



Innovative and Sustainable Groundwater Management in the Mediterranean

D.1. 4 DATA MANAGEMENT PLAN

VERSION 1.0



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.

Disclaimer: The content of this publication is solely responsibility of the authors and it does not represent the view of the PRIMA Foundation.

DOI: [10.5281/zenodo.4577339](https://doi.org/10.5281/zenodo.4577339)

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	<p>Universitat Politècnica de València (UPV), Spain</p> <p>Helmholtz-Zentrum für Umweltforschung (UFZ), Germany</p> <p>Università degli Studi di Parma (UNIPR), Italy</p> <p>Boğaziçi Üniversitesi (BU), Turkey</p> <p>Centre de Recherches et des Technologies des Eaux (CERTe), Tunisie</p> <p>Technical University of Crete (TUC), Greece</p> <p>Associação do Instituto Superior Técnico para a Investigação e Desenvolvimento (IST-ID), Portugal</p>		

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	September 1, 2020	Actual March 3, 2021
Version Number	1.0		
Deliverable Type	report (R)	Dissemination Level	public (PU)
Authors	Janire Uribe Asarta J. Jaime Gómez-Hernández		
Reviewer(s)	All		

Document History

Version	Date	Stage	Reviewed by
0.1	2020/10/20	First draft	All
0.2	2020/11/20	Contributions from all partners have been added	Janire Uribe Asarta J. Jaime Gómez-Hernández
0.3	2020/12/07	Second draft	All
0.4	2020/12/18	Corrections and comments from all partners have been added	Janire Uribe Asarta J. Jaime Gómez-Hernández
1.0	2021/03/03	Final version	Janire Uribe Asarta J. Jaime Gómez-Hernández

Table of Contents

Project Information	2
Document Information	3
Document History	3
Glossary	6
Executive Summary	7
1. Introduction	8
2. Data Summary	9
2.1. WP 1: Innovative Project Management.....	9
2.2. WP 2: Innovative Monitoring and Data Analysis in the MED	10
2.3. WP 3: Innovative Smart Modelling in the MED.....	13
2.4. WP 4: Innovative Governance and Socio-Economic Assessment in the MED.....	15
2.5. WP 5: Innovative Remediation Strategies in the MED	17
2.6. WP 6: Innovative Decision Support in the MED	19
2.7. WP 7: Innovative Dissemination and Communication in the MED.....	21
3. FAIR data	24
3.1. Making Data Findable, including Provisions for Metadata	24
3.2. Making Data Openly Accessible	27
3.3. Making Data Interoperable	29
3.4. Increase Data Re-Use.....	30
4. Allocation of Resources	32
5. Data Security.....	33
6. Ethical Aspects.....	34
7. References	35

List of Tables

Table 1. Deliverables of WP1.....	10
Table 2. Deliverables of WP2.....	11
Table 3. Deliverables of WP3.....	14
Table 4. Deliverables of WP4.....	16
Table 5. Deliverables of WP5.....	18
Table 6. Deliverables of WP6.....	20
Table 7. Deliverables of WP7.....	22

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.
EEA	European Environment Agency.
GA	Grant Agreement
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.
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 aims to consolidate the data management strategy of the InTheMED project. The Data Management Plan describe the data that will be collected and produced within the project and provide a guide on how this data will be manage, storage, disseminated and preserved. This is considered to be a living document that is expected to evolve during the project. This current deliverable presents the DMP at month 6 of the three-year project.

1. Introduction

The Data Management Plan is the deliverable number D1.4 of the “Innovative and Sustainable Groundwater Management in the Mediterranean” Grant Agreement Number 1923 project.

This document seeks to provide guidance on how to manage the data that will be collected and produced within the InTheMED project. 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 main use will be as a support tool for the InTheMED Consortium partners. 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 Data Management Plan is considered to be a living document. This document is the first version of it, due at six months after kick-off, but delayed in its submission due to the SARS-CoV-2 pandemics, and there will be, at least, two more deliverable versions, at mid-term and at project end. At this early stage in the evolution of the project, there are several questions regarding the data that cannot be fully answered because they are still being debated, however, they will be detailed in next versions of this document. Moreover, the DMP will be accurately modified and updated when major changes occur such as production of new data, changes in consortium policies or changes in consortium composition and external factors, as mentioned in the “Guidelines on FAIR Data Management in Horizon 2020” [1].

2. Data Summary

This chapter of the DMP has the objective of presenting the numerous data that is going to be created and collected throughout the execution of the InTheMED project. The data collection of each Work Package is separately displayed.

Since it is the beginning of the project, few data have been collected and produced so far. Consequently, the size of the majority of the files is yet to be defined and it is not going to be determined in the present first version of the DMP.

2.1. WP 1: Innovative Project Management

Purpose

The first Work Package, of which the lead participant is the UPV, 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 will be assured by communication by telephone, email, Skype and by cloud-based project management tools such as Slack or Trello. There will be three meetings, a kick-off, a mid-term and a final meeting. The last one will serve also as a showcase of the project outputs to all stakeholders. Furthermore, annual coordination meetings will be held where advisory members will participate. The initial plan was that the meeting venues will rotate among the partner countries but given the SARS-CoV-2 induced situation, they are shifting to virtual 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 will be documented in the minutes of the meetings. The minutes will be saved as a word document (.docx) or as a PDF document (.pdf).
- Presentations: during the meetings, presentations will be shown in Microsoft PowerPoint format (.pptx).
- Deliverables: from this WP, four deliverables will be created. The lead participant of each deliverable is the responsible to write the document and deliver it to the rest of the

Consortium as a word document (.docx). After their reviewing, the final version will be saved as a PDF document (.pdf).

Table 1. Deliverables of WP1

Deliverable number	Deliverable title	Short name of lead participant	Type	Diss. level
D1.1	Consortium Agreement	UPV	R	CO
D1.2	First, second and third periodic reports	UPV	R	PU
D1.3	Documentation of kick-off, annual coordination and WP leader meetings	UPV	R	CO
D1.4	Data management plan	UPV	R	PU

Re-Use of Existing Data and its Origin

Re-use of existing data is not expected.

Data Utility

These data will be primarily useful for the InTheMED Consortium partners. In the case of the DMP, it will also be useful 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 Work Package, led by the UFZ, 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, our hypothesis is that understanding past groundwater history can help in the comprehension of aquifers current state and in the prediction of future evolution (WP5).

Type, Format and Expected Size of Data

- Groundwater quality and quantity data: optical high-frequency sensor, grab sampling and HRMA will be used to collect and generate groundwater quality and quantity time series for each case study, as solute concentrations and hydraulic heads, respectively. We will work with stakeholders to identify regional, national and global SMARTs indicators (Specific, Measurable, Achievable, Relevant and Time-bound). The data-driven analyses will be implemented using different empirical machine learning methods such as Artificial Neural networks (ANN) or other methods based on deep learning. The formats of the files are expected to be Microsoft Excel documents (.xlsx) or comma separated values files (.csv), text files (.txt), python files (.py) and MATLAB files (.mat).
- Geophysical data: the spatial distribution of the aquifer properties, both the solid and fluid phases, will be collected and generated. The formats of the files are expected to be Microsoft Excel documents (.xlsx) or comma separated values files (.csv), text files (.txt), ESRI shapefiles (.shp), print to text files (.prn), MATLAB files (.mat) and ASCII grid files (.asc).
- Publications: peer-reviewed publications are expected to be prepared for submission to relevant high-impact international journals. The final version will be saved as a PDF document (.pdf).
- Deliverables: in addition to the datasets mentioned above, the following deliverables will be created. The lead participant of each deliverable is the responsible to write the document and deliver it to the rest of the Consortium as a word document (.docx). After their reviewing, the final version will be saved as a PDF document (.pdf).

Table 2. Deliverables of WP2

Deliverable number	Deliverable title	Short name of lead participant	Type	Diss. level
D2.1	Report on the integrated and innovative high-resolution monitoring strategies in the different case studies	IST-ID	Other	PU
D2.2	Report on the existing historical groundwater data on the MED region	UFZ	R	PU
D2.3	Report on regional groundwater trends and their controlling factors	UFZ	R	PU

Table 2. Deliverables of WP2

Deliverable number	Deliverable title	Short name of lead participant	Type	Diss. level
D2.4	Reinforcement of the systematic monitoring and data sharing	UFZ	DEM	PU

Re-Use of Existing Data and its Origin

- Data from InTheMED project: WP2 and WP5 will work jointly in 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 will be 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 will be compiled and saved in presumably Microsoft Excel documents (.xlsx) or comma separated values files (.cvs), text files (.txt), comma separated values (.csv) and ESRI shapefiles (.shp). The Consortium partners' sites, as well as national, regional and global water quality databases will be studied 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: Sistema Nacional de Informação 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/WRRRL 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/ggm-global-groundwater-monitoring-network>).

Data Utility

The results obtained in the present WP2 will be essential for the correct progress of the tasks involved in the rest of the work packages, it mainly has a close connection with WPP3, WP5 and WP6. Consequently, all the members of the Consortium will be handed this information. It will also be 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 relates 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 Work Package, led by the UNIPR, will create simplified models, one for each case study, which are called by the name of surrogate models, meta-models or smart models. These smart models will be built on the basis of long-time historical data (WP2), 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 will be considered.

Type, Format and Expected Size of Data

- Procedures and computer codes: procedures and computer codes will be developed for the application of surrogate models to the five case studies. The code will be saved as Python documents (.py), MATLAB files (.mat), Microsoft Excel documents (.xlsx) or comma separated values files (.csv) and text files (.txt).
- Downscaling of future climate projections: for the case studies, a statistical downscaling and a bias correction of the model outputs provided by EURO-CORDEX and MED-CORDEX will be performed considering that enough climate historical data are available. In case of insufficient data, an assessment will be made by comparison of future projections with the

outputs computed for the historical period. Data archive will be saved as MATLAB structures (.mat) and csv files (.csv).

- Publications: peer-reviewed publications are expected to be prepared for submission to relevant high-impact international journals. The final version will be saved as a PDF document (.pdf).
- 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.
- Deliverables: from this WP, four deliverables will be developed in report format. The lead participant of each deliverable is the responsible to write the document and deliver it to the rest of the InTheMED Consortium as a word document (.docx). After their reviewing, the final version will be saved as a PDF document (.pdf).

Table 3. Deliverables of WP3

Deliverable number	Deliverable title	Short name of lead participant	Type	Diss. level
D3.1	Identification of the surrogate models to be applied in the case studies	UPV	R	PU
D3.2	Report on surrogate models in the case studies	UNIPR	R	PU
D3.3	Data Archive containing the downscaled climate projections in the case studies	UNIPR	R	CO
D3.4	Report on the results of the analysis of different scenarios in the case studies	UNIPR	R	PU

Re-Use of Existing Data and its Origin

- Future projections of climate variables and climate historical data: EURO-CORDEX and MED-CORDEX provide Regional Climate Models (RCM), at a scale of about 12,5 km, downscaled from General Circulation Models (GCM), at a scale of hundreds of kms. The scenarios named Representative Concentration Pathways (RCPs), adopted by the ICPP for the Fifth Assessment Report (AR5), in particular, the RCP4.5 and the RCP8.5, will be considered.

- Existing models: previous numerical models of the MED area five case studies are being examined for its use in this project. This information will be obtained from the Consortium universities.
- Data from InTheMED project: for the training and validation of the surrogate models, the information gathered in WP2 will be used.

Data Utility

The outcomes achieved in this Work Package will be highly beneficial to the ideal execution of the assignments involved in the WP6.

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

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

Purpose

The fourth Work Package, led by the BU, will develop a broad socio-economic assessment in the different MED case studies, in which the environmental and hydrological factors with anthropogenic demands and pressures will be integrated. Therefore, it is necessary to map the stakeholders and identify the key informants in each 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: one-to-one contact and interviews with the stakeholders will be conducted. Moreover, citizens, engineers, operators and higher-level decision makers will be approached through their civil, public and private organizations. These are expected to be PDF documents (.pdf).
- Conceptual models: participatory interactive sessions (Living Labs) will be organized with the key informants with the objective of creating conceptual models of the sustainable resource exploitation and degradation problems for each of the five case studies. PDF documents (.pdf) and Microsoft PowerPoint presentations (.pptx) files will be created.

- Numerical models: a simulation model will be built for the Konya aquifer, from the case study of Turkey. The format the input files that will be generated are expected to be Microsoft Excel documents (.xlsx) or comma separated values files (.csv), text files (.txt), ESRI shapefiles (.shp), Stella Architect files (.stmx and .isdb), geodatabases (.gdb), MATLAB files (.mat) and print to text files (.prn).
- 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 as well. On the one hand, the format of the output files will be, text files (.txt), Microsoft Excel documents (.xlsx) or comma separated values files (.csv) and Stella Architect files (.stmx and .isdb). And on the other hand, for and during the Living Lab these data files will be generated PDF documents (.pdf), Microsoft Excel documents (.xlsx) or comma separated values files (.csv) and Microsoft PowerPoint presentations (.pptx).
- Video tutorials: how-to-use video tutorials will be produced to make understanding the dynamics of social-economical groundwater model less arduous. These files are expected to be in AVI (.avi) or MP4 (.mp4) format.
- Publications: peer-reviewed publications are expected to be prepared for submission to relevant high-impact international journals. The final version will be saved as a PDF document (.pdf).
- Deliverables: in addition to the datasets mentioned above, the following deliverables will be created. The lead participant of each deliverable is the responsible to write the document and deliver it to the rest of the Consortium as a word document (.docx). After their reviewing, the final version will be saved as a PDF document (.pdf).

Table 4. Deliverables of WP4

Deliverable number	Deliverable title	Short name of lead participant	Type	Diss. Level
D4.1	Report on the social-economic system characterisation, stakeholder mapping and water governance for selected case studies	BU	R	PU
D4.2	Report on the participatory systems mapping and the conceptual model	BU	R	PU
D4.3	Report on the numeric simulation model including model input files	BU	R	PU

Table 4. Deliverables of WP4

Deliverable number	Deliverable title	Short name of lead participant	Type	Diss. Level
D4.4	Report on simulation-based scenario analyses and policy design	BU	R	PU

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 will be 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 useful for researchers and professionals working in natural resources modelling since it will demonstrate a participatory method of how to identify groundwater problems and how to incorporate an economic evaluation into numerical modelling. This data will also raise awareness in local citizens to current and future pressures and to improve groundwater management in the MED region. Furthermore, the results obtained from this WP will be necessary for the accomplishment of the tasks of the 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. A specific SWOT analysis of remediation will be carried out according to the main groundwater problems in each case study. The expected file formats are Microsoft Excel

documents (.xlsx) or comma separated values files (.csv), text files (.txt), Microsoft Word documents (.docx) and PDF document (.pdf).

- Meeting minutes: Living Labs and workshops will be organized in which case-study partners and stakeholders will participate to discuss, taking advantage of the knowledge and experience of the latest, sustainable integrated water management. The decisions and comments discussed during these meetings will be documented in the minutes of the meetings. The minutes will be saved as a word document (.docx) or as a PDF document (.pdf).
- Publications: peer-reviewed publications are expected to be prepared for submission to relevant high-impact international journals. The final version will be saved as a PDF document (.pdf).
- Deliverables: from this WP, the following deliverables will be created. The lead participant of each deliverable is the responsible to write the document and deliver it to the rest of the Consortium as a word document (.docx). After their reviewing, the final version will be saved as a PDF document (.pdf).

Table 5. Deliverables of WP5

Deliverable number	Deliverable title	Short name of lead participant	Type	Diss. level
D5.1	Report on site characterization and hot spot identification	CERTE	R	PU
D5.2	Report on the procedure for the capacity building and the selection of the main hot spots	CERTE	R	PU
D5.3a	Report on the appropriate innovative remediation process developed	CERTE	R	PU
D5.3b	Report on the results of the upscaling recommendations of the innovative remediation processes and reuse strategies	CERTE	R	PU

Re-Use of Existing Data and its Origin

- Data from InTheMED project: the data collection of WP2 and the socio-economic analysis of WP4 will be considered in the present WP.
- Publications and water authorities: the data for WP5 will be collected from existing relevant publications, non-scientific publications and water authorities' databases.

Data Utility

The information obtained thanks to this WP will be 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 in the MED

Purpose

The sixth Work Package, led by the TUC, 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 (.csv), text files (.txt) and ESRI shapefiles (.shp).
 - Geospatial data: area boundaries, elevations and other topologic data that will be defined later during the project will be needed. The data that will be used for the multi-objective optimization will be digitalized, if needed, according to their spatial information. The data will be collected and generated in the format of Microsoft Excel documents (.xlsx) or comma separated values files (.csv), ESRI shapefiles (.shp), geodatabases (.gdb), ASCII grids files (.asc) and geoJSON files (.geojson).
 - 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: from this WP, the following three deliverables will be created. The lead participant of each deliverable is the responsible for its completion and deliver it to the rest of the InTheMED Consortium. Two of them are reports, which will be written in a word document (.docx). After their reviewing, the final version will be saved as a PDF document (.pdf). If the Atlas will be visualized within the Fuzzy WebDSS page, then its format will depend on the format of the Fuzzy WebDSS maps. Otherwise, it could be printed in hard copies from respective formats of images (i.e. .jpg, .tiff, etc).

Table 6. Deliverables of WP6

Deliverable number	Deliverable title	Short name of lead participant	Type	Diss. level
D6.1	Report on the development of the innovative DSS tool	TUC	R	PU
D6.2	Report on the results of the DSS for the case study sites	TUC	R	PU
D6.3	Atlas of the maps produced using the DSS	TUC	DEC	PU

Re-Use of Existing Data and its Origin

- Data from InTheMED project: the information produced in the previous WP (WP5) will be utilised for the implementation of the Fuzzy WebDSS. 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 will be studied. To perform this task, different sources of information will be 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 collection that is in this WP will be valuable 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 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 Work Package, number seven and led by the IST-ID, has the responsibility to develop novel communication strategies and dissemination materials to share the results of the project 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 InTheMED project to the whole MED region.

Type, Format and Expected Size of Data

- Meeting minutes: Living Labs and workshops for the community and stakeholders will be held. The comments, suggestions and so on discussed during these meetings will be documented in the minutes of the meetings. The minutes will be saved as a word document (.docx) or as a PDF document (.pdf).
- Posters, leaflets, brochures, factsheets, videos and presentations: it is foreseen to attend two scientific conferences (mid-term and final) to share main outcomes obtained by the work done in the previous six WP with the scientific community and public and private

organisms related to aquifer management. Moreover, there will be several exhibitions, seminars, webinars and workshops at national, MED and international level. For these, posters, leaflets, brochures, presentations, videos, factsheets and other promotional materials will be produced. The format will presumably be PDF documents (.pdf), Microsoft PowerPoint presentations (.pptx). AVI (.avi) and MP4 (.mp4) format.

- Newsletters: yearly newsletters will be made and sent to relevant members of the community. These are expected to be PDF documents (.pdf).
- Deliverables: in addition to the datasets mentioned above, the following deliverables will be created. The lead participant of each deliverable is the responsible for its completion and deliver it to the rest of the Consortium. All of them are reports, except for the project website which is currently being developed. The reports will be written in a word document (.docx) and after their reviewing, the final version will be saved as a PDF document (.pdf).

Table 7. Deliverables of WP7

Deliverable number	Deliverable title	Short name of lead participant	Type	Diss. level
D7.1a	Project website	IST-ID	DEC	PU
D7.1b	Communication and Dissemination Plan	IST-ID	R	PU
D7.2	Communication and dissemination activities	IST-ID	R	PU
D7.3	Report on synergies with groundwater initiatives in the Euro-MED region	UFZ	R	PU
D7.4	Report on mid-term workshop	IST-ID	R	PU
D.7.5	Report on the InTheMED final scientific conference	IST-ID	R	PU
D7.6	Exploitation plan	IST-ID	R	PU

Re-Use of Existing Data and its Origin

- Data from InTheMED project: all the research data gathered and produced by the previous WPs will be considered.
- Groundwater management approaches in the MED area: a review about the state-of-the-art innovative aquifer management approaches will be 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 re-usable research data. In the present chapter, it is outlined how this objective will be achieved 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]. Some issues will be more thoroughly defined in future versions of the DMP.

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, will be identifiable and locatable by means of a persistent and unique identifier.

These open InTheMED results will be deposited in Zenodo [4], an Open Access repository. Zenodo [4] 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 and 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 will be specified throughout the execution of the project and they will include a version number or a time stamp in ISO8601 format. Nevertheless, their structure can be anticipated to be as follows:

- For the deliverables: Project acronym, time stamp (YYYYMMDD), deliverable code, and version number.

E.g. “InTheMED_20201018_D14_v1”

- For the dataset files: Project acronym, time stamp (YYYYMMDD), WP number and “DS”, “INT”, which stands for dataset and interviews/questionnaires, respectively, or other proper resource identifier followed by a short description of its content.

E.g. “InTheMED_20201018_WP2_DSGroundwaterTimeSeries”

Regarding source software, Semantic Versioning Schema will be followed to number its release.

Search Keywords

Keywords will be added to optimise the possibilities of finding the outcomes and deliverables of the project. They will be defined following the terminology used in the various scientific fields addressed in the project and they will be descriptive to the content of each InTheMED result.

Metadata and Metadata Standards

In order to guarantee that data is easily locatable and EU funding is acknowledged bibliographic metadata will be 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 not 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 will follow the Dublin Core™ 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 Core™ Metadata Element set [8], 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, will be open by default with the exception of confidential deliverables, which are D1.1, D1.3 and D3.3. Likewise, meeting minutes and interviews and questionnaires made to stakeholders will be considered confidential. Confidential data will be stored in private servers of the Consortium members and only the Consortium will have access to it; hence, it will not be available for the public. Moreover, internal documentation will be posted and shared through the back-office system available at the InTheMED project's own website.

Currently, the InTheMED website is available, <https://inthemedprima.com/>, it has been developed by IST-ID, where information about the project, team and partners, highly relevant project results (datasets, deliverables and publications) will be linked to the repository, where they will be stored and from where they will be freely available for download. Moreover, it will exhibit web-based visualizations of the developed methods and models. Additionally, it is envisaged to 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 time series, which are the datasets attained from the WP2, will join the current of the industrial internet of things (IIOT) and they will be connected with the major portals sharing groundwater data, such as IGRAC (GGMN) and GEMStat.

Open InTheMED results will be deposited in an open access and public domain repository as soon as possible, especially, the data needed to validate the results in scientific publications. The European Commission funded site, Zenodo [4], located at CERN, has been selected as

InTheMED repository. The InTheMED community, where all the data and documentation will be collected, can be found following this link: <https://zenodo.org/communities/inthemed/>. This is a general-purpose open access repository that is technically compliant with open data requirements of OpenAIRE portal and Horizon 2020, hence, the process of reporting the InTheMED results in OpenAIRE is automatic.

In relation to peer-reviewed publications, they will be available in the website, as well as stored in the repository and submitted to open access relevant high-impact international scientific journals. There are two main ways to ensure open access to these publications, Green Open Access (self-archiving) or Gold Open Access (open access publishing). Presumably, InTheMED scientific publications will be published in Green Open Access, so the author's version or the final peer-reviewed manuscript will be deposited in the repository by the authors the same time as the publication date or within six months of publication if an embargo period is imposed by the publisher.

If source software is developed by InTheMED, it will be openly stored in the GitHub portal [10] along with its deposit in the repository.

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

If IPR issues arises 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. In these, the details of when the data will become available will be found. 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

There are no special software tools needed to access to data deposited as open data files in a repository nor in 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. However, most of the data is produced in ordinary and general electronic format that does not require specific software, except for some codes that may require particular compilers, such as MODFLOW, Python, MATLAB, feflow and Stella Architect.

Documentation of open source software required to access the data and developed by InTheMED will be made available on the project website and on the repository.

Access Requirements

The InTheMED results hosted in the selected repository, Zenodo [4], and the project website will be freely available without restriction to the public.

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.

Standard Vocabularies and Metadata Vocabularies, Standards and Methodologies

Following a standard vocabulary or ontology is advantageous to ensure the interdisciplinary interoperability of the InTheMED results. ISO/DIS 772 Hydrometry — Vocabulary and symbols [9] and ISO 19115-1:2014 Geographic information – Metadata – Part 1: Fundamentals [10] are some candidates, although the chosen vocabularies will be reported in future versions on the DMP.

Specifications regarding metadata have already been made in the second section of this chapter, “Making data findable”.

3.4. Increase Data Re-Use

Data Licensing and Date of Data Release

As the Article 26 of the GA states “results are owned by the beneficiary that generates them” and in case two or more beneficiaries have jointly contributed to the creation of 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 published through the aforementioned means. In order to protect the ownership, InTheMED datasets and publications will be release under a Creative Commons License, presumably Creative Commons Attribution-NonCommercial-ShareALike 4.0 (CC BY-NC-SA 4.0) [11] will be used. Regarding software licensing, copyleft licenses will be considered, such as GNU GPL [12].

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 chosen Open Access repository. Moreover, they will be made available in the project website as well. They will be release on publication date or within six months of publication if an embargo period is imposed by the publisher. The same measures will be taken concerning research data and related metadata, as it is 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 the selected repository 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 re-usable for at least one year after the end of the project, being 28 February 2023 the estimated project termination date.

Furthermore, Zenodo [4] commit 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 in the project website for at least five years, two more than the end of the project.

Data Quality Assurance Processes

InTheMED results will be corrected and validated by the partner involved in their production, then, the Consortium will perform a quality control in relation to the accuracy in following the FAIR conditions established in the DMP. After executing this data quality assurance process, the results will be made publicly available.

4. Allocation of Resources

Costs for Making Data FAIR

There are no costs anticipated to make open InTheMED results FAIR. Data archiving at the Zenodo [4] is completely free of charge. Likewise, licensing with Creative Commons [11] is free as well.

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

Responsibilities for Data Management

Data management is a shared duty between all InTheMED partners. The Coordinator of the project, UPV, is concerned about the overall data management at project level, while at WP level, the leader of each WP is responsible of the produced data. Each partner will handle their own datasets, they will be in charge of backing up and depositing their data in the repository. Nevertheless, all partners must inform and discuss with the rest of the Consortium before making any InTheMED result open. Moreover, the Coordinator will perform a quality control to assure that the FAIR conditions stated in the DMP are abided. Later, as the project activities progress, a short manual on how to upload the data to Zenodo [4] will be provided by the Coordinator to the rest of the Consortium.

The partner IST-ID is responsible for dissemination actions and website development, where the open InTheMED results will be available.

Costs for Making Data FAIR

There is no cost for long-term preservation of the InTheMED results and metadata in Zenodo [4].

5. Data Security

Provisions for Data Security

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

Long-Term Safe Preservation and Curation in Certified Repositories

The InTheMED results shared by Zenodo [4] will be safely preserved and curated indefinitely, as it is described in the fourth section of the present chapter, "Increase data re-use".

With regards to the project website maintenance, the domain will be 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.

6. Ethical Aspects

The interviews and questionnaires made to stakeholders will be considered as confidential, hence, they will not be publicly available. They will be stored in private servers of the Consortium and available only for the participants of the project.

InTheMED project does not involve personal data collection nor processing. Therefore, no other ethical issue has been found with regard to data management.

7. 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 Core™ Metadata Initiative (November 2020). *Dublin Core™ Metadata Initiative*. Retrieved from <https://www.dublincore.org/>
- [6] European Commission (2007). *INSPIRE*. Retrieved from <https://inspire.ec.europa.eu/>
- [7] Dublin Core™ 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] ISO/DIS 772 Hydrometry — Vocabulary and symbols. Retrieved from <https://www.iso.org/standard/73230.html>
- [10] ISO 19115-1:2014 Geographic information – Metadata – Part 1: Fundamentals. Retrieved from <https://www.iso.org/standard/53798.html>



- [11] 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/>
- [12] Free Software Foundation. *GNU General Public License*. Retrieved from <https://www.gnu.org/licenses/gpl-3.0.en.html>