AWESOME

WATER-ECOSYSTEM-FOOD

DATA MANAGEMENT PLAN

October, 2020

Programme Call:	PRIMA Call 2019 Section 1 Farming RIA
Project Number:	1942
Project Title:	AWESOME
Partners:	POLIMI (Project Coordinator), AUEB, YVC, UH, AF, RWTH, FEEM
Work-Package:	WP1
Deliverable #:	D1.3
Deliverable Type:	Document
Contractual Date of Delivery:	31 October 2020
Actual Date of Delivery:	23 October 2020
Title of Document:	Data Management Plan (DMP)
Author(s):	Andrea Castelletti, Elena Matta
Content of this report:	The Data Management Plan describes the management procedures and standards for data collection and generation, security, budget and ethical issues. It also presents the current metadata catalogue of the project.
Availability:	This report is public.

Document revisions		
Author	*Revision content*	*Date*
Elena Matta	D1.3_v01 – First draft of the data management plan (DMP) and of the metadata catalogue	Sept. 27, 2020
Matteo Vincenzo Rocco	Review of D1.3_v01	Oct. 10, 2020
Elena Matta	D1.3_v02 – Second draft of the DMP	Sept. 27, 2020
Ruslana Rachel Palatnik	Review	Oct. 15, 2020
Elena Matta	D1.3_v03 – Third draft of the DMP	Oct. 15, 2020
Andrea Castelletti	Review	Oct. 19, 2020
Elena Matta	D1.3_v04 – Fourth draft of the DMP and refinements of the metadata catalogue (v01)	Oct. 23, 2020

Table of Content

LIST OF ACRONYMS

Abbreviations

AB: Advisory Board
AHD: Aswan High Dam
CA: Consortium Agreement
CS: Case Study
DDP: Deliverable Development Plan
DoA: Description of Action (Annex I of the Grant Agreement)
DM: Deliverable Manager
EU: European Union
GA: Grant Agreement
GAs: General Assembly
GERD Grand Ethiopian Renaissance Dam
GDP: Gross Domestic Product
GHG: Greenhouse Gas Emissions
HPP: Hydropower Plant
IEA: International Energy Agency
IRENA: International Renewable Energy Agency
MB: Management Board
MED: Mediterranean
Mx: Month number
PC: Project Coordinator
PI: Principal Investigator
PO: Project Officer
PR: Project Review
PV: Photovoltaic
RES: Renewable Energy Sources
RO: Reverse Osmosis
RP: Reporting Period
SSP Shared Socio-economic Pathways
QC: Quality Control
QM: Quality Management
UoK: University of Khartoum
WP: Work Package

EXECUTIVE SUMMARY

The AWESOME Data Management Plan (DMP) defines the general policy and approach to data management in the project and assesses the related issues at the administrative and technical level, following the FAIR (Findable, Accessible, Interoperable and Reusable) principles of the European Union (EU) guidelines. The DMP will evolve during the project lifetime. Next versions will refine the data policy aspects and will present the datasets collected and generated by the AWESOME project in more detail. It is worth specifying that the AWESOME Grant Agreement (GA) and Consortium Agreement (CA) take precedence over this document, which does not replace by any means the contractual obligations among partners, and between partners and the PRIMA Foundation.

INTRODUCTION

Deliverable D1.3 is the AWESOME Data Management Plan (DMP) and is designed to set out key operational procedures to handle and manage research data, to ensure data security and quality as well as to foster data exchange and cooperation following the FAIR (findable, accessible, interoperable and re-usable) principles. The purpose of the DMP is to support the data management life cycle for all data that will be collected, processed or generated by the project. Indeed, the DMP describes the policy and the procedures concerning the acquisition, storage, classification, management, protection, and distribution of project data. This activity will also take care of timely and effective publication of relevant data and project results through the production of peer-reviewed articles (open access) as well as the active participation in international conferences, and - when applicable - as open data in online repositories, to improve project impacts and post-project legacy. The AWESOME ethical procedures are briefly presented as well in this document. At this early stage, the nature and extent of the datasets that the project will collect and generate can be assessed in a preliminary way. This DMP is thus an initial version that will be periodically updated during the project lifetime and the document stored in the internal project repository. It is advised to update the DMP as necessary and formally at the official reporting to PRIMA in M22 and in M42).

The document is structured after the European guidelines in the following way: Section 1 is the Data Summary and describes the purpose behind data collection, process and generation, and their relation to AWESOME objectives, while Section 2 explains how the FAIR principles are applied to the project, including tools, methodology and licences adopted. Section 3 presents how the project resources are allocated for data management, while Section 4 is dedicated to data security and storage. The final sections (5-7) are dealing with the ethical and legal aspects, and further issues that may arise. The document ends with the Annex 1, which contains the preliminary metadata catalogue, and is followed by the Annex 2 with the full list of references of the Annex 1.

1. DATA SUMMARY

In a context of increasing population growth in the Mediterranean (MED) region and, thus, increasing water, energy and food demands, the main objective of AWESOME is developing a decision-analytic platform based on a multi-level, integrated modelling approach to address the water-ecosystem-food (WEF) Nexus across a hierarchy of spatial scales: from the macroeconomic development to regional planning down to the single farm. The adoption of integrated and participatory approaches in a transboundary basin such as the River Nile Valley enables the project researchers to explore tradeoffs, synergies, and nested interdependencies across sectors and scales and to generate shared economic, environmental, and societal benefits. AWESOME will develop alternative WEF planning portfolios composed of regional policies, river-basin strategic planning solutions, and innovative technological options demonstrated at the local scale.

For all the reasons above, data management turns out to be very important for the successful achievement of the AWESOME goals. At this early stage, it is important to assess the nature and extent of the datasets the project will collect, process and generate, even if in a preliminary way. This initial version of DMP and metadata catalogue serves as fundamental step for model setup and integration for all the involved partners, to foster discussions and implementation, to then timely and successfully fulfill the project tasks. The initial version of the metadata catalogue (V01) can be found at the end of this document in the Annex 1, while the linked references in Annex 2. As mentioned in the introduction, periodic updates of the DMP and metadata catalogue will follow during the project period. The metadata catalogue is available as a Google Sheet[1] shared among the principal investigators (PIs) and is stored as an Excel file in the internal project repository[2] (further details in the next section, §2) following the guidelines for naming and version numbering after updates described in the Project Management Plan (PMP, D1.1). The reference directory in the AWESOME repository is “AWESOME_public/Data/Metadata/”, while the first version of the metadata file is named as “AWESOME_D13_POLIMI_WP1_V01D_Metadata.xls”.

Even if AWESOME data are described in detail in the Annex 1, a summary of the main data information follows here. Firstly, data will be collected, serving as inputs for the models at the different scales i.e. top-down: MED Region, Nile River Basin, Egypt, test site(s) of the different aquaponics/hydroponics technologies (probably in Cairo and El Gouna, Egypt). The input data will be processed in the models and used for the local experiments. Secondly, data will be generated by AWESOME and turn into outcomes of the different models and work package (WP) tasks. Therefore, data will be:

- **Collected** from multiple sources – e.g. literature, global datasets and prospects, environmental agencies, universities, research centres – to identify demographic, climatic,

1 https://docs.google.com/spreadsheets/d/1E9kqONldIWXKC7aG_lU3kNIsoGpA1U9w2XgCCsXD_Uo/edit#gid=80699478

2 https://131.175.15.9/cgi-bin/

economic, energetic, land use and hydrological aspects and indicators, and to characterize hydrologic regime, agriculture developments, energy production, social and economic systems, terrestrial and aquatic ecosystems. Data collected are structured in the metadata table according to the topic, to the different spatial and temporal scale (e.g. climate data for a certain country/region, recorded at a certain gauge, over a certain period, with a certain frequency) and to the correspondent WP involved/interested in the data. Information concerning the source, the format and the file naming are also given, as long as currently applicable. The definition of the modelling baseline and variation scenarios are under internal discussion at the moment among WP2-3-4, e.g. to download data from the global models agreeing on the most proper combinations/matrixes of SSPs and RCMs (respectively Shared Socioeconomic Pathways and Regional Climate Models) for the project aims and to align the different model setups accordingly.

- **Generated** from AWESOME models – i.e. macroeconomic, energetic, climatic and demographic models at the MED scale, decision analytic framework (DAF) model for the Nile River Basin – and from innovative technological solutions at the local scale, like hydroponics and aquaponics in the pilot sites. As for data collection, also for data generation the definition of the modelling baseline and variation scenarios, as well as model boundaries, are currently under internal discussion among WP2-3-4, with the aim to align the different model setups accordingly. Data generated from models will be used in fact to simulate scenarios and future projections, to then create planning portfolios adopting different water, ecosystems and food management policies, aimed to assess the impacts of changes in the system. Existing data, available in public repositories or from institutional partners or in literature will be re-used whenever possible, while new generated data will be stored securely in the AWESOME internal repository, as addressed later. Generated data will be presented through the publication of results in peer-reviewed publications (open access) and in form of oral contributions and posters at international conferences. If applicable, some generated data will be published on open access repositories like *Zenodo*[3], (for details, s. §2 of this document).

2. FAIR DATA

As announced in the PMP (D1.1), AWESOME will manage data through an internal repository[4], curated and hosted by Politecnico di Milano (POLIMI), where all technical information (e.g. officially released documents, contractual information, templates, meeting minutes) about the project are stored in a structured way in shared directories (Fig. 1). Data are also stored and exchanged among project partners using the same repository, which has a total capacity of 16,607 GB. The access to the internal repository is password-protected and restricted to designated project partners, as

[3] Free and open digital archive built by CERN and OpenAIRE. (https://zenodo.org/)
[4] https://131.175.15.9/cgi-bin/

described in §4.2 of the PMP. In case of need of higher storage space, a cloud service external to the project (*Dropbox Business*[5]) can be made available by POLIMI.

The final project outcomes, deliverables and scientific publications will be published and maintained on the AWESOME website[6] and eventually on *Zenodo*, on which will be automatically become visible on the *OpenAIRE* portal[7]. It will be internally discussed whether to store some data in open access on *Zenodo* as well and to which extent, leaving the final decision to the Principal Investigators (PIs) of each WP. Information about relevant project updates will be announced on the AWESOME website and on the AWESOME Twitter account (@AWESOME_PRIMA), as described fully in D7.1 (Multi-stakeholders Outreach Plan).

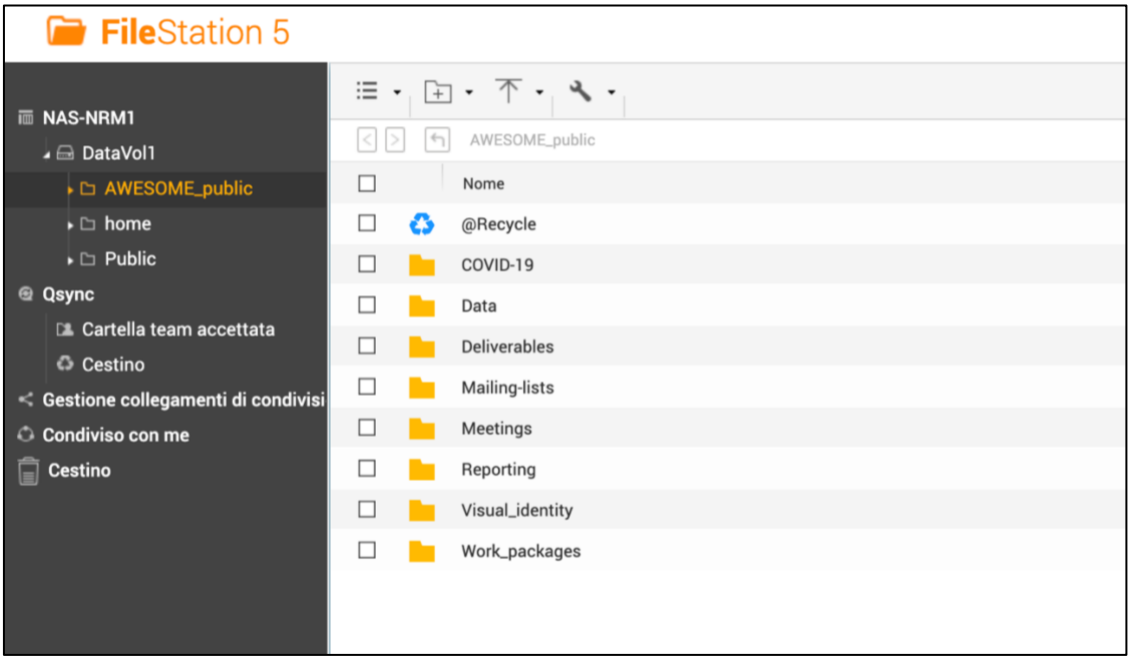

Figure 1 Overview of the shared directories in the AWESOME internal repository hosted by POLIMI

2.1. Making data findable, including provisions for metadata

In addition to what mentioned before, AWESOME would adopt the *Zenodo* repository to publish project outcomes and datasets, if applicable. Using this repository, all the public data of the project will be provided with a Digital Object Identifier (DOI) and a common dataset of metadata (based on Dublin Core[8]). Keywords would be sourced by the standard dictionaries, like the USGS water dictionary[9]. Versions of each dataset would be numbered and report main and minor changes. Main version change would occur after significant changes in the data (e.g. change in the data structure,

[5] https://www.dropbox.com/business

[6] https://awesome-prima.eu/

[7] European Commission Portal for reporting H2020's scientific publications. (https://www.openaire.eu/)

[8] A set of "core metadata" for simple and generic resource descriptions (http://dublincore.org).

[9] Water Science Glossary of Terms, compiled by USGS (https://water.usgs.gov/edu/dictionary.html).

massive correction or update, changes in the procedure for data collection or generation), while minor version changes after data updates or limited correction. Any changes would also be mentioned in the description metadata field. Metadata identification, description and naming conventions are reported, whenever applicable, in the attached metadata catalogue (Annex 1). Annex 1 will evolve with the progress of the project, as the data availability and their format will become clearer during the course of data collection activities.

2.2. Making data openly accessible

According to art. 29.3 of the GA, it is not required to the AWESOME Consortium to make research data openly accessible. However, some categories of data are already with open access in the internet (e.g. global datasets, global and regional projections such as CORDEX; s. Annex 1 for detailed information).

Data generated in AWESOME could be published in the *Zenodo* repository, together with associated metadata, if applicable and agreed among the involved institutions, and if relevant for validation of scientific publications and/or deliverables, as described in the foreword of §2 and §2.1. Data collected in AWESOME from various institutions and agencies, and thus with closed access, might be published on *Zenodo* as well, but only after written agreements with the correspondent data sources. Data collected and generated, along with all relevant project documents, will be securely stored in the internal project repository hosted by POLIMI (foreword of §2, and Fig.1). In general, data will be stored using standard formats specified in Annex 1 for each dataset.

Finally, it will be internally discussed among PIs whether to publish any open source software and tool developed within AWESOME on public software code repositories, like Github[10]. E.g. POLIMI has already published the so-called ClimateScenarioAnalysisToolbox[11], where some scripts (in Matlab and R) for downscaling climate scenarios are freely available.

2.3. Making data interoperable

The AWESOME project covers several disciplines and scientific foci; therefore, it is of highest priority to find a common language between partners and stakeholders, to integrate data and information from the different domains. In order to provide a common understanding on data within the project itself, the use of the Dublin Core Metadata Element set vocabulary will be adopted, which is a basic, domain-agnostic standard which can be easily understood and implemented, and as such is one of the best known and most widely used metadata standards. It is sponsored by the Dublin Core Metadata Initiative and was published as ISO Standard 15836 in February 2009 (reference and link in the footnote of §2.1).

Whenever possible and useful, more discipline-specific metadata will be also adopted, as those defined by OGC[12] for geospatial data, or further economic/agriculture-related glossaries.

[10] Online project hosting platform (http://github.com/).

[11] https://github.com/mxgiuliani00/ClimateScenarioAnalysisToolbox

[12] Open Geospatial Consortium (http://www.opengeospatial.org/).

2.4. Increase data re-use (through clarifying licences)

Data reuse for the public is not required by AWESOME (art. 29.3 of the GA). Nevertheless, there are few points that should be addressed and reminded to ensure a smooth management of research data and research results.

According to art. 26 of the Grant Agreement (GA), research data and software are owned by the PI (beneficiary) that generates them. In case of joint ownership of results, each PI must agree (in writing) on the allocation and terms of exercise of their joint ownership, stipulating a joint ownership agreement, to ensure compliance with their obligations under the GA. The PRIMA Foundation may assume ownership of results to protect them (art. 26.4 of GA for further details). Notwithstanding the above, owners of open results arising from the AWESOME project are encouraged to release their work under a Creative Commons license, preferably Creative Commons Attribution 4.0 CC-BY-4.0 licence[13].

Concerning the dissemination and exploitation of results, as written in art. 29 of the GA, each PI must 'disseminate' its results by disclosing them to the public by appropriate means (i.e. peer-reviewed journal articles) and ensure open access – free of charge, online access for any user. In particular, each PI that intends to disseminate its results must give advance notice to the other PIs of at least 45 days, together with sufficient information on the results to be disseminated, while any other PI may object within 30 days of receiving notification. Further, 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 (art. 29 of the GA for further details). Any update of relevant project outcome will be published on the AWESOME website (and on the AWESOME Twitter account) as addressed in detail in D7.1 (Multi-stakeholder Outreach Plan) and open access publications will be advertised on various partner institutions websites, as well as on scientific social networks like *Research Gate*[14].

In regard to research data and results ownership, each PI may transfer ownership of its results. It must however ensure that its obligations under art. 26.2, 26.4, 27, 28, 29, 30 and 31 of the GA also apply to the new owner and that this owner has the obligation to pass them on in any subsequent transfer. This does not change the security obligations in art. 37, which still apply. Nevertheless, the PRIMA Foundation has the right to object to any transfers or licensing, if the conditions explained in art. 30.3 apply.

Quality assurance concerning accuracy and completeness of metadata will be performed with the overall supervision and responsibility of WP1, but with mandatory participation and collaboration by the other WP and Case Study (CS) Leaders, since they are responsible for data collection, process and generation within their tasks – in agreement with the project ethical obligations, as addressed later in §5. In detail, data quality assurance will be performed through the following steps:

- Collected data

[13] Creative Commons Attribution-ShareAlike 4.0 International Public License (https://creativecommons.org/licenses/by-sa/4.0/legalcode).

[14] ResearchGate is an open scientific social network for researchers (www.researchgate.net).

1. Storage of raw datasets, without any further processing, in a dedicated folder (under "AWESOME_public/Data/WP…/…") in the project internal repository;
2. Data check and editing for assuring positional, attribute and temporal quality, completeness and consistency, under the responsibility of the project partner interested in the data ("Contributor"), as listed in Annex 1;
3. Compilation of metadata updating the metadata catalogue of Annex 1 (both generic and domain specific where applicable) reporting a brief summary with the editing done; where the metadata file can be found at AWESOME_public/Data/Metadata/AWESOME_D13_POLIMI_WP1_V01D_Metadata .xls" and is available as shared Google Sheet (see §1 of this document);
4. Storage of the final version of the datasets in a dedicated folder in the internal repository under the directory "AWESOME_public/Data/Collected/WP…/…";
5. Uploading of the datasets on *Zenodo* repository, if compliant with agreements among PIs and eventual limitation detailed in Annex 1, as well as if relevant for maintaining also after the project lifetime.

- Generated data
    1. Compilation of metadata updating the metadata catalogue of Annex 1 (both generic and domain specific where applicable) reporting a brief summary with the editing done; where the metadata file can be found at AWESOME_public/Data/Metadata/AWESOME_D13_POLIMI_WP1_V01D_Metadata .xls" and is available as shared Google Sheet (see §1 of this document);
    2. Storage of the final version of the datasets in a dedicated folder in the internal repository under the directory "AWESOME_public/Data/Generated/WP…/…";
    3. Uploading of the datasets on *Zenodo* repository, if compliant with agreements among PIs and eventual limitation detailed in Annex 1, as well as if relevant for maintaining also after the project lifetime.

Any update concerning data collection and/or generation within the AWESOME project as well as any editing of the metadata catalogue should be promptly communicated to the Project Coordinator (PCo) and to the PIs by short email notice.

3. ALLOCATION OF RESOURCES

The metadata catalogue reported in Annex 1 identifies, for each dataset, the responsible WP. Costs are included in the tasks related to data collection and generation and cannot be listed separately. Costs of the project internal repository are covered by internal POLIMI resources. As already mentioned in §2, *Dropbox Business* accounts will also be activated if needed, in order to have more space where to store and share raw datasets and simulation outputs. Eventual costs related to these accounts will be changing according to the numbers of accounts and disk space needed for each of them and they will be covered by project budget.

4. DATA SECURITY

The tools mentioned in the foreword of §2 of this document are hosted by POLIMI and the internal project repository on the POLIMI server is protected by firewall and institutional security policies. In detail:

- The internal repository is relying on POLIMI storage facilities and accessibility is reserved, protected by username and password known only by the selected users (PIs), for both upload and download functionalities;
- *Zenodo* repository is hosted at CERN and it is subject to its rules for data security as reported at https://zenodo.org/policies

All datasets maintained on the POLIMI server will be periodically subject to incremental backup in order to avoid data loss. WP1 (POLIMI) has the responsibility of management, organisation and preservation of the AWESOME repository.

Finally, in regard to security, AWESOME does not involve any activity raising security issues and does not handle EU classified information, neither as background nor as result (as in §5.2 of the DoA on security issues).

5. ETHICAL ASPECTS

As written in §5.1 of DoA, AWESOME will involve stakeholders that will participate in a series of meetings both in person and online, as well as in questionnaires performed by WP6. The purpose is to bring their local and regional expertise into the project and to share knowledge and develop capacity in the approaches and tools used in the project. Confidentiality will be assured, in that:

- No personal information will be collected from participant stakeholders other than contact information and a brief description of their work, and this information will not be shared beyond the limits of the project without the explicit written consent of the stakeholders.
- No quotes made by the stakeholders in the online discussion of the workshops will be published in print or on the internet without their explicit written consent.
- No further ethical issues are anticipated, but should they arise, the project consortium will ensure that EU legislation, international guidelines and the ethical and legal requirements of the countries involved in the project are adhered to.

Ethical aspects related to data management are following the obligation to comply with ethical and research integrity principles of art. 34.1 of the GA, for which the PIs must follow the ethical principles (including the highest standards of research integrity) and the applicable international, EU and national law. The PIs must ensure that the activities, also the data-related ones, under the action have an exclusive focus on civil applications. In addition, the PIs must respect the fundamental principle of research integrity — as set out in the European Code of Conduct for Research Integrity[15].

[15] European Code of Conduct for Research Integrity of ALLEA - All European Academies (http://ec.europa.eu/research/participants/data/ref/h2020/other/hi/h2020-ethics_code-ofconduct_en.pdf).

This implies compliance with the fundamental principles detailed in the art. 34.1 of the GA, applied here to data, which are wrapped up here in short:

- reliability in ensuring the quality of research data;
- honesty in developing, undertaking, reviewing, reporting and communicating research data in a transparent, fair and unbiased way;
- respect for colleagues, research participants, society, ecosystems, cultural heritage and the environment;
- accountability for the research data from idea to publication, for data management and organisation, and well as for their wider impacts meaning that the PIs must ensure that persons carrying out research tasks and processing data follow the good research practices and refrain from the research integrity violations described in this Code.

This summary of the ethical aspects here does not change the other obligations under the GA or obligations under applicable international, EU or national law, all of which still apply.

Finally, it is worth specifying that informed consent will come along with data sharing and long-term preservation, in case of questionnaires dealing with personal data implemented within the project.

6. OTHER ISSUES

At the time of the DMP deployment, the AWESOME project does not make use of any other national/funder/sectorial/departmental procedures for data management. This will be reported in case of occurrence on a later stage of the project (further versions of the DMP).

ANNEX 1: METADATA CATALOGUE

The first version of the AWESOME metadata catalogue is reported in Table 1, which will be updated and further completed during the project lifetime. The references mentioned in Table 1 are extensively reported in Annex 2.

Table 1 – AWESOME metadata catalogue, version 01 (V01).

Location	Topic	Description	Reference partner	Spatial coverage	Temporal coverage	Source	Quality	Contributor (WP)	Data format	File name composed by	Notes
MED Area, Egypt	Climate - historical meteorological variables (precipitation, Temperature, Radiation, Humidity, Wind Speed and Direction, etc.)	Station data from national agencies	POLIMI	River Nile Basin		Existing at National Water Research Center?		2			Not yet available
MED Area	Climate - historical meteorological variables	Gridded reanalysis/satellite datasets (e.g. Copernicus ERA5, CHIRPS, etc)	POLIMI	Worldwide/ MED Area	1950-2020 (ERA5), 1981-2015 (CHIRPS)	Various, e.g. CHIRPS: Funk et al. (2015)	Good	2	xls, nc (netcdf)		(Pre-) processing needed
MED Area	Climate - past conditions	Climate model simulations over control period	POLIMI	Worldwide	1981-2005	CORDEX	Good	2	nc (netcdf)		(Pre-) processing needed

Location	Topic	Description	Reference partner	Spatial coverage	Temporal coverage	Source	Quality	Contributor (WP)	Data format	File name composed by	Notes
MED Area	Climate - future conditions	Climate model simulations over future horizons for different scenarios	POLIMI	Worldwide	2006-2100	AWESOME	Good	2, 3			(Pre-) processing needed
Egypt	Energy Demand Projection	The projection of yearly total demand of energy carriers specific for the defined nodes if it is needed	POLIMI	Country/ Sub-Country Node specific	ModelLifeTime (Yearly)	Econometric Models using the endogenous results of other models such as CGE model; IRENA Report: Renewable Energy Outlook for Egypt	Scenario specific: depends on the definition of scenario and the quality of the economic model output	2	xlsx, csv	Energy carrier, year, demand	
Egypt	Renewable Energy Resource (RES) Availablity	Availability of RES based on the technology and location	POLIMI	Based on the Location of the RES, if the model is multi-node, or the average for the country in case of a single node model	Can be considered constant for the whole model lifetime	RenewableNinja: an open-source model to estimate the hourly availability of wind and solar sources for a specific region, hydromodels from other groups	High	2	xlsx, csv	Location, year	Based on the number of nodes, we may need to make estimations because of high variability of the data based on location

Location	Topic	Description	Reference partner	Spatial coverage	Temporal coverage	Source	Quality	Contributor (WP)	Data format	File name composed by	Notes
Egypt	Techno-Economic Parameters for Energy Production Techs	Full detail of current and future techno-economic parmeters of generation technologies, e.g. efficiencies, capital, fixed and operation & maintenance cost, the price of fuels	POLIMI	Country	If the model period is not so long, can be considered constant	General information: average techno-economic parameters of the world, country Specific information	General information: Good	2	xlsx, csv	Parameter, technology, location, year	
Egypt	History - Current Energy Situation	Historic - Current Energy generation mix with full details for calibration of the model	POLIMI	Country\Nodes	Year	International Energy Agency (IEA), Egyptian Electricity Holding Company	Good	2	html --> xlsx	Energy Production, technology, year	General Categories are available not high details
Egypt	Fossil Fuels (FF) Resource Availability	Defines maximum extractable FF from domestic reservoirs	POLIMI	Egypt	Yearly Whole Timeline	EGPC, BMI (fitch group company)	High	2	xlsx	FF Type, Year, Availability	

Location	Topic	Description	Reference partner	Spatial coverage	Temporal coverage	Source	Quality	Contributor (WP)	Data format	File name composed by	Notes
World or at least MED Area	Water types for crops production	For each crop what water types can be used and what is the substitutability between them	POLIMI	Country or regional level	1990-2016	POLIMI		2, 3	tif; asc	Crop, water type, year	Data depending on croplands data availability. At the moment only year 2000 croplands are available at global scale
World or at least MED Area	Watneeds (crop model)	Precipitation and Evapotranspiration data	POLIMI	Global (10 km resolution)	1980-2018	CRU CL 2.0 New et al. (2002), Harris et al. (2014)	Good	2	netcdf	Variable, year, month	
World or at least MED Area	Watneeds (crop model)	Maximum available soil moisture	POLIMI	Global (1 km resolution)		HWSD, FAO (2012) Nachtergaele et al. (2009)	Good	2	tif		
World or at least MED Area	Watneeds (crop model)	Maximum infiltration rate	POLIMI	Global (1 km resolution)		BGR & UNESCO (2008)	Good	2	tif		
World or at least MED Area	Watneeds (crop model)	Crop parameters	POLIMI	Crop specific		Portman et al. (2010); Siebert and al., (2012)	Good	2	xls		

Location	Topic	Description	Reference partner	Spatial coverage	Temporal coverage	Source	Quality	Contributor (WP)	Data format	File name composed by	Notes
World or at least MED Area	Watneeds (crop model)	Watneeds Output: Yearly green and blue water, precipitation, runoff and deep percolation fluxes for 26 main crops	POLIMI	Global (10 km resolution)	2000; 2016	AWESOME	Good	2	tif, asc	Crop, water type, year	Data depending on croplands data availability; as of Oct. 22 (2020) only year 2000 croplands are available at the global scale
World or at least MED Area	Watneeds (crop model)	Watneeds Output: Monthly green and blue water, precipitation, runoff and deep percolation fluxes maps for 4 main crops (namely wheat, rice, maize and sugarcane)	POLIMI	Global (10 km resolution)	2000; 2017	AWESOME	Good	2	tif, asc	Crop, water type, year	Data depending on croplands data availability. At the moment only year 2000 croplands are available at global scale

Location	Topic	Description	Reference partner	Spatial coverage	Temporal coverage	Source	Quality	Contributor (WP)	Data format	File name composed by	Notes
Egypt	Population	Key demographic indicators for different region, subregion, country or area and for different periods	AUEB	Worldwide	1950-2100	UN World Population Prospects 2019	Good	2, 3	xlsx, csv	Prospect acronym, variable acronym, version, variable name	
Egypt	Population - SSP	Projections of SSP population, urbanization, and GDP projections	AUEB	Worldwide	various	International Institute for Applied Systems Analysis (IIASA)	Good	2, 3	xlsx, csv		
Egypt	Economic drivers - SSP	Capital accumulation, saving rates, relationship between saving rates and investment rates, education, female participation, energy, total factor productivity	AUEB		1980-2100	CEPII	Good	2, 3	xlsx, csv		

Location	Topic	Description	Reference partner	Spatial coverage	Temporal coverage	Source	Quality	Contributor (WP)	Data format	File name composed by	Notes
Egypt	Input/Output (IO) tables with matching environmental and social satellite accounts	720-line item environmental indicators covering GHG emissions, labour inputs, air pollution, energy use, water requirements, land occupation, N and P emissions, primary inputs to agriculture (including 172 crops)	AUEB		1990-2015	KGM & Associates Pty. Ltd.		3			
MED Area	Macro-economic model	GDP and other economic indicators	YVC&HU	Country or regional level	2015-2050	AWESOME	Good	3	xls		Not yet available
MED Area	Macro-economic model	Agricultural output and prices, and water use by scenario	YVC&HU	Country or regional level	2015-2050	AWESOME	Good	3	xls		Not yet available

Location	Topic	Description	Reference partner	Spatial coverage	Temporal coverage	Source	Quality	Contributor (WP)	Data format	File name composed by	Notes
Nile at different locations (Lake Tana, Roseires, Sennar, AHD, GERD, Tekeze, Upper Atbara and Setit, Khashm El Girba, Jebel Aulia, Merowe)	Hydrography - Hydraulic infrastructure and hydropower station	Reservoir technical details, level to storage tables, releases, withdrawals, demands, power, reservoir operations, evaporation rates, water balance	POLIMI	Nile reservoirs	Existing + planned (mostly monthly constant data)	Papers and technical reports (complete references in the Annex 2; e.g. Wheeler et al. 2016)		4	pdf, xlsx	Table name, reservoir/HPP name, version	Constant/ unique set data
Nile at different locations (Victoria, Kyoga, Albert, Torrents, Sobat, Tana, Kessie, Border, Dinder, Rahad, Atbara)	Hydrography - literature data	Historical and stochastic time series: streamflow	POLIMI	Nile gauging station network	1912-2003 (monthly)	Papers and technical reports (complete references in the Annex 2; e.g. Jeuland and Whittington, 2014)		4	xlsx	Variable name, station name, source, version	Missing data
Nile Basin	Hydrography - meso scale model	DAF model simulation and indicators time series	POLIMI	Nile River Basin (from GERD to AHD or Nile Delta)	At least 30 years past + projections until 2050 ?	AWESOME	To evaluate after simulations	4	csv, ascii	Variable acronym, station name, pathways, version	

Location	Topic	Description	Reference partner	Spatial coverage	Temporal coverage	Source	Quality	Contributor (WP)	Data format	File name composed by	Notes
Egypt, Israel	Soil data - soil properties	Sand parameters (grain size, grain shape, mineral composition)	RWTH	Country		The Hebrew University of Jerusalem, literature		5	docx, xlsx		Not yet available
Egypt, Israel	Reverse Osmosis (RO) water salinity data	RO saline water source (salinity of influx water [ppt or ppm], amount, salinity and quality of brine per m^3 fresh water, management of brine)	RWTH	Local	Present	Various, e.g. Center for Applied Research on the Environment and Sustainability (CARES) at AUC, RO operators, ministry of the environment, literature		5	docx, xlsx		Not yet available
Egypt	RO energy parameter	RO / saline water source (energy consumption per m^3)	RWTH	Local	Present	Various (e.g. CARES at AUC, RO operators, literature)		5	docx, xlsx		Not yet available
Egypt	Energy data	Photovoltaic (PV) data (technical and economic data: e.g. average	RWTH	Country	Present	Various (e.g. CARES at AUC, literature)		5	docx, xlsx		Not yet available

Location	Topic	Description	Reference partner	Spatial coverage	Temporal coverage	Source	Quality	Contributor (WP)	Data format	File name composed by	Notes
		energy generation at location x, cost)									
Egypt	Water balance marine aquaculture	Base data from marine aquacultures	RWTH	Local	Present	Various (e.g. literature, fisheries office, operators)		5	docx, xlsx		Not yet available
Egypt- Case study site	Water consumption	Estimations of the consumption per experiment/ system	ZG	Case study level	Based on the experiments' duration; measured on a weekly basis	AWESOME	To evaluate after experiment	CS	xlsx	Variable estimated, experiment number, specification of time	
Egypt- Case study site	Electricity consumption	Estimations of the consumption per experiment/ system	ZG	Case study level	Based on the experiments' duration; measured on a weekly basis	AWESOME	To evaluate after experiment	CS	xlsx	Variable estimated, experiment number, specification of time	

Location	Topic	Description	Reference partner	Spatial coverage	Temporal coverage	Source	Quality	Contributor (WP)	Data format	File name composed by	Notes
Egypt- Case study site	Environmental factors measurements (air and water Temperature, pH, DO, Relative humidity, EC, Lux)	Measurements of the environmental conditions of all experiment/ system	ZG	Case study level	Based on the experiments' duration; measured 4 times per day	AWESOME	To evaluate after experiment	CS	xlsx	Variable estimated, experiment number, specification of time	
Egypt- Case study site	Morphological measurements	Plant morphological data and of plant appearance data (fresh shoot system mass, fresh root system mass, number of leaves, stem diameter, stem length, head fresh weight, head diameter, head dry weight) for all experiment/ system outcomes	ZG	Case study level	Based on the experiments' duration (2-3 samples each time); measured on a weekly basis	AWESOME	To evaluate after experiment	CS	xlsx	Plant ID, experiment number, time	

Location	Topic	Description	Reference partner	Spatial coverage	Temporal coverage	Source	Quality	Contributor (WP)	Data format	File name composed by	Notes
Egypt- Case study site	Harvest gains	Fish and plant production per unit water	ZG + RWTH	Case study level	Based on the experiments' duration	AWESOME	To evaluate after experiment	5, CS	xlsx	Plant ID, experiment number, time, water expenditure, harvest gains	
Egypt- Case study site	Investment cost	The cost of purchasing and building the systems in relation to the production achieved	ZG + RWTH	Case study level	Based on the experiments' duration	AWESOME	To evaluate after experiment	5, CS	xlsx	System, experiment number, bill of quantities	
Egypt- Case study site	RO expenditure and outcome	How water use improves quantitatively if salty water is used in HP or AP before it goes into the RO	RWTH	Case study level	Based on the experiments' duration	AWESOME	To evaluate after experiment	5	xlsx	Discharge [m^3], salinity [ppt, ppm]	
Egypt- Case study site	PV expenditure and outcome	Quantification of reduced fossil energy consumption (electricity)	RWTH	Case study level	Based on the experiments' duration	AWESOME	To evaluate after experiment	5	xlsx	Electricity generation, total electricity consumption	

ANNEX 2: ADDITIONAL DATA REFERENCES

The references related to the metadata of Annex 1 are reported hereafter:

- Baum, Z., Palatnik, R. R., Kan, I., & Rapaport-Rom, M. (2016). Economic Impacts of Water Scarcity under Diverse Water Salinities. *Water Economics and Policy*, 2(1). https://doi.org/10.1142/S2382624X15500137
- Basheer, M., Wheeler, K. G., Elagib, N. A., Etichia, M., Zagona, E. A., Abdo, G. M., & Harou, J. J. (2020). Filling Africa's Largest Hydropower Dam Should Consider Engineering Realities. *One Earth*, 3(3), 277–281. https://doi.org/10.1016/j.oneear.2020.08.015
- BGR, & UNESCO. (2015). *The Global map of groundwater vulnerability to floods and droughts: explanatory notes - UNESCO Digital Library*. https://unesdoc.unesco.org/ark:/48223/pf0000232431
- Business Monitoring International-fitch solution, Retrieved October 22, 2020, from fithchsolutions.com
- Crespo Cuaresma, J. (2017). Income projections for climate change research: A framework based on human capital dynamics. *Global Environmental Change*, *42*, 226–236. https://doi.org/10.1016/j.gloenvcha.2015.02.012
- Dellink, R., Chateau, J., Lanzi, E., & Magné, B. (2017). Long-term economic growth projections in the Shared Socioeconomic Pathways. *Global Environmental Change*, *42*, 200–214. https://doi.org/10.1016/j.gloenvcha.2015.06.004
- Digna, R. F. M. O. (2020). Optimizing the Operation of a Multiple Reservoir System in the Eastern Nile Basin Considering Water and Sediment Fluxes [IHE Delft]. https://ihedelftrepository.contentdm.oclc.org/digital/collection/phd1/id/5457
- Egyptian petroleum cooperation, Egypt. Retrieved October 22, 2020, from egpc.com
- Fouré, J., Bénassy-Quéré, A., & Fontagné, L. (2013). Modelling the world economy at the 2050 horizon. *Economics of Transition*, *21*(4). https://doi.org/10.1111/ecot.12023
- Fouré, J., Quéré, A. B., & Fontagné, L. (2012). *CEPII - The Great Shift: Macroeconomic projections for the world economy at the 2050 horizon*. http://www.cepii.fr/CEPII/en/publications/wp/abstract.asp?NoDoc=4179
- Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., & Michaelsen, J. (2015). The climate hazards infrared precipitation with stations - A new environmental record for monitoring extremes. *Scientific Data*, *2*(1), 1–21. https://doi.org/10.1038/sdata.2015.66
- *Electricity statistics – Data services - IEA*. (n.d.). Retrieved October 22, 2020, from https://www.iea.org/subscribe-to-data-services/electricity-statistics
- IRENA. (2018). *Renewable Energy Outlook: Egypt*. International Renewable Energy Agency, Abu Dhab
- Harris, I., Jones, P. D., Osborn, T. J., & Lister, D. H. (2014). Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset. *International Journal of Climatology*, *34*(3), 623–642. https://doi.org/10.1002/joc.3711
- *Models, tools, and data - Models, Tools, and Data - IIASA*. (n.d.). Retrieved October 22, 2020, from https://iiasa.ac.at/web/home/research/modelsData/models-tools-data.html
- Jeuland, M., & Whittington, D. (2014). Water resources planning under climate change: Assessing the robustness of real options for the Blue Nile. *Water Resources Research*, 50(3), 2086–2107. https://doi.org/10.1002/2013WR013705

- Jeuland, M., Wu, X., & Whittington, D. (2017). Infrastructure development and the economics of cooperation in the Eastern Nile. *Water International*, 42(2), 121–141. https://doi.org/10.1080/02508060.2017.1278577
- Jiang, L., & O'Neill, B. C. (2017). Global urbanization projections for the Shared Socioeconomic Pathways. *Global Environmental Change*, *42*, 193–199. https://doi.org/10.1016/j.gloenvcha.2015.03.00
- KC, S., & Lutz, W. (2017). The human core of the shared socioeconomic pathways: Population scenarios by age, sex and level of education for all countries to 2100. *Global Environmental Change*, *42*, 181–192. https://doi.org/10.1016/j.gloenvcha.2014.06.004
- Lenzen, M., Kanemoto, K., Moran, D., & Geschke, A. (2012). Mapping the structure of the world economy. *Environmental Science and Technology*, *46*(15), 8374–8381. https://doi.org/10.1021/es300171x
- Lenzen, M., Moran, D., Kanemoto, K., & Geschke, A. (2013). Building EORA: a global multi-region input-output database at high country and sector resolution. *Economic Systems Research*, *25*(1), 20–49. https://doi.org/10.1080/09535314.2013.769938
- Liersch, S., Koch, H., & Hattermann, F. F. (2017). Management Scenarios of the Grand Ethiopian Renaissance Dam and Their Impacts under Recent and Future Climates. *Water*, 9(10), 728. https://doi.org/10.3390/w9100728
- NBI. (2016). *Nile Basin Water Resources Atlas*. ISBN: 978-9970-444-02-1
- Nachtergaele, F., Van Velthuizen, H., Verelst, L., Batjes, N., Dijkshoorn, K., Van Engelen, V., Fischer, G., Jones, A., Montanarella, L., Petri, M., Prieler, S., Teixeira, E., Wiberg, D., & Shi, X. (2009). *Harmonized World Soil Database - Version 1.1*.
- New, M., Lister, D., Hulme, M., & Makin, I. (2002). A high-resolution data set of surface climate over global land areas. *Climate Research*, *21*(1), 1–25. https://doi.org/10.3354/cr021001
- Palatnik, R. R. (2019). The Economic Value of Seawater Desalination—The Case of Israel. In *Economy-Wide Modeling of Water at Regional and Global Scales* (pp. 193–208). Springer, Singapore. https://doi.org/10.1007/978-981-13-6101-2_9
- Portmann, F. T., Siebert, S., & Döll, P. (2010). MIRCA2000-Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling. *Global Biogeochemical Cycles*, *24*(1). https://doi.org/10.1029/2008gb003435
- Rady, Y. Y., Rocco, M. V., Serag-Eldin, M. A., & Colombo, E. (2018). Modelling for power generation sector in Developing Countries: Case of Egypt. *Energy*, *165*, 198–209. https://doi.org/10.1016/j.energy.2018.09.089
- Siebert, S., & Döll, P. (2010). Quantifying blue and green virtual water contents in global crop production as well as potential production losses without irrigation. *Journal of Hydrology*, *384*(3–4), 198–217. https://doi.org/10.1016/j.jhydrol.2009.07.031
- Van der Krogt, W. N. M., & Ogink, I. H. J. M. (2013). Development of the Eastern Nile Water Simulation Model Main Report
- Wheeler, K. G., Basheer, M., Mekonnen, Z. T., Eltoum, S. O., Mersha, A., Abdo, G. M., Zagona, E. A., Hall, J. W., & Dadson, S. J. (2016). Cooperative filling approaches for the Grand Ethiopian Renaissance Dam. Water International, 41(4), 611–634.https://doi.org/10.1080/02508060.2016.1177698
- Wheeler, K. G., Hall, J. W., Abdo, G. M., Dadson, S. J., Kasprzyk, J. R., Smith, R., & Zagona, E. A. (2018). Exploring Cooperative Transboundary River Management Strategies for the Eastern Nile Basin. Water Resources Research, 54(11), 9224–9254. https://doi.org/10.1029/2017WR022149