Project deliverables

Deliverable #D6.3

Data Management Plan (DMP)

AGREEMAR

Adaptive agreements on benefits sharing for managed aquifer recharge in the Mediterranean region

Deliverable #D6.3

Data Management Plan (DMP)

Author(s)

Catalin Stefan (TUD)

Executive summary

This document contains the Data Management Plan (DMP) of the AGREEMAR project. The DMP describes the practices adopted for data management according to general guidelines set by the European Commission and in compliance with the Cooperation Agreement signed by the consortium. The document follows the structure of the template for Horizon2020 Data Management Plan (DMP), v2.0 from 15.02.2018 (available for download under: https://ec.europa.eu/research/participants/data/ref/h2020/other/gm/reporting/h2020-tpl-oa-data-mgt-plan-annotated_en.pdf). The document is published in its first version as project deliverable D6.3 and is intended to be updated during the course of the project as a living document.

Work package	Work package 6. Project management and coordination
Deliverable number & title	D6.3. Data Management Plan (DMP)
Partner responsible	Technische Universität Dresden (TUD)
Deliverable author(s)	Catalin Stefan (TUD)
Quality assurance	Jana Glass (TUD)
Planned delivery date	30.11.2022
Actual delivery date	24.11.2022
Citation	Stefan, C. AGREEMAR Deliverable D6.3: Data Management Plan. Available online at https://www.agreemar.inowas.com/deliverables.
Dissemination level	PU (Public)

Revision history

Version	Date	Author	Remarks
v0.1	18.11.2022	Catalin Stefan (TUD)	First version
v0.2	24.11.2022	Jana Glass (TUD)	Revision and quality assurance
v1.0	24.11.2022	Catalin Stefan (TUD)	Final version for publication

Abstract

This document contains the Data Management Plan (DMP) of the AGREEMAR project. The DMP describes the practices adopted for data management according to general guidelines set by the European Commission and in compliance with the Cooperation Agreement signed by the consortium. The document follows the structure of the template for Horizon2020 Data Management Plan (DMP), v2.0 from 15.02.2018 (available for download under: https://ec.europa.eu/research/participants/data/ref/h2020/other/gm/reporting/h2020-tpl-oa-data-mgt-plan-annotated_en.pdf). The document is published in its first version as project deliverable D6.3 and is intended to be updated during the course of the project as a living document.

Table of contents

Data Management Plan (DMP)

1. Introduction

AGREEMAR is a research project that proposes an integrated and coordinated methodology to assess and map the feasibility of managed aquifer recharge (MAR) in the Mediterranean basin and safeguard its success through the implementation of an adaptive governance framework. This will be achieved through the integration of site feasibility mapping, groundwater flow modelling, governance frameworks and stakeholders' engagement.

This document contains the Data Management Plan (DMP) of the AGREEMAR project. The DMP describes the practices adopted for data management according to general guidelines set by the European Commission and in compliance with the Cooperation Agreement signed by the consortium. The document follows the structure of the template for Horizon2020 Data Management Plan (DMP), v2.0 from 15.02.2018 (available for download under: https://ec.europa.eu/research/participants/data/ref/h2020/other/gm/reporting/h2020-tpl-oa-data-mgt-plan-annotated_en.pdf). The document is published in its first version as project deliverable D6.3 and is intended to be updated during the course of the project as a living document.

2. Data summary

2.1 Purpose of data collection / generation and its relation to project objectives

The main purpose of data collection is to generate site feasibility maps for MAR, to construct numerical groundwater flow models and to design adapted governance models to be applied at the four demo regions of the project. Additionally, data collected through interviews and direct surveys will be used to understand the socio-environmental conditions and to enable the implementation of the project. The generation of these key outcomes is linked to the overall project objective to optimize the hydrological balance in the Mediterranean countries by promoting technical and non-technical solutions for subsurface water storage and reuse. To achieve this, the project will focus on:

- Validation, optimisation and up-scaling adaptive and innovative water management strategies, such as MAR solutions, and use of non-conventional water sources to augment aquifer storage (O1).

- Improving the cross-sectoral uptake of MAR for climate change adaptation and ensuring the adoption of integrated governance models that will guarantee long-term, safe and efficient implementation, based on environmental, social and economic indicators (O2).

- Strengthening the institutional and managerial capacities of stakeholders to take up the integrated approach for planning and implementation of MAR (O3).

- Adoption of participative approaches to reduce barriers and fortify linkages among water resources managers and water users thus reducing conflicts and increasing social trust (O4).

- Demonstration of how innovative approaches for planning and implementation of MAR will lead to better use of freshwater and preservation of natural ecosystem services (O5).

2.2 Types and formats of the data generated / collected by the project

2.2.1 Types of data

- Data from self-administered online questionnaires
- Data collected from interviews with stakeholders and MAR experts
- Geospatial environmental datasets and time series of climate parameters
- Modelling data generated by water balance and groundwater models
- Georeferenced site feasibility maps

2.2.2 Formats of data

- MS Excel (.xlsx), comma-separated values (.csv) and text files (.txt) for time series and statistical datasets
- Shapefile and raster image formats (.shp, .shx, .dbf etc.) for maps and GIS files
- Specific formats for input / output files to be used by the modelling software MODFLOW, AQUATOOL (GeoJSON, .bda, .dat, .dbf, .pm, .prj etc.)

2.3 Reuse of the existing data

The project will reuse existing data generated by previous initiatives and also historical data available at the demo regions. Available datasets will be used to generate MAR feasibility maps, water management and groundwater flow models.

2.4 Origin of the data

Data will be obtained from public repositories of international, national and regional agencies, through direct contact with responsible authorities, by remote sensing or direct observations.

2.5 Expected size of the data

To be evaluated during the course of the project. The expected size depends on the methodological approach, i.e., number of feasibility criteria selected and complexity and scale of groundwater models.

2.6 Beneficiaries of the data ('data utility')

- Members of the AGREEMAR consortium
- General, regional and local stakeholders at the demo regions of the project
- PRIMA research community and European Commission services and agencies
- Broader research community
- General public

3. FAIR data

The AGREEMAR project strives to follow the FAIR principles (data shall be findable, accessible, interoperable, and re-usable) for the collection and generation of its research data.

3.1 Making data findable

A specific template will be created by the project coordinator and shared with the consortium partners in order to describe data collected and generated by the AGREEMAR project. The template will be stored in the project data repository and will be curated by the project coordinator. Whenever possible, the template will follow existing best practices for environmental and ecological metadata management.

A naming convention will be defined for datasets used in the project in order to facilitate quick recognition of the content and enable joint data usage regardless of format and language barriers.

Especially for self-generated data, a clear versioning system will be applied by adding a suffix to the name of the file, with documentation provided in the metadata library in case of using different versions of the same dataset. (e.g., consequent scenarios of the same model etc.).

3.2 Making data openly accessible

To share data among AGREEMAR consortium partners, the coordinator implemented a dedicated data repository. This is a group drive that is setup on the servers of TUD and that can be accessed by all partners via personalised login credentials. An internal data structure will include three types of access: a) data that can be openly made available to the general public; b) data that can be accessed and processed by all project

partners (not available to general public), and c) data that can be accessed under special conditions (according to provisions made in the Consortium Agreement and following the requirements set by the data provider).

The access to the internal AGREEMAR data repository will be possible through different protocols such as CIFS/SMB, NFS3, NFS4 and SSH and personalised TUD login credentials (already created for to all project partners at the start of the project). The most common access connection is through a SFTP programme using the data gateway sftp dgw.zih.tu-dresden.de and the authentication method using the TUD login and password.

The project-generated datasets that can be made openly available to the public will be published via the project website in a visible, dedicated section (e.g., '/Resources'). Each data output will be equipped with data utilisation license that will specify the conditions for further use and dissemination. Further certified data repositories will be explored for making data openly accessible. Specific outputs such as the MAR feasibility maps will be submitted for publication in the Global Groundwater Information System (GGIS) managed by the International Groundwater Resources Assessment Centre (IGRAC). GGIS is an online platform supporting the sharing of groundwater data and information worldwide: https://www.un-igrac.org/global-groundwater-information-system-ggis and includes a portal dedicated to MAR and MAR suitability mapping: https://www.un-igrac.org/ggis/mar-portal.

To access the data, conventional software for environmental applications can be used, such as GIS software (e.g., QGIS - https://www.qgis.org) for shapefiles and raster images, MS Office suite and statistical software like R for time series, MODFLOW-specific interfaces for MODFLOW files etc. At the core of the numerical groundwater flow simulations is the free, web-based INOWAS platform (https://www.inowas.com) which has detailed, online documentation for all the simulation tools (https://www.inowas.com/tools/).

3.3 Making data interoperable

To allow the interoperability of stored and created data, common data formats such as CSV, txt, tiff or shp are used. Those data formats can be opened by a range of software tools and allow the exchange of data and results.

The web-based INOWAS platform used for groundwater modelling in the AGREEMAR project already implemented data formats and standards for easy interoperability. This includes the generation of GeoJSON files for model input but also the import of data to the platform using the OpenMetrics standard (https://www.openmetrics.io).

The data and metadata will be named in English using standard vocabulary to facilitate interoperability.

3.4 Data re-use

This section will be compiled in the course of the project once the different datasets will be available to the AGREEMAR project and the specific requirements of the data providers will be known.

4. Data management

4.1 Resources allocated to data management

The AGREEMAR data repository is a service provided free of charge by the Center for Information Services and High Performance Computing (ZIH) of TUD. The service is provided to all TUD staff members and can be used for the management of research data in joint projects. The repository will be available for the entire duration of the project, after which the consortium will explore alternative solutions such as certified data repositories.

The research output resulted from the project will be published, whenever possible, under open-access conditions. Each partner is responsible to cover the costs for such publications if compliant to the national funding conditions.

4.2 Responsibilities for data management

The responsibility for data management lies with the project coordinator TUD who will be responsible for:

- Setting up the data repository according to technical project requirements
- Create accounts for project partners and facilitate access to the repository
- Maintain the good functionality of the repository
- Create templates and safeguard the compliance with data naming and versioning scope
- Support partners to manage datasets hosted in the repository.

The project coordinator is not responsible for interruptions of the data repository services that are due to force majeure. The responsibility of the content (datasets) belongs to the partners who uploaded the data.

5. Data security

The following data security provisions are available for AGREEMAR data repository according to the conditions offered by the service provider:

- Secure access only from TUD campus or through VPN
- Authentication using provided user name and password
- Data mirror to a second location
- Snapshots available depending on setting (by default: one snapshot every 4 hours from 8 a.m. to 4 p.m. on work days, these are kept for two days; one snapshot per day for 12 more days; 13 snapshots each week; 6 more snapshots per month)
- Backup once per week
- Storage up to 5 TB data and 5,000,000 files.

For data published in external repositories, the specific provisions of the respective systems will apply.

6. Ethical aspects

Not applicable.

Acknowledgement