundwater quality) and dynamic forms (e.g., river flows or water pressures). The policies include, among others, the EU directives related to surface and groundwater, drinking water, wastewater and treated water, ISO/CEN standards as well as UN/WHO recommendations. The collected data will be processed to extract key information under the IDEATION schema/ontology and stored as linked data to facilitate further processing and comparative studies.

For each policy the following information are collected:

- ID
- Name
- Issuing organisation
- Geographical Scope
- Type
- Scientific domain
- Type of water Water management focus
- Parameter change rate assumption
- Scope
- Impact
- Description
- Status
- Effect
- Date
- Resource
- Recommendations
- Parameter

- Change rate
- Collection of the metadata of outputs from past or ongoing research and innovation projects, funded by the European Union and other entities, dealing with inland water monitoring and management. These outputs include reports, datasets, software, hardware (e.g., sensors), platforms, research infrastructures, and end-user products and services. The analysis will document metadata from various funding programs (e.g., FP7, H2020, HEU, LIFE, Interreg) and initiatives like ERICs and open-access platforms such as WISE and Copernicus. Data collection will involve mapping projects, conducting online surveys, and analyzing information from digital systems.

For each project, the following information are stored:

- Project ID
- Project Acronym
- Project Name
- Funding Authority
- Call for proposals
- Focus Area
- Begin Date
- End Date
- Status
- Budget
- Description
- Resource
- Related policies
- Inland water types

2. Research and technological tools

- Systematic literature analysis of scientific results and reviews, related to monitoring, forecasting, and simulation of the inland water cycle. The results will be classified for the type of inland water studied, the focus of applications, and the geographical scope. A gap analysis will be conducted to identify what is missing, as well as study the relevance to policy making.

For each scientific result, the following information are stored:

- Type

- Year
 - Authors
 - Title
 - Venue/Journal
 - DOI
 - Research type
 - Water System
 - Technical Focus
 - Abstract
 - Link with Policy
 - Link with Projects
 - Link with Tools
- Collection of data models, AI technologies, simulation tools, intelligent systems, and digital twins, within the context of inland water. Specifically, data models (based on mathematical or computational foundations) may include hydrological models for predicting flow and quality, social models regarding water usage, hydraulic models for urban distribution etc. AI-based technologies may include machine learning algorithms for processing hydrological data or for analysing satellite images. Simulation tools use the data models to predict the outcome of different events. Intelligent systems consider cyber-physical architectures with sensors, actuators and software for monitoring, event detection, control, event management, risk assessment and decision support. Finally, Digital Twins are software that couples real data with models and simulators, to assist in evaluating scenarios and decisions on a realistic (virtual replica) of the system. This catalogue will include connections to endpoints to data and services made available via data spaces, such as the Wataverse and DSBA⁶.

For each data model, the following information are considered:

- Name
- Technology
- Type
- Data used as input
- Produced datasets (openly available)

⁶ <https://data-spaces-business-alliance.eu/dsba-releases-technical-convergence-discussion-document/>

- Demo (video if available)
- Paper (if available)
- Project ID (if available)
- Service description

3. Sensing and Open Data services

- Catalogue of inland water open data and sensing sources: selection of different elements that can enable the creation of digital twins relevant to inland water systems. This involves identifying the open data sources which include relevant environmental, socio-economic, and citizen data. These data may originate from sensors, local or remote (e.g., satellite images, weather stations, quality measurements), from surveys and research (e.g., census, land use, economic indicators), as well as from platforms or apps (e.g., social media posts, crowd-sourcing, citizen science). For each element, its quality, update frequency and accessibility will be evaluated, along with any legal restrictions, and recorded in its metadata, along with its endpoint.

For each open data, the following information are collected:

- Source organisation
 - Name
 - Data source
 - Quality
 - Availability
 - Update frequency
 - Legal restrictions
 - Description
 - Resource
 - API Protocol
- Catalogue of data and services from national meteorological organisations: identification and cataloguing of data and services provided by meteorological organisations that are relevant to inland waters, including weather forecasts, hydrological data, climate data, snow and ice data, evaporation rates. For selected European countries, the availability of data and their accessibility and interoperability will be assessed, as well as its relevance in the context of a digital twin.

For each collected data, the following information are saved:

- Source
- Organisation
- Name
- Type
- Area
- Quality
- Availability
- Update frequency
- Legal restrictions
- Description
- Resource
- API Protocol

3.2.3 Reference Architecture

The Reference Architecture will include the definition of architectural specifications and data models, the collection of good practices, the guidelines for the design and validation of physical simulation models and AI-algorithms, as well as technical recommendations for specific implementations and deployments of inland water DTs.

- Selection of frameworks tailored to individual tasks, and the specification of data structures to ensure seamless integration and maximum interoperability with the DTO and the DestinE Digital Twins
- Collection of data interoperability good practices and FAIR principles
- Guidelines for the design and validation of physical simulation models and AI-algorithms enabling the Digital Twin of inland waters: data preparation, model choice, validation procedures, adherence to regulations, and a focus on continual enhancement

3.2.4 Publications

Scientific information refers to peer-reviewed scientific research articles. Open dissemination strategies have been chosen to optimise a large use of the results of IDEATION in all addressed target groups.

The following publications are expected within the project's lifecycle:

- Peer-review journals: IDEATION partners envisage to publish at least 10 papers.

- Type of Access: Open Access self-archiving (Green Access) or Open Access publishing (Gold open access)
- Conferences (≥ 12 papers are envisaged to be published in conference proceedings)
 - Type of Access: conferences with open access strategies for the proceedings will be prioritised.

All academic Partners have endorsed Open Access and they will publish relevant research outputs on the Open Research Europe (ORE) platform as appropriate. Articles will be submitted via the single-page submission system provided by ORE.

Prepublication checks will be scanned to ensure that all policies and ethical guidelines are adhered to. Once the article has passed the pre-publication checks, the preprint version will be published, enabling immediate viewing and citation. Subsequently, experts will be selected and invited by ORE to review the paper. Their reviews and names will be published alongside the article, together with the authors' responses and comments from registered users. All versions of an article will be linked and independently citable. Finally, articles that pass peer review are sent to major indexing databases and repositories and made available on the project website.

ORE articles will be published under a Creative Commons by licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited, and leaves the copyright of the article with the current copyright holder. In IDEATION, data associated with ORE articles will be made available under the terms of a Creative Commons Public Domain Dedication. In doing so, data reuse will be facilitated and problems of attribution stacking when combining multiple datasets each authored by multiple authors avoided.

4 FAIR Data

Following the Horizon Europe Programme Guide, IDEATION will fully embrace the open access / open research data policy of Horizon Europe by providing online access to scientific information that is free of charge to end users and that is reusable, following the FAIR principles (see Figure 2): Findable, Accessible, Interoperable, and Reusable data. Moreover, IDEATION follows the MELODA5 dimensions to improve the quality of the Open Data (evaluate the Reputation and Dissemination actions over the data). Regarding the exploitation of the results, they will be made available in open source. Nevertheless, it may be necessary to apply open access restrictions to subsets of data derived from IP protection or security measures. In case of restrictions, they will be assessed and handled by CETAQUA and the whole consortium in *T6.4: Exploitation plan* and *T6.5: IP Management*.

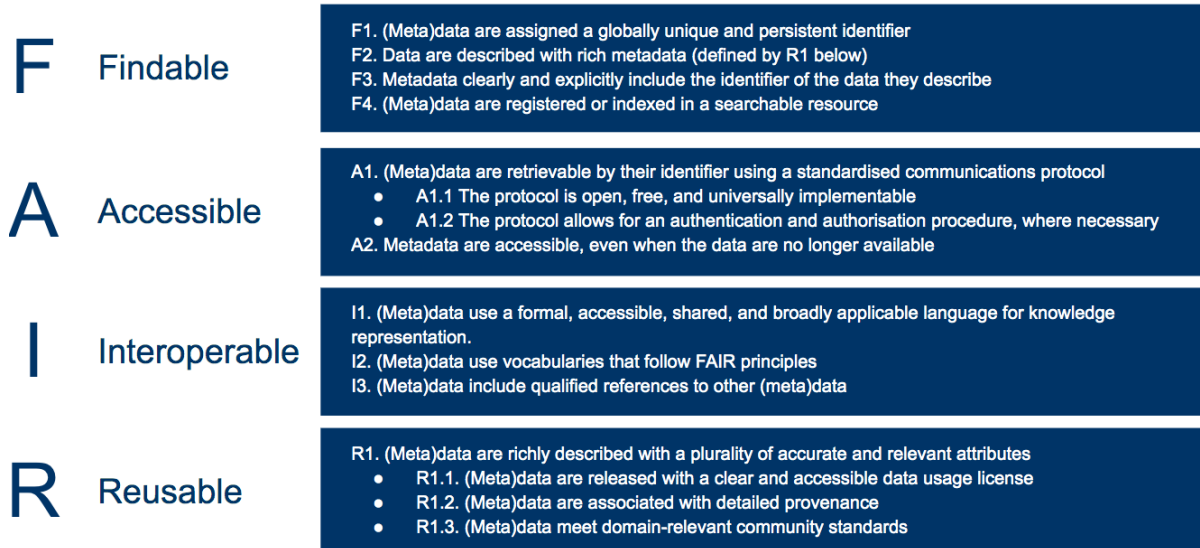


Figure 2 - FAIR principles

These principles are not themselves a standard, but rather a guide to data producers. The objective is to reach FAIRness of data which is a prerequisite for proper data management.

In the following sections, we will explain how the IDEATION consortium is going to ensure these FAIR principles concerning the data used and produced during IDEATION research activities.

4.1 Making Data Findable – Including Provisions for Metadata

The first step in (re)using data is to find them. Making data findable means that the data should be easy to find for humans and computers). This requires metadata to be made available online as an easily searchable source, and that data is assigned a persistent identifier specified in the metadata. Other means to achieve findability include having rich

metadata describing the data and having the metadata online in a searchable resource such as a catalogue or data repository.

Data discoverability (and metadata provision) IDEATION will generate all the metadata information in DCAT-AP format from the corresponding datasets and articles generated and published in Open Data Portals for consultation. To support the development of MSFs and keep track of their evolutions, ENG will provide an open digital platform for online collaboration and information sharing (i.e., Decidim). Finally, the Open-Source CMS that best suits the results produced is currently under selection and will be made available to the consortium after the necessary customizations to store all the data and information collected in OpenKIWAS, which will be also made available online through the IDEATION website.

Persistent identifiers (PIDs) will be used by all IDEATION partners to ensure the findability of research outputs, including data. They will be globally unique and long-lasting references to digital objects (e.g., data, publications, and research outputs) or non-digital objects (e.g., researchers, research institutions, grants).

To enhance the **findability of research outputs**, and their potential reuse, standardised metadata frameworks will be adopted, ensuring that data and other research outputs are accompanied by rich metadata that provides them with context. IDEATION will generate all the metadata information in DCAT-AP format from the corresponding datasets and articles generated and published in Open Data Portals for consultation.

To support the **development of MSFs** and keep track of their evolutions, Decidim (ENG), an open digital platform for online collaboration and information sharing.

The data collected during the project for the OpenKIWAS will be integrated into a Content Management System (CMS), a fundamental tool for organising and making information accessible in a more efficient and intuitive way. In fact, a Content Management System is not just a repository of data, but a platform that allows content to be structured, managed and presented according to specific usage needs. There are numerous types of CMS, each with unique characteristics. Of these, Drupal is a particularly interesting example, due to its open source nature and the wide range of possibilities it offers.

Drupal is distinguished above all by the strength of its community, a global network of developers who have been collaborating for years to create and constantly improve plugins, modules and extensions that extend the functionality of the platform. This aspect guarantees not only great flexibility, but also continuous technological evolution. One of the most appreciated advantages of Drupal is the presence of an access control system, known as the Access Control List, which allows access to content to be restricted according to the permissions assigned to users. This functionality makes it ideal for projects where security and customised data management are priorities.

Another strength of Drupal is its configurability. Depending on the needs of the project, it is possible to completely customise the way data is entered and made available to end users. This means that the system can be adapted to precisely reflect the specific needs of its users, providing a more fluid and organised experience. During the data import process, special attention will be paid to data transformation and cataloguing. In this phase, the raw data, represented as simple textual lists, will be reconfigured and converted into structured entities. This operation will not only improve the overall organisation of the system, but will also allow the full potential of the platform to be exploited. The entities created will be enriched with additional information and specific relationships, making the data more navigable, accessible and immediate for the user. Thanks to this structure, it will be possible to create different pages corresponding to different entities within the CMS. For instance, it will be possible to have pages dedicated to projects, project outputs, scientific results, and so on. This approach will allow users to navigate more precisely, accessing, filtering, or displaying only data related to specific categories of interest. The flexibility of Drupal is expressed above all in the possibility of creating customised views for each entity category. Views are a powerful model that not only allows data to be organised and presented clearly and dynamically, but also allows filters to be applied to deepen each detail. For instance, a user may choose to view only the results of a particular project or access a detailed list of scientific outputs filtered by year or author.

This system not only simplifies access to and display of information, but also encourages dynamic and participative content management, allowing users of the platform to actively contribute to its updating and growth, while respecting the predefined structure and organisation.

Finally, metadata about each dataset completed will be collected by the task leader in collaboration with the project coordinator (T1.3: Data requirements planning and management) through an online survey allowing to gather in a single operation both the metadata necessary to deposit the dataset into Zenodo if needed (see Annex III) and the metadata needed to update IDEATION DMP (see section 2.3.5). Moreover, some datasets will be made available through the IDEATION website, which will increase the discoverability of the data used and/or produced by IDEATION.

If a specific tool is needed to access/read/use the data put in a dataset, the partner owning this dataset will be requested to indicate it in the metadata table which will accompany each dataset (Annex I).

Naming Conventions

General documents produced by IDEATION partners shall be named so that the following information can be distinguished whenever possible:

- Project acronym: IDEATION (where applicable).
- Title of the document: descriptive of the contents of the document -

- Date of creation: DD-MM-YYYY (the date should be changed for different versions).
- Version number: version numbers should start at 0 (in case of early drafts) and continue in 1.0 and be incremented just by the author of the document by 0.1 for minor revisions, and by 1.0 for major revisions/new releases.

For Datasets we recommend the following convention:

- A prefix indicating the WP and task in which the dataset has been used/created
- A short and explicit name for the dataset
- An integer indicating the version of the dataset: **IDEATION_Data_WPx_Tx.x_Name_Vx**

For the specific case of deliverables, the name should follow the scheme: ***DX.Y: [Title of the Deliverable as set in the GA] vx.y.[docx/pdf]***

Search keywords for re-use optimization

The visibility and discoverability of the data that can be shared through the IDEATION website should increase by selecting the most appropriate keywords for SEO optimization. The partners are encouraged to provide adequate rich metadata within the data sets in order to ease the interpretation of the data and to increase the identification, discoverability, re-use and preservation thereof (See Annex I)

For those data sets that will be made available through Zenodo, automatically searching keywords are created by analysing the metadata info provided during the upload (see Annex II). It is therefore crucial that each partner owning a dataset and filling in the metadata survey takes special care to provide several keywords to describe his dataset, in order also to ease data discoverability.

Data clear versioning

Zenodo repository standardization ensures that data is stored under specific structure to be easily identified on a historical basis. It automatically assigns a DOI (Digital Object Identifier) to every record uploaded to its repository. This DOI ensures that IDEATION research outputs are persistently and uniquely identifiable, making it easier to cite and share.

Zenodo supports versioning as well, assigning a new DOI to the updated record, while maintaining a parent DOI for the entire collection.

Each time a data owner would like to opt for a restricted access to the dataset, he will be asked to explain why.

In such a case, a scientist who would see the dataset in Zenodo and who would be interested in having access to it, will have the possibility to contact the main author and motivate his request for access to the data.

4.2 *Making Openly Data Accessible*

As already explained, IDEATION will provide online access to scientific information that is free of charge to end users to facilitate sharing and re-use of results. Selected data and results will be shared with the scientific community and other stakeholders through publications in scientific journals and presentations at conferences, as well as through the open access to these data via the open Zenodo repository and/or the project website. As often as possible, and following the project policy concerning the IPR and open access, themselves directly linked to the rules enacted in the Grant Agreement and in the Consortium Agreement, the partners will be encouraged to give open access to the research data.

Access to restricted data

Some of the main restrictions assessed to consider the datasets as confidential and not accessible are:

- Protect IPR regarding new processes, products or technologies where the disclosure of the research data could have a negative impact in this competitive advantage (for example the grant of a patent is rejected due to an earlier disclosure of information);
- Security Concerns (it can affect strategic interests or assets of the EU);
- Personal Data that might have been collected in the project and can't be anonymised (i.e. stakeholders' lists).
- Other legitimate interests not covered under the above points

For data with access restrictions, a controlled access mechanism will be implemented. This may include:

- Access Request Forms: Interested parties may need to fill out a request form detailing their intended use of the data.
- Data Use Agreements: Users may be required to sign agreements outlining the terms and conditions of data use, ensuring compliance with legal and ethical standards.
- Secure Access Platforms: Data may be hosted on secure platforms that manage access permissions and track data usage.

Data access methods and tools

If a specific tool is needed to access/read/use the data put in a dataset, the partner owning this dataset will be requested to indicate it in the metadata table which will accompany each dataset (Annex I), including when the dataset is accessed thanks to an API or query language through an end point.

Data availability



In order to maximise the impact of IDEATION data, the project will facilitate sharing of results and deliverables within and beyond the consortium following open dissemination strategies:

- Open Access self-archiving (Green Access) or Open Access publishing (Gold open access) for publications in peer-review journals
- Regarding conferences, those with open access strategies for the proceedings will be prioritised.
- All academic partners have endorsed Open Access and they will publish relevant research outputs on the Open Research Europe (ORE) platform as appropriate. ORE articles will be published under a Creative Commons by licence.
- Zenodo, an open research data repository by European OpenAIRE, will be used as an open tool for collecting, organising, citing, and sharing IDEATION data.

Metadata:

All metadata generated by the IDEATION project will be made openly available and licensed under CC0. This ensures that metadata can be freely accessed, reused, and distributed by anyone without restrictions.

The metadata will include comprehensive information to enable users to access the data. This will include:

- Persistent Identifier (PID): Such as a DOI, which links to the dataset.
- Access Instructions: Details on how to access the data, including any necessary authentication or authorization steps.
- Contact Information: For further assistance or inquiries about the dataset.

The data will be made available and findable after the end of the project, in accordance with standard research data management practices. Metadata will be guaranteed to remain available even after the data itself is no longer accessible. This ensures that the existence and basic description of the data can be discovered, which is critical for transparency and reproducibility.

Comprehensive documentation will be provided, including references to any software required to access or read the data. This documentation will cover:

- Software Requirements: Names and versions of software needed.
- Usage Instructions: Detailed instructions on how to use the software to access and process the data.
- Open Source Code: Whenever possible, the relevant software will be included or linked as open-source code, ensuring that users can freely access, use, and modify the software as needed.



4.3 *Making Data Interoperable*

Data usually needs to be integrated with other data and to interoperate with applications or workflows for analysis, storage, and processing. Data should be readable for machines and humans without the need for specialised or ad hoc algorithms, translators, or mappings. This requires the adoption of a common language every participant uses, materialised in the adoption of common APIs for data exchange, the definition of common data models and common mechanisms for traceability of data exchange transactions and data provenance.

FAIR principles will be followed in the design of the IDEATION reference architecture to facilitate integration with the DTO and guarantee interoperability and the DestinE Digital Twin. A specific task in WP4 (T4.1) will define an integration layer with the corresponding tools, defining data formats, ensuring smooth data exchange, and addressing security. A set of interoperability guidelines and specifications on protocols, standards, APIs and data formats to be used, as well as guidelines for efficient integration will be released in M12 and M21.

To allow the adoption of IDEATION at a broader level, specific data exchange and interoperability issues will be addressed in T4.2 Collection of data interoperability good practices and FAIR principles. Thus, the project will lay down principles and guidelines that are firmly rooted in the idea of making data FAIR, as well as evaluate the general reusability, including Reputation and Dissemination actions over the data (MELODA5). These principles emphasise the importance of data being easily discoverable, readily accessible, and capable of seamless integration with external third-party systems. In addition to FAIR principles, IDEATION also considers other interoperability standards, i.e., WaterML⁷ and international initiatives like FIWARE⁸, IDSA⁹, GAIA-X¹⁰, BDVA¹¹ and all together under the DSBA¹², ensuring that the IDEATION digital twin ecosystem is collaborative, efficient, and poised for effective data exchange. Whenever possible as an implementation of these principles the generated datasets will include on one hand geographic information and attached to the metadata a description of the data model that makes easy for reusers to understand the elements of the data sources.

According to the catalogue of data models, AI technologies, simulations, intelligent systems and DT related to inland water that will be produced within T3.2.2, comprehensive guidelines will be defined to guide the development and validation of physical simulation models and AI algorithms (T4.3). These models and algorithms are the building blocks of the IDEATION digital twin for inland waters, and their accuracy, reliability, and applicability are

⁷ <https://www.ogc.org/standard/waterml/>

⁸ <https://www.fiware.org/>

⁹ <https://internationaldataspaces.org/>

¹⁰ <https://gaia-x.eu/>

¹¹ <https://bdva.eu/>

¹² <https://data-spaces-business-alliance.eu/>

of utmost importance. The guidelines span from data preparation to model selection, validation procedures, and compliance with regulatory standards. Moreover, they focus on continuous improvement and emphasise trustworthiness in the adoption of AI-based tools. Finally, during Phase I a robust, scalable, and interoperable architecture will be designed as input for the Phase II of the project (T4.4).

4.4 Enabling Data Reuse

After the project termination, the datasets will be made available for re-use through uploads to the dataset repository for the project. Documentation to validate data analysis and facilitate data re-use (e.g. readme files with information on methodology, codebooks, data cleaning, analyses, variable definitions, units of measurement, etc.) will be provided for each dataset and this will be updated in the individual DMPs (see Annex I)

As mentioned in chapter 2.2 Open Science Practice, ORE articles will be published under a Creative Commons by licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited, and leaves the copyright of the article with the current copyright holder. In IDEATION, data associated with ORE articles will be made available under the terms of a Creative Commons Public Domain Dedication. In doing so, data reuse will be facilitated and problems of attribution stacking when combining multiple datasets each authored by multiple authors avoided.

Protocols and templates will be defined to standardize data collection methods; this will ensure consistency in data formats, units of measurement, and variable definitions. Collected data will be submitted to validation checks, including automated scripts to detect anomalies (e.g., missing values) and manual reviews to confirm the accuracy of critical data. Systematic data cleaning processes may be applied to handle incomplete, inconsistent, or erroneous entries. In addition, data management will follow ethical guidelines, ensuring that datasets containing sensitive information are anonymized and stored securely.

Some data models generated out of the created data would be made available through the Smart Data Models (SDM¹³) initiative allowing the generation of compatible datasets. For those models that have generated data models at the SDM, validation can be performed for those datasets in json format automatically thanks to the json schema made available at SDM.

¹³ <https://smartdatamodels.org/>



5 Other Research Outputs

In addition to the management of data, IDEATION also considers and plans the management of other research outputs that will be generated or re-used throughout their projects. In fact, the OpenKIWAS, as already mentioned, collects project result outputs and their metadata, originating from different funding schemes including FP7/H2020/HEU, ERDF, LIFE, CEF, DEP, and Interreg. European Research Infrastructure Consortia (ERIC) will be also involved, with particular reference to those strictly related to IDEATION: The European Marine Biological Resource Centre (EMBRC), The European Multidisciplinary Seafloor and water column Observatory (EMSO), as well as the European contribution to the Argo programme (Euro-Argo) for its transversal objectives. Information from open-access European digital information systems and platforms related to inland waters will be analysed, including technical/operational programs such as the Water Information System of Europe (WISE)¹⁴ and the EU's Copernicus program for Earth Observation¹⁵, whose endpoints will be collected, and mapped into the catalogue. The catalogue will also include connections to endpoints to data and services made available via data spaces, such as the EU Wataverse project (ID: 101070262 - 2022-2025)¹⁶.

¹⁴ <https://water.europa.eu/>

¹⁵ <https://www.copernicus.eu/en>

¹⁶ <https://cordis.europa.eu/project/id/101070262>

6 Allocation of Resources

In order to deliver an estimation of the costs of making data FAIR, we will rely on future consortium discussions to identify the relevant project outputs and decide on ways and means of their open access (if applicable) or what data is useful to keep and for how long as well as for understanding costs and potential value. In the deliverable *D1.1: Quality assurance and Risk Management Plan v1*, roles and responsibility in data management has been given. For the sake of having a self-consistent deliverable, the management bodies and corresponding members are recalled in Table 2.

Table 2 - Management bodies

Management body	Members
Coordination Team	Coordinator, Financial controller <ul style="list-style-type: none"> • <i>Project coordinator</i>: Eloisa Vargiu (CETAQUA) • <i>Financial controller</i>: Public Programme Dept. (CETAQUA)
Scientific Technical Committee	One per partner plus the project coordinator: <ul style="list-style-type: none"> • Lydia Vamvakeridou-Lyroudia (KWR) • Gerasimos Antzoulatos (CERTH) • Natalia Zamora (BSC) • Demetris Eliades (UCY) • Roberto Di Bernardo (ENG) • Angeles Tejado (FIWARE) • Ana de León (WE) • Giovanni Coppini (CMCC) • Tobias Jonas (WSL) • Albert Chen (UNEXE)
Work Package Leaders	Coordinator and Work Package leaders (WPL): <ul style="list-style-type: none"> • <i>WP1</i>: Ana Paz Agudo (CETAQUA) • <i>WP2</i>: Stefania Munaretto (KWR) • <i>WP3</i>: Dimitrios Kouzapas (UCY) • <i>WP4</i>: Mirko Gallone (ENG) • <i>WP5</i>: Franco Crivello (CETAQUA) • <i>WP6</i>: Maria Mirachtsi (WE)
Thematic Managers	<ul style="list-style-type: none"> • <i>Innovation</i>: Demetris Eliades (UCY) • <i>AI Manager</i>: Natalia Zamora (BSC) • <i>IPR & Exploitation</i>: Miquel Sarrias (CETAQUA) • <i>Data</i>: Roberto Di Bernardo (ENG) • <i>Stakeholders engagement</i>: Stefania Munaretto (KWR) • <i>Communication and dissemination</i>: Maria Mirachtsi (WE)

In particular, the The Data Manager (DM) is responsible for:



1. Monitor data handling procedures and the FAIR principles: Findability, Accessibility, Interoperability, and Reusability.
2. Monitor data generated in IDEATION that will be acquired, collected, pre-processed for consistency, and included in the project repository.
3. Monitor data exploitation, curation, and sharing in IDEATION in accordance with the project Data Management Plan, DMP.
4. Ensure, together with the PC, the full compliance with the General Data Protection Regulation, GDPR, (Regulation (EU) 2016/679).

The Project Coordinator (CET), Data Manager (ENG) and the Dissemination Manager (WE) have a particular responsibility to ensure that data shared through the IDEATION website are easily and freely available, but also that backups are performed, and that proprietary data are secured.

7 Data Security

7.1 Data Publishing and Storage

As reported in *D1.1: Quality assurance and Risk Management Plan v1* for the coordination of the work, from M1, CETAQUA put at partners disposal a Shared Drive hosted by CETAQUA under the security constraints of the Veolia Group, to which CETAQUA belongs. During the live cycle of the IDEATION project, the data will be first stored in a cloud-based server for data storage. Profiles in trustworthy repositories are expected to be created for long-term preservation of the research data openly shared and available to any stakeholder interested in our research (excluding those affected by restrictions mentioned in the FAIR section).

The website of IDEATION is available for dissemination purposes (<https://ideation-euproject.eu/>, managed by WE).

Finally, as commented above, Zenodo, an open research data repository by European OpenAIRE, will be used as an open tool for collecting, organising, citing, and sharing IDEATION data. Therefore, all the datasets which are stored in Zenodo will be covered by Zenodo policies on longevity (<http://about.zenodo.org/policies/>).

7.2 Quality control of the data

During the project, quality control of the data will be the responsibility of ENG (WP4 leader) and CETAQUA (project coordinator). All project data will follow the FAIR principles (T4.2), their usage is approved by the respective water utilities and regulators at each site for all data sources included in the project. All the research data will be of the highest quality, have long-term validity and will be well documented in order other researchers to be able to get access and use them after 5 years.

The same attention to data protection will be placed in publication and dissemination of data and analysis to maintain the correct balance between scientific exploitation and data security.

A continuous, responsible and secure data handling will be kept during the project's life cycle in order to align with data privacy, security, accessibility requirements and in short:

- **Confidentiality:** Protecting sensitive data and ensuring that only authorised parties have access. This aligns with Horizon Europe's requirement to manage data "as open as possible, as closed as necessary." The IDEATION Data Manager and the Project Coordinator will ensure the full compliance with the General Data Protection Regulation, GDPR, (Regulation (EU) 2016/679) in the case of handling personal data.
- **Integrity:** Maintaining the accuracy and consistency of data throughout its lifecycle. Ensuring data quality is key, especially in projects where data may be reused by other researchers or policymakers.

- **Availability:** Ensuring that data is accessible to authorised users when needed. Horizon Europe projects often emphasise making data available for reuse under the FAIR principles, so addressing availability is a natural fit.

Any database containing data deemed to be sensitive will be encrypted using the current industry-standard level of encryption. Personal data, if any, will be anonymized upon completion of the project. In accordance with the Data Protection Act 2018, all research data will not be kept for longer than it is needed. The exchange of data between partners will be handled by a separate sharing service. This is a secure service, and all data must be encrypted using the current industry-standard to ensure maximum security during transmission.

Specific responsibilities regarding Data Protection are identified in the Consortium Agreement:

Where necessary, the Parties shall cooperate in order to enable one another to fulfil legal obligations arising under applicable data protection laws (the Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data and relevant national data protection law applicable to said Party) within the scope of the performance and administration of the Project and of this Consortium Agreement.

In particular, the Parties shall, where necessary, conclude a separate data processing, data sharing and/or joint controller agreement before any data processing or data sharing takes place.

8 Ethical Aspects

The IDEATION Grant Agreement addresses the ethical considerations related to the objectives, methodology, and anticipated impact of the project's results, ensuring compliance with ethical principles and relevant legislation. Ethics and Data Protection Monitoring, along with the technical robustness and ethical considerations of AI-based tools, will be detailed in the respective deliverables: Research Ethics and Data Protection Monitoring Report V1 and V2.

The interaction with third parties and the processing of their personal data will be performed via informed consent documents In accordance with the [EU General Data Protection Regulation](#).

A Personal Data Consent Form is used each time IDEATION partner organisations need to collect, store and manage personal data from people who participate in project events or other activities.

Deliverables *D1.7: Research Ethics and Data Protection Monitoring Report V1* at M12 and *D1.8: Research Ethics and Data Protection Monitoring Report V2* at M24 will go into details of ethics and data protection.

9 CONCLUSIONS

The IDEATION Data Management Plan (DMP) establishes a framework for managing data throughout the project's lifecycle, adhering to the FAIR principles (Findable, Accessible, Interoperable, and Reusable) and ensuring compliance with ethical standards and EU regulations.

In this first iteration, the DMP focuses on Phase I activities (M1–M12) until M6, explaining the types of data that will be primarily handled (publicly available information, including policies, scientific research outputs, datasets, and tools relevant to inland waters and digital twins). Special attention is given to metadata quality and data interoperability standards to facilitate future integration.

IDEATION emphasises open access to results, publishing datasets and outputs in repositories like Zenodo, with assigned DOIs for persistent identification and easy dissemination. It also adheres to high ethical standards, ensuring data privacy, secure handling, and compliance with GDPR. Ethical aspects of AI-based tools and their adoption are also addressed.

The next version of the deliverable at M15 will update and expand the current one, incorporating feedback and addressing data management challenges as the project progresses.

ANNEX I - Dataset Template

Metadata provided for each dataset to feed IDEATION DMP

Partner	Partner name
Dataset	Reference Dataset number / name
Reference	Work Package WPx
Owner/	Referring person who takes responsibility and his/her organization
Type of Data	Raw (Datasets, images, presentations, publications, software, etc...) Processed, Simulations, Metadata, Protocols, Methodologies...
Description and purpose	Data description and purpose of the data collection / generation and its relation to the objectives of the project.
Data Origin	Primary or secondary data, simulated, open data
Data Size	Preliminary estimation of the expected size of the dataset
Data Format	E.g., Text (.docx, PDF, etc...), Tabular (.xls, etc...), Image (.jpg, .TIFF, etc...), Video (.mov, .mp4, etc...), Scripts, codebases, or models
Tools	What methods or software tools are needed to access the DATA?
Utility & Re-use	Clarify to whom the data might be useful
F-Specific Metadata	Keyword(s) that categorize data to make it linked/searchable
F-Naming convention	What naming conventions do you follow? E.g., Convention internal to the organization; Project convention, etc
F-Standards	Reference to existing standards in topic area governing data collection, aggregation, storage and sharing. Adaptation of dataset to community standards to maximize interoperability with other researchers. Potential license restrictions. Discoverability. Need for aggregation and anonymization.
A- Availability	Public or Sensitive (restricted under conditions set out in the GA)
A- Accessibility	Source & Location, Access Restrictions (Note if Authentication needed, Embargo period, etc)
R-Archiving and preservation	Procedure for long-term preservation, length of preservation, an estimation of costs and how this will be covered. Data recovery methods and secure storage
R- Software tools	What methods or software tools are needed to access the data?
Ethical aspects	Are there any ethical or legal issues that can have an impact on data sharing? Is informed consent for data sharing and long-term preservation included in questionnaires dealing with personal data?
Data Security	

ANNEX II - Zenodo Template

File information								
File Name							Required	
Upload of the file (maximum size: 50 GB)	File addition, removal or modification are not allowed after the publication of the upload, because a Digital Object Identifier (DOI) is registered with DataCite for each upload.							
Communities								
Community name	IDEATION						Recommended	
Upload type								
Dataset type	Publication <input type="checkbox"/>	Publication type						Required
		<input type="checkbox"/> Book	<input type="checkbox"/> Preprint	<input type="checkbox"/> SW documentation				
		<input type="checkbox"/> Book section	<input type="checkbox"/> Project Deliverable	<input type="checkbox"/> Thesis				
		<input type="checkbox"/> Conference paper	<input type="checkbox"/> Project Milestone	<input type="checkbox"/> Technical note				
		<input type="checkbox"/> Journal Article	<input type="checkbox"/> Proposal	<input type="checkbox"/> Monitoring paper				
		<input type="checkbox"/> Patent	<input type="checkbox"/> Report	<input type="checkbox"/> Other				
Poster <input type="checkbox"/>	Presentation <input type="checkbox"/>	Dataset <input type="checkbox"/>	Video/Audio <input type="checkbox"/>	Software <input type="checkbox"/>	Lesson <input type="checkbox"/>	Other <input type="checkbox"/>		
Image <input type="checkbox"/>	Image type							
	<input type="checkbox"/> Figure	<input type="checkbox"/> Plot	<input type="checkbox"/> Photo					
	<input type="checkbox"/> Diagram	<input type="checkbox"/> Drawing	<input type="checkbox"/> Other					

Basic information		
Digital Object Identifier (DOI) (Optional)	If a publisher already assign a DOI to the dataset (for scientific publication for example), indicate it. If not, no need to fill in this field as Zenodo will automatically register a new DOI.	Optional
Publication Date		Required
Title		Required
Authors	Family Name, given names	Required
	Affiliation	Required
	ORCID (It provides a persistent digital identifier that distinguishes you from every other researcher and, through integration in key research workflows such as manuscript and grant submission, supports automated linkages between you and your professional activities ensuring that your work is recognized" https://orcid.org/)	Optional
Description		Required
Version/Language		Optional
Keywords		Optional
Additional notes		Required

ANNEX III - Template for General Datasets

Dataset	Description / Purpose	Owner	Data Origin	Type	Format	Size	Tools ¹⁷	Use ¹⁸	Type of Access: open / restricted	If restricted, explain why	Specific repository ¹⁹
WP1 – T1.1:	Administrative and financial management - CETAQUA										
WP1 – T1.2:	Quality assurance and risk management - CETAQUA										
WP1 – T1.3:	Data requirements planning and management - FIWARE										
WP1 – T1.4:	Establish and manage an external advisory board - CETAQUA										
WP2 – T2.1:	Stakeholder mapping - WE										
WP2 – T2.2:	MSFs for understanding stakeholder needs, gaps, and solutions requirements - KWR										
WP2 – T2.3:	Definition of reference use cases CETAQUA										
WP2 – T2.4:	MSFs for understanding stakeholders' priorities, acceptable solutions, and policy implementation Recommendations - KWR										
WP3 – T3.1.1:	Catalogue of EU and international policies and standards for inland waters - UNEXE										
WP3 - T3.1.2:	Catalogue of projects and information systems outputs - CETAQUA										
WP3 – T3.2.1:	Catalogue of scientific research results - UCY										
WP3 – T3.2.2:	Catalogue of data models, AI technologies, simulations, intelligent systems and Digital Twins related to inland water - UCY										
WP3 – T3.3.1:	Catalogue of inland water open data and sensing sources - CERTH										

¹⁷ (methods or software tools needed to access / use the data)

¹⁸ (For whom could these data be useful, outside the Project consortium?)

¹⁹ (Does your organisation use a specific repository for data? if yes, which one?)



IDEATION

www.ideation-project.eu

WP3 – T3.3.2: Catalogue of data and services from national meteorological organisations - CMCC
WP4 – T4.1 Technical specifications and data models towards seamless integration with DTO and DestinE Digital Twins - BSC
WP4 – T4.2 Collection of data interoperability good practices and FAIR principles - FIWARE
WP4 – T4.3 Guidelines for the design and validation of physical simulation models and AI-algorithms enabling the Digital Twin of inland waters - CETAQUA
WP4 – T4.4 Designing an optimal inland water Digital Twin high-level reference architecture - ENG
WP5 – T5.1 Mapping between functional and no-functional requirements and current solutions - CERTH
WP5 – T5.2: SWOT analysis - UNEXE
WP5 – T5.3: Technical governance for managing inland waters - UNEXE
WP5 – T5.4: IDEATION monitoring and assessment – CETAQUA
WP5 – T5.5: Final recommendations and roadmap - CETAQUA
WP6 - T6.1: Building a communication and dissemination strategy - WE
WP6 – T6.2: Communication and dissemination activities - WE
WP6 – T6.3: Set-up and promotion of networking activities - WE
WP6 – T6.4: Exploitation plan - CETAQUA
WP6 –T6.5: IP Management - CETAQUA





IDEATION to prepare the development of the digital twin of the inland waters (rivers, lakes, reservoirs, wetlands, snow, and ice) addressing activities to be developed and to make it integrated and interoperable with the DTO for a unified digital twin of ocean and waters.



IDEATION



**Funded by
the European Union**



IDEATION



Funded by
the European Union