



Research and Innovation action

H2020-SC5-2017

Data Management Plan (DMP) update

Deliverable D5.3

Version N°1.1

Authors: BSC, CapGemini



This project received funding from the Horizon 2020 programme under the grant agreement n°776787.

Disclaimer

The content of this deliverable reflects only the author's view. The European Commission is not responsible for any use that may be made of the information it contains.

Document Information

Grant Agreement	776787
Project Title	Sub-seasonal to Seasonal climate forecasting for Energy
Project Acronym	S2S4E
Project Start Date	01/12/2017
Related work package	WP 5 - Operational Climate Service: Decision Support Tool (DST) implementation
Related task(s)	Task 5.1 Data Management Plan - definition of data protocols and formats
Lead Organisation	BSC
Submission date	31/05/2019
Dissemination Level	PU

History

Date	Submitted by	Reviewed by	Version (Notes)
22/05/2019	BSC	CAPGEMINI, BSC	1.0 First draft
30/05/2019	BSC	CAPGEMINI	1.1 Final version

Table of content

Table of content.....	4
Summary.....	5
Keywords.....	5
1 Introduction.....	6
1.1 Purpose of the document.....	6
1.2 Applicable and reference documents.....	6
The applicable documents are listed in the table below:.....	6
2 Data collection and storage	6
2.1 Type of data collected.....	6
2.2 Data collection method and results.....	6
3 Data sharing	7

Summary

This document is the second revision of the Data Management Plan. The first version was delivered in *D5.2 Data Management Plan (DMP) presentation* at M6, the third one is due at M36 in *D5.4. Data Management Plan (DMP) project*. The purpose of this document is to present the management of data in the scope of the S2S4E project.

Keywords

Data, privacy, preservation, open access, storage

About S2S4E

The project seeks to improve renewable energy variability management by developing a tool that for the first time integrates sub-seasonal to seasonal climate predictions with renewable energy production and electricity demand.

Our long-term goal is to make the European energy sector more resilient to climate variability and extreme events.

Large-scale deployment of renewable energy is key to comply with the emissions reductions agreed upon in the Paris Agreement. However, despite being cost competitive in many settings, renewable energy diffusion remains limited largely due to seasonal variability. Knowledge of power output and demand forecasting beyond a few days remains poor, creating a major barrier to renewable energy integration in electricity networks.

To help solve this problem, S2S4E is developing an innovative service to improve renewable energy variability management. The outcome will be new research methods exploring the frontiers of weather conditions for future weeks and months and a decision support tool for the renewable industry.

More information:

www.s2s4e.eu

1 Introduction

1.1 Purpose of the document

The objective of this document is to review the policies and technical solutions about the collected and generated data during the project, across the different work packages, the data being scientific or not (personal data, results from surveys, etc.). In the following pages, we only consider new developments or adjustments compared with the first version of the DMP (D5.2, M6). The general definitions and policies are in the previous deliverable.

1.2 Applicable and reference documents

The applicable documents are listed in the table below:

Id	Deliverables	WP
D5.2	Data Management Plan	WP5
D8.1	POPD – Requirement No. 1: Information sheet of users	WP8
D8.2	POPD – Requirement No.2: Guideline for personal data management	WP8

2 Data collection and storage

2.1 Type of data collected

As described in the previous version of the DMP (deliverable D5.2), the data collected during the project will be of two types: scientific data and personal data. The personal data corresponds to the logins to the DST and the information provided by the users during the surveys conducted by *WP2 Definition of user needs and the role of S2S forecasts in decision-making processes*. Regarding the Decision Support Tool (DST) users' data, it is foreseen to collect them after the launching of the DST following the data protection policy.

2.2 Data collection method and results

During the reporting period, only scientific data has been collected, using ssh connections and the web api from the European Centre for Medium-Range Weather Forecasts (ECMWF).

So far, around 4TB of data has been collected from the following datasets: era5 reanalysis, ECMWF system5 C3S, NCEP CFS v2 reanalysis, for variables surface temperature, minimum and maximum temperature, sea level pressure, surface wind module, precipitation, solar surface radiation. The period downloaded is from 1993 to 2019 (ECMWF System5), 1999 to 2019 (NCEP) and 1993 to 2018 (era5). All files are downloaded at the highest temporal frequency available (1hourly for era5, 6hourly for the other systems) and weekly and monthly means are computed from them to produce the final data ingested by the DST.

All this data is kept in the GPFS Native RAID storage at BSC described in the previous deliverable D5.2.

3 Data sharing

For easy exchange of small temporary scientific data between partners, EUDAT B2DROP has also been used. This allows easy access to the data, and ensures the long term availability of the data without having to worry about giving access to the storage of the partners.

The automatic download of the hydrological data from SMHI to the BSC archive for further ingestion in the DST has now been set up and will be done through an ftp server. The expected data size is and the data will be downloaded once a week for the subseasonal and once a month for the seasonal. The subseasonal will represent 100MB per week and the monthly data 100MB per month.