Innovative and Sustainable Groundwater Management in the Mediterranean

D1.4 DATA MANAGEMENT PLAN

VERSION 2.0

Acknowledgment: This project is part of the PRIMA Programme supported by the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No 1923.

DOI: 10.5281/zenodo.6402256

Project Information

Project Title	Innovative and Sustainable Groundwater Management in the Mediterranean		
Project Acronym	InTheMED	Grant Agreement Number	1923
Program	Horizon 2020		
Type of Action	Water RIA – Research and innovation Action		
Start Data	March 1, 2020	Duration	36 months
Project Coordinator	Universitat Politècnica de València (UPV), Spain		
Consortium	Universitat Politècnica de València (UPV), Spain (Coordinator) Helmholtz-Zentrum für Umweltforschung (UFZ), Germany Università degli Studi di Parma (UNIPR), Italy Boğaziçi Üniversitesi (BU), Turkey Centre de Recherches et des Technologies des Eaux (CERTE), Tunisie Technical University of Crete (TUC), Greece Associacao do Instituto Superior Tecnico para a Investigaçao e Desenvolvimiento (IST-ID), Portugal		

Document Information

Deliverable Number	D1.4	Deliverable Name	Data Management Plan		
Work Package number	WP1	Work Package Title	Innovative Project Management in the MED		
Due Date	Contractual	August, 2021	Actual	March, 2022	
Version Number	2.0				
Deliverable Type	Report (R)	Dissemination Level	Public (PU)		
Authors	Janire Uribe-Asarta Vanessa A. Godoy J. Jaime Gómez-Hernández				
Reviewer(s)	All				

Document History

Version	Date	Stage	Reviewed by
1.1	2022/02/24	First draft	All
1.2	2022/03/22	Inputs from all partners have been added	Janire Uribe-Asarta Vanessa A. Godoy
1.3.	2022/03/28	Second draft	Janire Uribe-Asarta Vanessa A. Godoy J. Jaime Gómez-Hernández
2.0	2022/03/31	Final version	Janire Uribe-Asarta Vanessa A. Godoy J. Jaime Gómez-Hernández

Table of Contents

List of Tables

List of Figures

Glossary

BU	Boğaziçi Üniversitesi.
CERTE	Centre de Recherches et des Technologies des Eaux.
CO	Confidential.
DEC	Websites, patents filling, press and media actions, videos, etc.
DMP	Data Management Plan.
EEA	European Environment Agency.
GA	Grant Agreement.
GEMStat	The global water quality database.
IGRAC	International Resources Assessment Centre.
INSPIRE	Infrastructure for Spatial Information in the European Community.
IST-ID	Associação do Instituto Superior Técnico para a Investigação e Desenvolvimento.
IPR	Intellectual Property Rights.
MED	Mediterranean.
NGO	Non-governmental organizations.
OTHER	Software, technical diagram, etc.
PI	Principal Investigator.
PU	Public.
R	Document, report.
SME	Small and Medium-sized enterprises.
TUC	Technical University of Crete.
UFZ	Helmholtz-Zentrum für Umweltforschung.
UNIPR	Università degli Studi di Parma.
UPV	Universitat Politècnica de València.
WISE	Water Information System for Europe.
WP	Work Package.

Executive Summary

The overall objective of the InTheMED project is to implement innovative and sustainable management tools and remediation strategies for MED aquifers (inland and coastal) in order to mitigate anthropogenic and climate-change threats by creating new long-lasting spaces of social learning among different interdependent stakeholders, NGOs, and scientific researchers in five field case studies. These are located at the two shores of the MED basin, namely in Spain, Greece, Portugal, Tunisia, and Turkey.

InTheMED will develop an inclusive process that will establish an ensemble of innovative assessment and management tools and methodologies including a high-resolution monitoring approach, smart modelling, a socio-economic assessment, web-based decision support systems (DSS) and new configurations for governance to validate efficient and sustainable integrated groundwater management in the MED considering both the quantitative and qualitative aspects.

This Deliverable is the second version of the Data Management Plan (DMP) and an update of the one presented in March 2021. The DMP aims to consolidate the data management strategy of the InTheMED project. The DMP describes the data collected and produced within the project and provides a guide on how this data will be managed, stored, disseminated and preserved. This updated version answers specific topics that remained unresolved in the first version. The DMP is considered to be a living document that is expected to evolve during the project.

1. Introduction

The Data Management Plan (DMP) is the deliverable number D1.4 of the "Innovative and Sustainable Groundwater Management in the Mediterranean" Grant Agreement Number 1923 project.

This document is the second version of the DMP presented in March 2021, and an update of the data management guidelines considering the progress made in the InTheMED project to date. Moreover, the updated version of the DMP answers specific topics that remained unresolved in the first version, such as those related to the size of the data.

Following the recommendations provided by the European Commission, this deliverable has been prepared considering the template of the "Guidelines on FAIR Data Management in Horizon 2020" [1], and it covers:

- Type, format and size of collected, processed and generated data.
- Handling of the data during and after the end of the project.
- Definition of the dissemination level of the data: open access or confidential.
- Methodologies and standards to be applied.
- Curation and preservation of the data during and after the end of the project.

Its primary use will be as a support tool for the InTheMED Consortium partners to efficiently handle data and comply with the Open Research Data Pilot initiative of the European Commission. As it is a public document, it will also serve to inform other researchers, professionals and the broad society about where to find the data and results of the InTheMED project.

The DMP is considered a living document, meaning that it will be accurately modified and updated when significant changes occur, such as the production of new data, changes in consortium policies or changes in consortium composition and external factors. At least, there will be one more version at the project end.

2. Data Summary

The second chapter of the DMP has the objective of presenting the data created and collected throughout the execution of the InTheMED project. The data collection of each Work Package (WP) is separately displayed.

2.1. WP 1: Innovative Project Management

Purpose

The first WP, of which the lead participant is the UPV team, is responsible for the administrative and financial aspects of the project as well as for following the progress and fulfilment of the objectives and deliverables.

Type, Format and Expected Size of Data

- Meeting minutes: Efficient communication between partners and good control on work progress are being assured by contact by telephone, email, video calls and cloud-based project management tools such as Slack or Trello. We have held a kick-off meeting, a Project Management Board meeting, two Steering Committee Meetings, a Scientific Advisory Board meeting and a mid-term meeting. UPV team will continue to organise the necessary meetings during the second half of the project and schedule a final meeting. The final meeting will also serve as a showcase of the project outputs to all stakeholders. Furthermore, annual coordination meetings are being held where advisory members participate. The initial plan was that the meeting venues would rotate among the partner countries but given the SARS-CoV-2 induced situation, the meetings of the first half of the project were shifted to virtual meetings. Given the evolution of the pandemic, an attempt will be made to start face-to-face meetings. In these meetings, the activities carried out and the risks in reaching the InTheMED objectives will be discussed. The decisions and comments discussed during these meetings are being documented in the minutes of the meetings. The minutes are saved as a text file (.docx) or as a PDF document (.pdf). The size of each document is about 300 KB. During these meetings, presentations are shown in Microsoft PowerPoint format (.pptx).

- Deliverables: Table 1 summarises the deliverables that have already been created within WP1. The lead author of each deliverable is responsible for writing the document and

delivering it to the rest of the Consortium as a text file (.docx). After reviewing, the final version is saved as a PDF document (.pdf).

Table 1. Deliverables of WP1

Deliverable Number and Title	Lead Participant	Type	Diss. Level	Delivery date	Format and size
D1.1 Consortium Agreement	UPV	R	CO	2021/05/14	.pdf (1,218 kB)
D1.2 First, second and third periodic reports	UPV	R	PU	2021/05/14	.pdf (996.1 kB)
D1.3 Documentation of kick-off, annual coordination and WP leader meetings	UPV	R	CO	2021/09/15	.pdf (500 kB)
D1.4 Data management plan v1	UPV	R	PU	2021/03/03	.pdf (779.4 kB)

Re-Use of Existing Data and its Origin

Re-use of existing data is not expected.

Data Utility

These data are primarily useful for the InTheMED Consortium partners. In the case of the DMP, it is also valuable for researchers, professionals and the general public interested in the activities related to the InTheMED project.

2.2. WP 2: Innovative Monitoring and Data Analysis in the MED

Purpose

The second WP, led by the UFZ team, has a threefold objective that is the acquisition of historical MED data, the analysis and sharing of collected data, and the enrichment of data availability, using the High-Resolution Monitoring Approach (HRMA), which allows near real-time monitoring, in case studies characterized by limited data collections. The characterisation of the five case studies will strengthen the understanding of groundwater functioning and long-term groundwater trends and, therefore, contribute critical data for the smart models (WP3). Moreover, the hypothesis is that understanding past groundwater history can help comprehend aquifers' current state and predict future evolution (WP5).

- Groundwater quality and quantity datasets: CERTE team has installed three sensors and has connected them to a platform that displays real-time measures in the three selected sites in Tunisia. They also receive daily three Comma-separated values (.csv) files summarizing the hourly measurements of the qualitative parameters of the three sensors via an FTP (File Transfer Protocol) server with a size of 4,7 KB per file. As well they have long-time series data of groundwater level at 35 observation points, water quality data from the last decade at seven observation points, and five-year averages of the groundwater exploitation rate from 1980 until 2015. These data were collected from the Regional Commissariat of Agricultural Development of Nabeul and from the General Directorate of Water Resources and are in Microsoft Excel file format (.xlsx) with a total size of 311 KB. Moreover, the sampling campaigns carried out during the project allowed them to identify and collect an updated information on the quality of groundwater and rivers in some monitoring points in the study area. These data are in Microsoft Excel file format (.xlsx) with a size of 20 KB. UPV team has historical data of groundwater level monitoring in 10 observation points. These data were provided by the Spanish Water Authority and are in a Comma-separated values (.csv) file with a size of 24.1 KB. IST-ID team has daily historical data of water electrical conductivity at four borehole locations for a two-month period. The data was provided by SOMINCOR (Sociedade Mineira de Neves Corvo SA) and is stored in a Comma-separated values (.csv) file with a size of 3 KB. UFZ team is developing a database of groundwater level time series of countries in the Mediterranean region. Aiming to geographically cover as much extension as possible of the region, including countries of the study cases and others, data have been collected for Spain, Portugal, Italy, Turkey, Greece, France and Tunisia with more than 12,000 piezometers. They are being processed to have a homogeneous format of monthly standardized water table depth time series in comma-separated values. The final database, including raw and processed data and metadata is expected to have a size of around 10 GB.

- Geophysical data: IST-ID has access to 63 two-dimensional resistivity profiles distributed within the perimeter of the mining site. These data are stored in .dat files and comprehend apparent resistivities, initial electrical conductivity models and inverted profiles

with deterministic approaches. In total these files have a size of 53.1 MB. An additional file with the start and end of profile for 22 profiles is also available. This is a text file (.txt) and has a size of 2KB.

- Geophysical models: IST-ID team inverted with geostatistical DC inversion method all the 22 geophysical profiles. **¡Error! No se encuentra el origen de la referencia.** summarizes the size of all the products generated with the inversion method. These products comprise geostatistical simulation files in GEOEAS (Geostatistical Environmental Assessment Software) format. The code of the inversion method was developed in MATLAB and has a series of functions stored as scripts in .m files of variable size.

Table 2. Size of the inverted profiles with geostatistical DC inversion method

Profile	Size (MB)
P15	20.8
P16	14.3
P17	18.6
P18	12.1
P19	12.1
P20	11.6
P21	18.6
P22	20.8
P23	9.95
P24	20.8
P25	9.95
P26	9.95
P27	4.48

P28	4.48
P29	2.27
P30	4.48
PN20	12.1
PN21	11.7
PN22	4.48
PN23	4.48
PN24	13.9

An experimental test with the UNIPR team to assess the reliability of modelling a contamination plume from geophysical data in a laboratory experiment is also being performed. We expect the data and models to be delivered in the same type of format.

- Deliverables: In addition to the datasets mentioned above, table 3 summarises the deliverables within WP2. The lead author of each deliverable is responsible for writing the document and delivering it to the rest of the Consortium as a text file (.docx). After reviewing, the final version is saved as a PDF document (.pdf).

Table 3. Deliverables of WP2

Deliverable Number and Title	Lead Participant	Type	Diss. Level	Delivery date	Format and size
D2.1 Report on the integrated and innovative high-resolution monitoring strategies in the different case studies	IST-ID	Other	PU	2021/08/31	.pdf (1.4 MB)
D2.2 Report on the existing historical groundwater data on the MED region	UFZ	R	PU	Not delivered yet	-
D2.3 Report on regional groundwater trends and their	UFZ	R	PU	Not delivered	-

controlling factors				yet	
D2.4 Reinforcement of the systematic monitoring and data sharing	UFZ	DEM	PU	Not delivered yet	-

Re-Use of Existing Data and its Origin

- Data from InTheMED project: WP2 and WP5 are jointly working on the definition of groundwater problems and the identification of hotspots and key stakeholders.

- Existing groundwater quality and quantity data: existing groundwater quality and quantity datasets of the five case studies are being investigated in order to select, based on the causal DPSIR (Drivers-Pressure-State-Impact-Responses) framework adopted by the European Environmental Agency, those time series that are more relevant for the transferability and reproducibility of the InTheMED concept to other countries. These long-time historical series are being compiled and saved in Microsoft Excel documents (.xlsx), comma separated values files (.cvs), text files (.txt), and ESRI shapefiles (.shp). The Consortium partners' sites, as well as national, regional and global water quality databases are being explored such as:

  - Greek database: Hellenic Republic. Decentralized Administration of Crete (https://www.apdkritis.gov.gr/en/open-data).
  - Spanish database: Redes de seguimiento. Ministerio de Agricultura, pesca y alimentación. Ministerio para la transición ecológica. (https://sig.mapama.gob.es/redes-seguimiento/).
  - Portuguese database: Sistena Nacional de Infromação de Recursos Hídricos. (https://snirh.apambiente.pt/).
  - French database: Portail national d'accès aux données sur les eaux souterraines. (https://ades.eaufrance.fr/).
  - Tunisian open database: (http://www.onagri.nat.tn/)

- WISE/WRRL and EEA: (https://www.eea.europa.eu/data-and-maps/indicators/nutrients-in-freshwater/nutrients-in-freshwater-assessment-published-6).
- GEMStat: (https://gemstat.org/).
- IGRAC: Global Groundwater Monitoring Network (GGMN) (https://www.un-igrac.org/special-project/ggmn-global-groundwater-monitoring-network).

Data Utility

The results obtained in WP2 are essential for the correct progress of the tasks involved in the rest of the work packages. It mainly has a close connection with WP3, WP5 and WP6. Consequently, all the members of the Consortium have access to this information. It is also of great interest to researchers and professionals in the field of hydrology, hydrogeology and other earth sciences, as well as, to public or private institutions related to water management and use. Furthermore, it may interest the general public, NGOs and those who work in the field of agriculture.

2.3. WP 3: Innovative Smart Modelling in the MED

Purpose

The third WP, led by the UNIPR team, is creating simplified models, one for each case study, which are called surrogate models, meta-models or smart models. These models are built based on long-time historical data, detailed numerical models and expertise from the partners. Its aim is to provide specific answers to the stakeholders regarding the analysis of alternative scenarios and making decisions under uncertain future conditions, where the effects of future climate and anthropogenic changes are considered.

Type, Format and Expected Size of Data

- Input data for the training and validation of the smart models: The UPV team used as input data the recharge, pumping, month, day, year, and hydraulic head from a numerical model ran in MODFLOW 2005 by using its Python interface (Flopy). These data were saved as Microsoft Excel file documents (.xlsx) with a size of 5.40 MB. The UNIPR team used as input

data the precipitation, temperature and groundwater level for the development of the data-driven surrogate model for the Tunisia study case. Precipitation, temperature and groundwater level data were saved in Microsoft Excel file format (.xlsx) with sizes 328 KB, 171 KB and 328 KB, respectively. The TUC team used as input data the observed hydraulic head in dry and wet hydrological period and surface elevation by using space-time geostatistics in R. Input data were saved as Microsoft Excel file documents (.xlsx) with a size of 2 MB.

- Procedures and computer codes for the training and validation of the smart models: The UPV team used Python scripts (.py) with a size of 29 KB to implement their smart model. The UNIPR team used MATLAB scripts (.m) with a total size of 30 KB to implement the Tunisia smart model. The TUC team used R scripts with a size of 50 KB to implement their smart model.

- Smart models to be used externally and hosting: The UPV team has generated 15 models in .pkl format with an approximate size of 33 KB each ,the smart models were deployed via Flask on Heroku. The UNIPR team will deploy the Tunisia smart model, implemented in Excel spreadsheet, on Zenodo; the TUC team generated geostatistical models with an approximate size of 100 KB; the smart models will be deployed on Zenodo.

- Video tutorials: how-to-use video tutorials will be produced to explain the modelling tools developed in the present WP3. These files are expected to be in AVI (.avi) or MP4 (.mp4) format.

- Input data for the downscaling of future climate projections at the case-study scale: The UNIPR collected the precipitation and temperature data provided by 17 Regional Climate models. These data were saved as MAT-files (.mat) with total size of 1.02 GB for Requena-Utiel pilot site, 505 MB for Tympaki, 8.81 GB for Konya, 527 MB for Grombalia and 3.62 MB for Castro Verde. For Requena-Utiel study case, the input field data (precipitation and temperature) used to perform the bias correction of the climate models were saved as text file (.txt) and Microsoft Excel file (.xlsx) with a total size of 24.8 MB. For the other study cases, the data were saved as Excel file (.xlsx) with a total size of 2.02 MB for Tympaki, 20.92 MB for Konya, 1.70 MB for Grombalia and 7.83 MB for Castro Verde.

- Procedures and computer codes for the downscaling of future climate projections at the case-study scale: The UNIPR team used Matlab scripts (.m) with a total size of 78 KB to perform the gap-filling of the data field series and the downscaling and bias correction of the Regional Climate Models.

- Deliverables: Table 4 summarises the deliverables that are created within WP3. The lead author of each deliverable is responsible for writing the document and delivering it to the rest of the Consortium as a text file (.docx). After reviewing, the final version is saved as a PDF document (.pdf).

Table 4. Deliverables of WP3

Deliverable Number and Title	Lead Participant	Type	Diss. Level	Delivery date	Format and size
D3.1 Identification of the surrogate model to be applied in the case studies	UPV	R	PU	2021/09/02	.pdf (941.1 KB)
D3.2 Report on surrogate models in the case studies	UNIPR	R	PU	Not delivered yet	-
D3.3 Data Archive containing the downscaled climate projections in the case studies	UNIPR	R	CO	Not delivered yet	-
D3.4 Report on the results of the analysis of different scenarios in the case studies	UNIPR	R	PU	Not delivered yet	-

Re-Use of Existing Data and its Origin

- Input data for models: The hydraulic heads for the Spanish case study, the Requena-Utiel aquifer, are freely accessible from the website of the Ministry for Ecological Transition and Demographic Challenge (Ministerio para la Transición Ecológica y el Reto Demográfico, MITECO) of the Spanish Govern: https://www.miteco.gob.es/es/cartografia-y-sig/ide/descargas/agua/red-piezometrica.aspx or https://www.miteco.gob.es/es/cartografia-y-sig/ide/descargas/basedatospiezometria_tcm30-533415.zip. For Grombalia case study, the precipitation and groundwater level data have been provided by CRDA (the Regional

commissariat of agricultural development of Nabeul) and temperature data by INM (national institute of meteorology). For the Tympaki case study, the groundwater level data have been downloaded from the Water Resources Portal of the Decentralized Administration of Crete, 2022 (https://www.apdkritis.gov.gr/en) and the National Groundwater Monitoring Systems of the Hellenic Survey of Geology & Mineral Exploration, 2020 (http://geodata.gov.gr/maps/?package=42b085a2-d390-4bd4-a9a0-6f1b3ec2ee9b&resource=646ab5b9-71ed-4b35-934f-78eec37978d1&locale=el).

- Future projections of climate variables and climate historical data: the precipitation and temperature projections have been acquired from 17 Regional Climate Models (RCM), which are part of the EURO-CORDEX PROJECT (https://www.euro-cordex.net/). The scenarios named Representative Concertation Pathways (RCPs), adopted by the ICPP for the Fifth Assessment Report (AR5), in particular, the RCP4.5 and the RCP8.5, are considered. The historical data used to perform the bias correction of the RCMs are precipitation and temperature collected from national repositories. For Requena-Utiel, the data have been downloaded from SAIH (The Automatic Hydrological Information System http://saih.chj.es/chj/saih/glayer?t=p). For Tympaki, the data have been provided by the automatic weather stations NOANN (Network of the National Observatory of Athens). For Konya, the data have been provided by MGM (Turkish State Meteorological Service). For Grombalia, precipitation and temperature data have been provided by CRDA (the Regional Commissariat of Agricultural Development of Nabeul) and by INM (National Institute of Meteorology). For Castro Verde, the data were collected from SNIRH (the National Information System for Water Resources). Temperature data provided by WATCH Forcing Data (WFD) were used to fill gaps in some historical data sets (https://catalogue.ceh.ac.uk/documents/ba6e8ddd-22a9-457d-acf4-d63cd34f2dda#:~:text=The%20WATCH%20Forcing%20data%20is,half%2Ddegree%20land%20grid%20points).

- Existing models: previous numerical models of the MED area of three case studies are being used. These models are provided by the Consortium universities.

- Data from InTheMED project: for the training and validation of the surrogate models, the information gathered in WP2 is also used.

Data Utility

The outcomes achieved in this WP are highly beneficial to the ideal execution of the assignments involved in WP6.

Additionally, this information will be precious to scientists developing methods and models in earth sciences.

2.4. WP 4: Innovative Governance and Socio-Economic Assessment in the MED

Purpose

The WP4, led by the BU team, is developing a broad socio-economic assessment in the different MED case studies, in which the environmental and hydrological factors with anthrophonic demands and pressures will be integrated. Therefore, it is necessary to map the stakeholders and identify the key informants in each of the five selected case studies to collaboratively characterize socio-economic systems and identify both existing and future sustainability issues.

Type, Format and Expected Size of Data

- Interviews: Stakeholders, such as farmers, wineries, local industries, operators, and higher-level decision-makers, have been approached through their civil, public and private organizations. The BU team created a semi-structured interview template to get to know them and understand their position regarding groundwater management and groundwater quantity and quality state. The format of this document is .docx, and its size is approximately 30 KB. This template was originally written in Turkish, then translated to English and shared with the partners from Spain, UPV team, and Tunisia, CERTE team, as guiding material for the interviews with stakeholders involved in each study area. The format of these questionnaires is .docx, .pdf an .xlxs. The size of these documents, including both the questions and the answers, is approximately 1.3 MB for the Spanish case study, 2.5 MB for the Turkish case study and for the Tunisian case study, some questionnaires have been already filled in and others are in progress.

- Living Labs: The BU team has conducted two participatory interactive sessions (Living Labs) with the stakeholders of the Konya aquifer. Both Living Labs were held in Konya city

centre on the 30th of September, 2021, and on the 17th of February, 2022. The first one aimed to identify the problems with stakeholders to model conceptualization and model formulation phases of the participatory socio-economic modeling and to identify key variables of interest, their historical development over time as perceived by the participants and their prominent policy proposals intended for sustainable groundwater use in Konya Closed Basin. The second living lab focused on participatory model conceptualization aiming the represent the problems identified and test the policy proposals suggested during the first workshop. The Milestones M4.1 and M4.2 summarise the outputs of these workshops. The format of these documents is .pdf and the size is approximately 13 MB.

- Conceptual models: During the living labs in Konya, key informants, as part of a larger stakeholder group, arrive at a common understanding of the resilience problems of their basin and decide on their priorities. After discussion, the conceptual model of the sustainable resource exploitation and degradation problems for the Konya aquifer is co-created by the informants and BU team. This way, the model realism and the model's local ownership are enhanced. The output is saved as PDF documents (.pdf) and Microsoft PowerPoint presentations (.pptx).

- Socio-economic simulation model: Socio-economic simulation model implements systems dynamics principles to represent the dynamic nature of the groundwater social-ecological system. The format of the output files will be text files (.txt), Microsoft Excel file documents (.xlsx), or comma separated values files (.cvs) and Stella Architect files (.stmx and .isdb) and the size is approximately 10 MB.

- Numerical model outcomes: Different alternative scenarios and policy combinations will be implemented and the model outcomes will be discussed, stakeholders will join us in the debate. On the one hand, the format of the output files will be text files (.txt), Microsoft Excel documents (.xlsx), or comma separated values files (.cvs) and Stella Architect files (.stmx and .isdb) and the size of the documents is approximately 15 MB.

- Deliverables: In addition to the datasets mentioned above, table 5 summarises the deliverables within WP4. The lead author of each deliverable is responsible for writing the

document and delivering it to the rest of the Consortium as a text file (.docx). After reviewing, the final version is saved as a PDF document (.pdf).

Table 5. Deliverables of WP4

Deliverable Number and Title	Lead Participant	Type	Diss. Level	Delivery date	Format and size
D4.1 Report on the social-economic system characterisation, stakeholder mapping and water governance for selected case studies	BU	R	PU	2021/09/10	.pdf (15.7 MB)
D4.2 Report on the partici-patory systems mapping and the conceptual model	BU	R	PU	2021/03/31	.pdf (2.1 MB)
D4.3 Report on the numeric simulation model including model input files	BU	R	PU	Not delivered yet	-
D4.4 Report on simulation-based scenario analyses and policy design	BU	R	PU	Not delivered yet	-

Re-Use of Existing Data and its Origin

- Data from InTheMED project: the collected data and trend analysis done in WP2 and the smart models defined in WP3 are being considered in the execution of the present WP. As well as the SWOT (Strengths, Weaknesses, Opportunities and Threats) analysis done in the WP5, in order to identify the main groundwater issues, hotspots and key stakeholders.

Data Utility

The information provided in this WP will be helpful for researchers and professionals working in natural resources modelling since it will demonstrate a participatory method of identifying groundwater problems and incorporating an economic evaluation into numerical modelling. This data will also raise awareness in local citizens of current and future pressures and improve groundwater management in the MED region. Furthermore, the results obtained from this WP are necessary to accomplish the tasks of WP6.

2.5. WP 5: Innovative Remediation Strategies in the MED

Purpose

The fifth Work Package, led by the CERTE, will strongly engage stakeholders with the definition of sustainable water governance focusing on the pollution issues. The main objective is the approval of particular remediation strategies for the selected case studies through a participatory system.

Type, Format and Expected Size of Data

- Inventory of water threats: A dataset identifying water supplies, wastewater production, stream flows, rainfall, groundwater levels, pumping rates and inventorying the water threats and hotspots in each of the five case study sites, needing immediate intervention, has been collected. A specific SWOT analysis of remediation has been performed according to the main groundwater problems in each case study. These data were progressively collected from the concerned partners in a Google Drive document using track changes mode with a size of 20.5 MB.

- Remediation strategies: The CERTE team created a survey template to acquire preliminary information on the nitrate pollution in Requena-Utiel, Spain, on the nanoparticle's and nitrate pollution in Tympaki, Greece, and on the mining activities remediation strategies in Castro Verde, Portugal. The format of the surveys is .docx, and its size is respectively, 127 KB, 129 KB and 58 KB. This template was shared with the partners from Spain, UPV team, Greece, TUC team, and Portugal, IST-ID team. The size of these documents, including both the questions and the answers, is approximately 152 KB for the Spanish case study, 60 MB for the Greek case study and 584 KB for the Portuguese case study. The results of these surveys will be used to evaluate the current water management strategies (in-situ) and to define innovative remediation strategies in each case study by identifying several cost-benefit parameters. For example, in Tunisia, the CERTE Team is developing a cost-benefit model under R-studio software dedicated to polluting industries well developed in the area. This model will be useful to optimize the best sustainable remediation scenarios based on specific inputs (investment, energy consumption/m3 treated water, labour cost, etc ..) and outputs (COD removal, quality parameters residue

concentrations, reuse rate, etc). An excel file is generated according to the specificity of water activity in each case study with a file size of 10 kB.

- Deliverables: Besides the datasets mentioned above, table 6 collects the deliverables within WP5. The lead author of each deliverable is responsible for writing the document and delivering it to the rest of the Consortium as a text file (.docx). After reviewing, the final version is saved as a PDF document (.pdf).

Table 6. Deliverables of WP5

Deliverable Number and Title	Lead Participant	Type	Diss. Level	Delivery date	Format and size
D5.1 Report on site characterization and hot spot identification	CERTE	R	PU	2021/09/02	.pdf (7.4 MB)
D5.2 Report on the procedure for the capacity building and the selection of the main hot spots	CERTE	R	PU	Not delivered yet	-
D5.3a Report on the appropriate innovative remediation process developed	CERTE	R	PU	Not delivered yet	-
D5.3b Report on the results of the upscaling recommendations of the innovative remediation processes and reuse strategies	CERTE	R	PU	Not delivered yet	-

Re-Use of Existing Data and its Origin

- Data from InTheMED project: the data collection of WP2 and the socio-economic analysis of WP4 are being considered in the present WP.

- Publications and water authorities: the data for WP5 are being collected from existing relevant publications, non-scientific publications, water authorities' databases and field work campaigns.

Data Utility

This data is valuable for the achievement of the tasks of WP6. Likewise, it will be helpful to educate Water, Agricultural and Environmental authorities, stakeholders and citizens and raise awareness about the current state of aquifers in the MED area and the measures that must be taken in order to amend them both in the present and in the future.

2.6. WP 6: Innovative Decision Support Systems in the MED

Purpose

The WP6, led by the TUC team, will establish an innovative web-based Decision Support System (Fuzzy WebDSS) tool to provide an optimal, sustainable and easy-to-understand and visualize solutions/decisions for groundwater resources management to aquifer managers, users and even the public in general.

Type, Format and Expected Size of Data

- Multi-objective optimization algorithm: a multi-objective optimization algorithm will be developed to optimize pumping, discharge, inflows, outflows and water reuse, to maximize benefits and reliability and to minimize water availability risks and deviations from desired performance levels. Different scenarios will be tested and evaluated given different land use and the impact of climate change. Various files will be generated during this assignment, their format is expected to be Python files (.py), FEFLOW files (.fem) and MATLAB files (.mat), Microsoft Excel documents (.xlsx) or comma separated values files (.cvs), text files (.txt) and ESRI shapefiles (.shp).

- Geospatial data: The geospatial data utilized comprises of georeferenced vector files (Shapefiles - .shp) – approximately 10mb on average, but depends on the density and complexity of the data, along with the accompanying .sbn, .shx .dbf files that contain attribute data, as well as .prj files containing the projection mapping data for each data frame implemented. Finally, raster files (over 20Mb) are utilized for fine data spatial interpolation and performing map algebra.

- Database: An environmental data management system has been created in order to gather and mine the environmental data. The design of the database includes the storage of

the collected data in a coherent environmental database management system (DBMS) for further use within the project. A critical issue that was taken into is the existence of queries for automatic updating, importing data, exporting data, selecting specific subsets and grouping records. The database was connected with an open-source business intelligence tool, which helps users question the database and visualize answers in useful optical formats that can be shared. Structured data can be exported to CSV, XLSX and JSON file formats. Furthermore, there is a sharing lightweight data option, through the business intelligence tool, using public links. Viewers of public links are able to update the question's filters (if any) on the current dataset.

- Fuzzy WebDSS maps: thematic maps will be produced from the DSS according to the various studied scenarios, and will be visualized in maps accessible through the Fuzzy WebDSS page. Data from the High-Resolution Monitoring, the smart numerical modelling technique and the optimization algorithm will be used to "feed" the DSS tool. The file format is not yet known, it depends on the coupling of the algorithm with the web-page. This issue will be resolved during the design of the DSS tool. For now, the formats of the files are expected to be image files (.jpg, .png) and Print to text files (.prn).

- Deliverables: Table 7 summarises the deliverables that are created within WP6. The lead author of each deliverable is responsible for writing the document and delivering it to the rest of the Consortium as a text file (.docx). After their reviewing, the final version is saved as a PDF document (.pdf).

Table 7. Deliverables of WP6

Deliverable Number and Title	Lead Participant	Type	Diss. Level	Delivery date	Format and size
D6.1 Report on the development of the innovative DSS tool	TUC	R	PU	2021/02/28	.pdf (4.2 MB)
D6.2 Report on the results of the DSS for the case study sites	TUC	R	PU	Not delivered yet	-
D6.3 Atlas of the maps produced using the DSS	TUC	DEC	PU	Not delivered yet	-

Re-Use of Existing Data and its Origin

- Data from InTheMED project: the information produced in the previous WP (WP5) are being utilised for the implementation of the Fuzzy WebDSS. At the end, this tool will combine the monitoring data in WP2 with the smart numerical modelling results achieved in WP3. Furthermore, the DSS will be tested under future conditions forecasted in WP4 and for the remediation alternatives studied in WP5.

- Experience in water resources governance: previous national and international experiences and initiatives regarding water resources management are being studied. To perform this task, different sources of information are being checked, such as water authorities, universities, the stakeholders themselves and parties involved in the management and use of water.

- Geospatial information: Additionally, one of the tasks of the present WP is to establish a GIS platform in which all the necessary and available geospatial information, from regional/national and MED databases, will be transmitted.

Data Utility

The data collected in this WP will be advantageous for the scientific community since a novel multi-criteria optimization methodology within a Fuzzy logic WebDSS will be designed. Moreover, as a groundwater management tool will be developed, it will be highly valuable for the parties and actors involved, such as, Water, Agricultural and Environmental authorities, individual farmers, farmer associations, stakeholders, SMEs and the tourism industry. Additionally, in order to gather the necessary data, interactive engagement of the public is needed, hence, it may also interest the wider society.

2.7. WP 7: Innovative Dissemination and Communication in the MED

Purpose

The last WP, number seven and led by the IST-ID team, has the responsibility to develop novel communication strategies and dissemination materials to share the project results with the stakeholders and the larger society. Moreover, it aims to create cross-country dissemination tools based on participatory workshops and public communication that

empower the replicability of the methodology of the InTheMED project to the whole MED region.

Type, Format and Expected Size of Data

- Posters, leaflets, brochures, factsheets and videos: General information regarding the project was created to disseminate InTheMED to relevant stakeholders and related EU projects. One virtual poster was created in English and displayed at the GREEN NIGHT event (an experience of Tunisia in the European Researchers). The virtual poster was developed in English and in a PDF file of size 943 KB. Also, one general factsheet about the project was created in English and delivered as a PDF file of size 309 KB. Two brochures about the Portuguese and Spanish case studies are also available in English and in the corresponding language of the case study. These brochures are available in PDF format. The additional brochures for the remaining case studies are being finalized and will follow the same format. Also, one video with a recording from the project PI was created in mp4 format with a file size of 44,308 KB. All these data are publicly available at the project website (https://inthemedprima.com/).

- Newsletters: Two newsletters are planned at the mid-term and end of the project. These newsletters will be delivered as PDF files and Sway files to allow more interaction with the readers. These newsletters will be available at the project website (https://inthemedprima.com/).

- Publications: The TUC team published the following article in Water Resources Research in December, 2020: Katzourakis, V. E., & Chrysikopoulos, C. V. (2021). Modeling the transport of aggregating nanoparticles in porous media. Water Resources Research, 57, e2020WR027946. https://doi.org/10.1029/2020WR027946, and the following article in Water in May, 2021: Stefanarou, A.S.; Chrysikopoulos, C.V. Interaction of Titanium Dioxide with Formaldehyde in the Presence of Quartz Sand under Static and Dynamic Conditions. Water 2021, 13, 1420. https://doi.org/10.3390/w13101420. The UNIPR team published the following article in Journal of Hydrology in December 2021: Secci, D., Tanda, M.G., D'Oria, M., Todaro, V., Fagandini, C., 2021. Impacts of climate change on groundwater droughts by means of standardized indices and regional climate models. J. Hydrol. 603, 127154. https://doi.org/10.1016/J.JHYDROL.2021.127154.

- Deliverables: Table 8 sums up the deliverables produced within WP7. All of them are reports, except for the project website, https://inthemedprima.com/, which includes the project results, increasing the number of potential users interested in our results. The lead author of each deliverable is responsible for writing the document and delivering it to the rest of the Consortium as a text file (.docx). After reviewing, the final version is saved as a PDF document (.pdf).

Table 8. Deliverables of WP7

Deliverable Number and Title	Lead Participant	Type	Diss. Level	Delivery date	Format and size
D7.1a Project website	IST-ID	DEC	PU	2020/08/03	-
D7.1b Communication and Dissemination Plan	IST-ID	R	PU	2021/04/26	.pdf (1.2 MB)
D7.2 Communication and dissemination activities	IST-ID	R	PU	Not delivered yet	-
D7.3 Report on synergies with groundwater ini-tiatives in the Euro-MED region	UFZ	R	PU	Not delivered yet	-
D7.4 Report on mid-term workshop	IST-ID	R	PU	2021/03/31	.pdf (1.0 MB)
D.7.5 Report on the InTheMED final scientific conference	IST-ID	R	PU	Not delivered yet	-
D7.6 Exploitation plan	IST-ID	R	PU	Not delivered yet	-

Re-Use of Existing Data and its Origin

- Data from InTheMED project: all the research data gathered and produced by the previous WPs are being considered and are uploaded at the website as the contents are produced.

- Groundwater management approaches in the MED area: a review about the state-of-the-art innovative aquifer management approaches was made to explore opportunities to

improve the models and DSS developed in the InTheMED project. Therefore, the Consortium will build a network of contacts with groundwater management authorities and major groundwater-oriented projects in the MED region, such as, Horizon2020, WaterJPI, LIFE+, Interreg, ENI-CBC-Med, COSME programme, SWIM and the UN. The data collection obtained through this activity will be collected in the deliverable number D.7.3 of this WP.

Data Utility

In this WP, the four main target groups of the InTheMED project are identified: scientific community developing models; Water, Agricultural and Environmental Authorities, who are responsible for planning and implementation of regional strategies; stakeholders such as SMEs, farmers and associations socio-economically involved and, finally, citizens, who will benefit from objective, trustworthy and comprehensible information.

3. FAIR Data

In line with the spirit of the European Commission, the data management plan has to follow the FAIR principles, which stands for findable, accessible, interoperable and reusable research data. The present chapter outlines how to achieve this objective by answering the set of questions that are found in the "Guidelines on FAIR Data Management in Horizon 2020" [1], the open data requirements [2] and the Open Access implementation guidelines [3].

3.1. Making Data Findable, including Provisions for Metadata

Standard Identification Mechanism

The open data produced and used in the InTheMED project, such as datasets, deliverables, publications and software, are identifiable and locatable by means of a persistent and unique identifier.

These open InTheMED results are being deposited in Zenodo [4], an Open Access repository. Zenodo is a generalist multidisciplinary repository that is developed under the European OpenAIRE program and is operated by Centre Européen pour la Recherche Nucléaire CERN, in Geneva, Switzerland. It automatically assigns a Digital Object Identifier (DOI) to the uploaded data. It has the option of DOI versioning, which enables the editing and updating of a record file that has already been published.

Naming Conventions and Version Numbers

File names are being specified throughout the execution of the project and they include a version number as follows:

- For the deliverables: Project acronym, time stamp (YYYYMMDD), deliverable code, title of the deliverable, and version number. E.g., "InTheMED_20201018_ D1.4_DataManagementPlan_v1".

When a deliverable is composed of more than one file it is identified with a letter for each part e.g., "InTheMED_D1.4a_DataManagementPlan_v1". When big changes are made, version number changes. E.g., from "InTheMED_D1.4_DataManagementPlan_v1" to

"InTheMED_D1.4_DataManagementPlan_v2". If the changes are essentially small corrections or additions, version number is change as follows: from Version 1.1 to Version 1.2 etc.

- For the dataset files: Project acronym, WP number and "DS", "INT", which stands for dataset and interviews/questionnaires, respectively, followed by a short description of its content. E.g., "InTheMED_WP2_DSGroundwaterTimeSeries".

- For models: Project acronym, WP number, study case name, and "SM", "NM", which stands for smart model and numerical model, respectively. E.g., "InTheMED_WP1_Requena_NM".

- Source software: Semantic Versioning Schema will be followed to number its release.

Search Keywords

Keywords are being added to optimise the possibilities of finding the outcomes and deliverables of the project. They are being defined following the terminology used in the various scientific fields addressed in the project and they are being descriptive to the content of each InTheMED result. It will be updated up to the end of the project, however, the following keywords were already been selected:

- Smart model
- Metamodel
- Surrogate model
- Groundwater modelling
- Geostatistical DC inversion
- Requena-Utiel aquifer
- Grombalia shallow aquifer
- Konya Closed Basin
- High resolution monitoring

- Groundwater level trends
- Groundwater level monitoring
- Groundwater management
- DSS – Decision Support System
- Groundwater overexploitation
- System dynamics model
- Group model building
- Living Lab
- Sustainability
- Mediterranean
- Climate change

Metadata and Metadata Standards

In order to guarantee that data is easily locatable and EU funding is acknowledged, bibliographic metadata are being provided for each open InTheMED result. As described in the Article 29.2 of the project Grant Agreement (GA), the "bibliographic metadata must be in a standard format and must include all of the following:

- the terms "PRIMA", "European Union (EU) and "Horizon 2020",
- the name of the action, acronym and grant number,
- the publication date, and length of embargo period if applicable, and
- a persistent identifier."

Furthermore, according to the Article 27.3 of the GA, "applications for protection of results (including patent applications) filed by or on behalf of a beneficiary must include the following: "The project leading to this application is part of the PRIMA programme supported

by the European Union"". The Article 28.2 of the GA declares that "if results are incorporated in a standard, the beneficiary concerned must ask the standardisation body to include the following statement in (information related to) the standard: "Results incorporated in this standard received funding from the PRIMA programme supported by the European Union"". The Article 29.4 of the GA states that "any dissemination of results must:

- display the PRIMA logo,
- display the EU emblem, and
- include the following text:

  "This project is part of the PRIMA programme supported by the European Union's Horizon 2020 research innovation programme".

Finally, as declared in the Article 29.5 of the GA, "any dissemination of results must indicate that it reflects only the author's view and that the PRIMA foundation is nor responsible for any use that may be made of the information it contains", therefore, the following must be included in publications: "Disclaimer: The content of this publication is solely responsibility of the authors and it does not represent the view of the PRIMA foundation".

The bibliographic metadata are following the Dublin CoreTM Metadata Schema standard [5]. Moreover, open InTheMED geospatial results will be accompanied by metadata according to the INSPIRE guidelines [6] (based on EN ISO 19115 and EN ISO 19119). Following the Dublin CoreTM Metadata Element set [7], these fifteen elements should at least be created:

- Contributor: name of the funding entity.
- Coverage: spatial or temporal topic of the data.
- Creator: the person responsible for making the data.
- Date: the year when the data is made publicly available.
- Description: a brief summary of the data.
- Format: file format, physical medium and dimensions of the data.

- Identifier: a unique string that identifies the data.
- Language: language of the data.
- Publisher: name of the entity that makes to data openly available.
- Relation: a related resource.
- Rights: property rights associated with the data, including intellectual property rights.
- Source: a related resource from which the data is derived.
- Subject: the topic of the data.
- Title: name of the data.
- Type: nature or genre of the data.

3.2. Making Data Openly Accessible

Dissemination Level and How will the Data be Made Accessible?

All the data generated and used during the InTheMED project, such as datasets, deliverables, publications and software, are open by default except for confidential documents, which will not be available for the public.

The confidential documents are Deliverables 1.1, 1.3 and 3.3, and the research data associated with those documents, project meeting minutes and interviews/questionnaires made to stakeholders. The confidential documents are stored in private serves of the Consortium members and in the InTheMED community in Zenodo under restricted access, only available for the project participants. Moreover, internal documentation is posted and shared through the back-office system in the InTheMED project's website and only the Consortium will have access to it.

IST-ID team developed the InTheMED website, https://inthemedprima.com/, where information about the project, team and partners, highly relevant project results (datasets, deliverables and publications) are linked to the repository, where they are stored and from where they are freely available for download. Moreover, it will exhibit web-based

visualizations of the developed methods and models. Additionally, we employ ICT tools such as social media to communicate InTheMED results to reach a larger audience and raise awareness among the wider society. The InTheMED social networks are:

- Twitter: @InTheMED_PRIMA
- LinkedIn: InTheMED PRIMA
- Facebook: Inthemed Prima

Open InTheMED results are deposited in open access and public domain repository, Zenodo, a European Commission-funded site located at CERN [4]. It is a general-purpose open access repository that is technically compliant with the open data requirements of the OpenAIRE portal and Horizon 2020 to enable the harvesting of metadata. Hence, the process of reporting the InTheMED results in OpenAIRE is automatic. The InTheMED community in Zenodo, see Figure 1, can be found following this link: https://zenodo.org/communities/inthemed/.

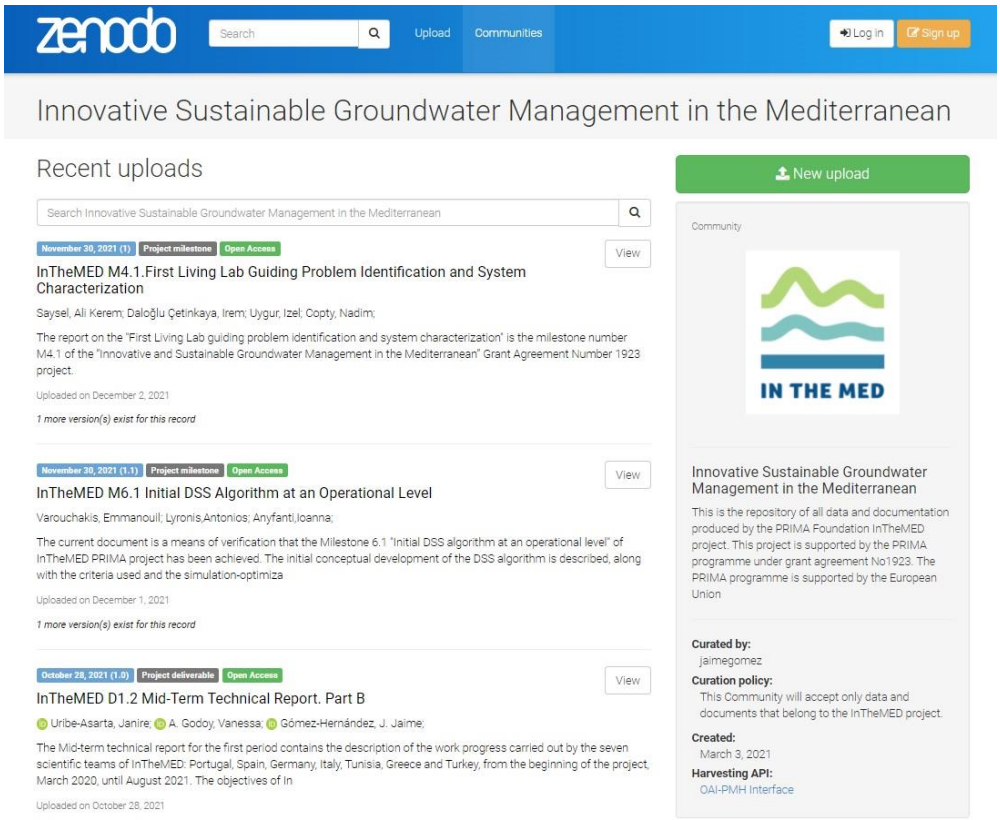

Figure 1. InTheMED community in Zenodo.

Open InTheMED time series, which are the datasets attained from the WP2, will join the current of the industrial internet of things (IIOT). Considering the European policy for open public data, we plan to make both raw and processed data public for Spain, Portugal, Greece and Italy. In the case of Tunisia and Turkey, as they have different data sharing policies, only the processed data for anomalies would be made public. Currently the best possible arrangement for this is still being processed. The time series will be connected to major portals sharing groundwater data, such as IGRAC (GGMN) and GEMStat.

In relation to journal and peer-reviewed publications, InTheMED will attempt to publish in Gold Open Access (open access publishing) journals, so the article is immediately published in open access mode, sometimes requiring the payment of a fee by the author. Nonetheless, InTheMED will presumably publish in Green Open Access (self-archiving) journals. Depending on the publication conditions of the journal, the author's version or the final peer-reviewed manuscript will be published directly in Zenodo or published in the ArXiv repository. In the latter case, the embargoed copy will be deposited in Zenodo, with embargoed access and indicating the link to ArXiv where the author's version will be freely available. All InTheMED publications should be promptly included in Zenodo after their acceptance.

The software developed by InTheMED will be openly stored in the GitHub portal [8] along with its deposit in Zenodo.

There are no provisions made in the Consortium Agreement concerning any project beneficiary to keep their data closed. At the moment, all project beneficiaries follow the guidelines of the open data pilot.

If IPR issues arise from a product patent or peer-reviewed publications and the open access to research data is restricted or embargoed, information about this data will be made publicly available through the repository and the project website by “README” text files. The details of when the data will become available will be found in these. Additionally, these “README” text files will be included along with datasets, where the content of the resource will be detailed.

Software Tools to Read or Reuse Data

No special software tools are needed to access data deposited as open data files in Zenodo or on the project website. These files will be downloadable from the repository via HTTP protocol using a standard web browser.

Regarding software and tools for the reuse of openly accessible data, once it is downloaded, it depends on the type and format of such data. Reports and most datasets are produced in a standard and general electronic format that does not require specific software to read or reuse them. However, some datasets are created with MATLAB platform. InTheMED code is written in Python and MATLAB languages, and we also use Stelle Architect software. We will prioritize the use of free and open-source software, such as Python, although it will not always be feasible. On the other hand, despite MATLAB being proprietary software, it is a widely used language in the scientific and academic community. Documentation of open-source software required to access the data developed by InTheMED will be made available on the project website and the repository.

Access Requirements

Open InTheMED results hosted in the InTheMED community in Zenodo and the project website will be freely available without restriction to the public.

Confidential InTheMED results will only be available for the project participants.

3.3. Making Data Interoperable

Allowing Data Exchange and Re-Use

One of the greatest fulfilments that InTheMED Project would like to achieve is to provide novel research data, methodologies and tools for a sustainable groundwater preservation and governance in the MED region. And, furthermore, ensure that those outcomes are replicable and transferable to other MED countries. Therefore, the InTheMED results will be free and openly available using standard file formats and open software applications will be used whenever is possible and the conversion of proprietary data files into standard formats will be encouraged.

3.4. Increase Data Re-Use

Data Licensing and Date of Data Release

As Article 26 of the GA states, “results are owned by the beneficiary that generates them” and if two or more beneficiaries have jointly contributed to creating an InTheMED result, both will be considered as shared owners of that product. All the InTheMED results will be shared within the Consortium and once the Consortium decides to make the research data public, it will be promptly published through the means mentioned earlier.

In order to protect the ownership, InTheMED datasets and publications will be released under a Creative Commons License, presumably Creative Commons Attribution-NonCommercial-ShareALike 4.0 (CC BY-NC-SA 4.0) [9] will be used. Copyleft licenses will be considered for software licensing, such as GNU GPL [10].

Assumedly, scientific publications will be published by Green Open Access journals, which permits us to retain the ownership of the result and deposit the author’s version or the final peer-reviewed manuscript in the InTheMED community in Zenodo. Moreover, they will be made available on the project website as well. They will be released promptly after publication or within six months of publication if the publisher imposes an embargo period. The same measures will be taken concerning research data and related metadata, as described in the second section of this chapter, “Making data findable”.

Data Useable by Third Parties

Open InTheMED results, including research data needed to validate the results in scientific publications and the scientific publications, deposited in Zenodo and the project website are usable by third parties during and after the end of the project.

Length of Time for Data Re-Usage

In compliance with the GA, Article 31.3 “Requests for access may be made – unless agreed otherwise – up to one year after the period set out in Article 3”. Thus, the research data will remain reusable for at least one year after the end of the project, being 28 August 2023 the estimated project termination date at this moment.

Furthermore, Zenodo commits to retaining the availability of published data for the lifetime of the repository, of at least 20 years.

Finally, InTheMED results will be available for re-use on the project website for at least five years, two more than the end of the project.

Data Quality Assurance Processes

InTheMED results are being corrected and validated by the PI of the partner team involved in their production. If two or more partner teams jointly create InTheMED results, all of them must participate in the correction and validation. Then, the Coordinator perform a quality control concerning the accuracy in following the FAIR conditions established in the DMP. After executing this data quality assurance process, the results are made publicly available on Zenodo and the project website.

4. Allocation of Resources

Costs for Making Data FAIR

There are no costs anticipated to make open InTheMED results FAIR. Data archiving at Zenodo is completely free of charge. There is no cost for long-term preservation of the InTheMED results and metadata in Zenodo. Likewise, licensing with Creative Commons is free as well.

Moreover, the maintenance of the project website, the possible cost of publishing scientific articles and any other costs related to providing open access to research data are eligible for reimbursement during the duration of the project if they fulfil the general conditions specified in Article 6 of the GA.

5. Responsibilities Towards Data Management

Sound data management demands the cooperation and coordination of all team partners at project and work package levels. The Project Coordinator, UPV, is concerned about the overall data management at project level. In contrast, at WP level, each team leader is responsible for their produced data. Each partner is handling their datasets; they backing up and depositing their data in the repository.

This chapter identified the responsibilities towards data management.

5.1. Project Coordinator

- Create the DMP in collaboration with the rest of the project partners,
- Create a user's guide on how to curate the InTheMED results,
- Monitor data management activities and remind deadlines to the rest of the partners,
- Create a template for project documents, deliverables, milestones and any other type of report, to keep a coherent and representative format for the InTheMED project and to maintain a general uniformity and good image,
- Review InTheMED results and perform a quality control to assure that the FAIR conditions stated in the DMP are abided before their author uploads them to Zenodo,
- Ensure that all partners comply with the provisions of the DMP.

5.2. Partner Team Leader

The PI of each WP must:

- Implement the DMP in their respective WP and results,
- Monitor data management activities regarding their WP and remind deadline to their colleagues,
- Use the template for all project documents,
- Review their InTheMED results before sending them to the Project Coordinator, WP1,

- Check the conditions for publishing open data, open-source software and study the Green Open Access publication conditions to upload the author's version or the final peer-reviewed manuscript to Zenodo or ArXiv,
- Upload their InTheMED results to Zenodo after receiving approval from the Project Coordinator.

5.3. Dissemination and Communication Leader

The PI of WP7 is the Dissemination and Communication leader and they must:

- Guarantee that the open access policy of the selected journals complies with the H2020 open data requirements,
- Monitor that the green access publications are deposited in the InTheMED community in Zenodo,
- Incorporate all the InTheMED results available in Zenodo to the project website,
- Upload their InTheMED results to Zenodo after receiving approval from the Project Coordinator.

6. Data Security

Provisions for Data Security

Three backups are being made to minimize the risk of data loss; two storage media and one off-site are recommended. The data are being saved on the Consortium's private servers and copied onto a second external location, such as local or portable devices, and finally, cloud storage are being used. For cloud storage solutions, the use of the services of each partner's institutions is encouraged whenever possible.

Long-Term Safe Preservation and Curation in Certified Repositories

The InTheMED results shared in Zenodo are safely preserved and curated indefinitely, as described in the third chapter, "FAIR Data".

Regarding the project website maintenance, the domain was acquired for five years, two more than the end of the project. Additionally, an effort will be made to keep the website running after this period.

7. Ethical Aspect

The interviews and questionnaires made to stakeholders are confidential; hence, they will not be publicly available. They are stored in the InTheMED community in Zenodo under restricted access and are only available for the project participants.

InTheMED project does not involve personal data collection or processing. Therefore, no other ethical issue has been found with regard to data management.

8. References

[1] European Commission, Directorate-General for Research & Innovation. (26 July 2016). *Guidelines on FAIR Data Management in Horizon 2020. Version 3.0.* Retrieved from http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

[2] European Commission, Directorate-General for Research & Innovation. (21 March 2017). *Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020. Version 3.2.* Retrieved from https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf

[3] European Research Council. (07 April 2017). *Guidelines on the Implementation of Open Access to Scientific Publications and Research Data in projects supported by the European Research Council under Horizon 2020. Version 1.1.* Retrieved from https://erc.europa.eu/sites/default/files/document/file/ERC%20Open%20Access%20guidelines-Version%201.1._10.04.2017.pdf

[4] European Organization For Nuclear Research, & OpenAIRE. (2013). *Zenodo*. Retrieved from https://www.zenodo.org

[5] Dublin CoreTM Metadata Initiative (November 2020). Dublin CoreTM Metadata Initiative. Retrieved from https://www.dublincore.org/

[6] European Commission (2007). INSPIRE. Retrieved from https://inspire.ec.europa.eu/

[7] Dublin CoreTM Metadata Initiative (November 2020). DCMI Metadata Terms. Retrieved from https://www.dublincore.org/specifications/dublin-core/dcmi-terms/

[8] github. GitHub. Retrieved from https://github.com/

[9] Creative Commons. Creative Commons Attribution-NonCommercial-ShareALike 4.0 (CC BY-NC-SA 4.0). Retrieved from https://creativecommons.org/licenses/by-nc-sa/4.0/

[10] Free Software Foundation. GNU General Public License. Retrieved from https://www.gnu.org/licenses/gpl-3.0.en.html