



Open Sea Operating Experience to Reduce Wave Energy Costs

Deliverable D8.5

Data Management Plan (first version)

Lead Beneficiary	TECNALIA
Delivery date	2016-07-27
Dissemination level	Public
Status	Approved
Version	1.0
Keywords	Open access, datasets, metadata, ZENODO



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654444

Disclaimer

This Deliverable reflects only the author's views and the Agency is not responsible for any use that may be made of the information contained therein

Document Information

Grant Agreement Number	654444
Project Acronym	OPERA
Work Package	WP 8
Task(s)	T8.2
Deliverable	D8.5
Title	Data Management Plan (first version)
Author(s)	Pablo Ruiz-Minguela (TECNALIA), Endika Aldaiturriaga (OCEANTEC)
File Name	OPERA_D8.5_Data management plan_TECNALIA_20160727_v1.0.docx

Change Record

Revision	Date	Description	Reviewer
0.0	14-04-2016	Initial outline	Pablo Ruiz-Minguela
0.5	21-06-2016	Full draft contents	WP8 partners
0.9	15-07-2016	Version for peer review	Sara Armstrong (UCC)
1.0	27-07-2016	Final deliverable to EC	EC

EXECUTIVE SUMMARY

This document, D8.5 Data Management Plan (DMP) is a deliverable of the OPERA project, which is funded by the European Union's H2020 Programme under Grant Agreement #654444. OPERA's main goal is to collect, analyse and share open sea operating data and experience to validate and de-risk four industrial innovations for wave energy opening the way to long term cost reduction of over 50%.

The Consortium strongly believes in the concepts of open science, and in the benefits that the European innovation ecosystem and economy can draw from allowing the reuse of data at a larger scale. Besides, there is a need to gather experience in open sea operating conditions, structural and power performance, and operating data in wave energy. Therefore, this project proposes to remove this roadblock by delivering for the first time, open access, high-quality open sea operating data to the wave energy development community.

The OPERA project participates in the Pilot on Open Research Data launched by the European Commission along with the H2020 programme. The use of a Data Management Plan is required for all participating projects.

The purpose of the Data Management Plan (DMP) is to provide an analysis of the main elements of the data management policy that will be used by the Consortium with regard to the project research data.

The DMP covers the complete research data life cycle. It describes the types of research data that will be generated or collected during the project, the standards that will be used, how the research data will be preserved and what parts of the datasets will be shared for verification or reuse. It also reflects the current state of the Consortium agreements on data management and must be consistent with exploitation and IPR requirements.

Research data linked to exploitable results will not be put into the open domain if they compromise its commercialisation prospects or have inadequate protection, which is a H2020 obligation. The rest of research data will be deposited in an open access repository.

The DMP is not a fixed document; on the contrary it will evolve during the lifespan of the project. This first version of the DMP includes an overview of the datasets to be produced by the project, and the specific conditions that are attached to them. The next versions of the DMP will get into more detail and describe the practical data management procedures implemented by the OPERA project.

The expected types of research data that will be collected or generated along the project lie in the following categories: 1) Environmental monitoring; 2) Mooring performance; 3) Biradial performance; 4) Power output; 5) Power quality; and 5) Offshore operations.



TABLE OF CONTENTS

EXECUTIVE SUMMARY	3
TABLE OF CONTENTS.....	4
LIST OF FIGURES	6
LIST OF TABLES	6
ABBREVIATIONS AND ACRONYMS	7
1. INTRODUCTION	8
1.1 OPERA motivation	8
1.2 Purpose of the Data Management Plan.....	9
1.3 Research data types in OPERA	10
1.4 Responsibilities.....	11
2. ENVIRONMENTAL MONITORING	13
2.1 Dataset reference and name.....	13
2.2 Dataset description	13
2.3 Standards and metadata	14
2.4 Data sharing	14
2.5 Archiving and preservation	15
3. MOORING PERFORMANCE.....	16
3.1 Dataset reference and name.....	16
3.2 Dataset description	16
3.3 Standards and metadata	16
3.4 Data sharing	17
3.5 Archiving and preservation	17
4. BIRADIAL TURBINE PERFORMANCE	18
4.1 Dataset reference and name.....	18
4.2 Dataset description	18
4.3 Standards and metadata	19
4.4 Data sharing	19
4.5 Archiving and preservation	19
5. POWER OUTPUT.....	20
5.1 Dataset reference and name.....	20



5.2 Dataset description	20
5.3 Standards and metadata	20
5.4 Data sharing	21
5.5 Archiving and preservation	21
6. POWER QUALITY	22
6.1 Dataset reference and name.....	22
6.2 Dataset description	22
6.3 Standards and metadata	22
6.4 Data sharing	23
6.5 Archiving and preservation	23
7. OFFSHORE OPERATIONS	24
7.1 Dataset reference and name.....	24
7.2 Dataset description	24
7.3 Standards and metadata	24
7.4 Data sharing	25
7.5 Archiving and preservation	25
8. REFERENCES	26
ANNEX I: KEY PRINCIPLES FOR OPEN ACCESS TO RESEARCH DATA.....	28
ANNEX II: SCIENTIFIC PUBLICATIONS	29

LIST OF FIGURES

Figure 1.1: Research data life cycle (Adapted from UK Data Archive [1])	9
Figure 1.2: Research data options and timing	10

LIST OF TABLES

Table 1.1: OPERA types of data.....	10
-------------------------------------	----

ABBREVIATIONS AND ACRONYMS

DMP	Data Management Plan
DoA	Description of Action
DOF	Degree of Freedom
EC	European Commission
ER1	Exploitable Result 1 (Floating OWC device and shared mooring configuration)
ER2	Exploitable Result 2 (PTO based on the biradial air turbine)
ER3	Exploitable Result 3 (Elastomeric tether for mooring systems)
ER4	Exploitable Result 4 (Advanced control algorithms for WECs)
ER5	Exploitable Result 5 (Open-sea testing and validation services)
ER6	Exploitable Result 6 (Guidelines and standards for ocean energy)
ER7	Exploitable Result 7 (Engineering and construction of wave energy farms)
ER8	Exploitable Result 8 (Seakeeping, installation and O&M services)
Hs	Significant wave height
Tp	Peak wave period
WP	Work Package



1. INTRODUCTION

1.1 OPERA MOTIVATION

The OPERA project participates in the Pilot on Open Research Data launched by the European Commission (EC) along with the H2020 programme. This pilot is part of the Open Access to Scientific Publications and Research Data programme in H2020. The goal of the programme is to foster access to research data generated in H2020 projects. The use of a Data Management Plan (DMP) is required for all projects participating in the Open Research Data Pilot.

Open access is defined as the practice of providing on-line access to scientific information that is free of charge to the reader and that is reusable. In the context of research and innovation, scientific information can refer to peer-reviewed scientific research articles or research data.

Research data refers to information, in particular facts or numbers, collected to be examined and considered and as a basis for reasoning, discussion, or calculation. In a research context, examples of data include statistics, results of experiments, measurements, observations resulting from fieldwork, survey results, interview recordings and images. The focus is on research data that is available in digital form.

The Consortium strongly believes in the concepts of open science, and in the benefits that the European innovation ecosystem and economy can draw from allowing the reuse of data at a larger scale.

Furthermore, there is a need to gather experience in open sea operating conditions, structural and power performance and operating data in wave energy. In fact, there has been very limited open sea experience in wave energy, which is essential in order to fully understand the challenges in device performance, survivability and reliability. The limited operating data and experience that currently exists are rarely shared, as testing is partly private-sponsored.

This project proposes to remove this roadblock by delivering for the first time, open access, high-quality open sea operating data to the wave energy development community.

Nevertheless, data sharing in the open domain can be restricted as a legitimate reason to protect results that can reasonably be expected to be commercially or industrially exploited. Strategies to limit such restrictions will include anonymising or aggregating data, agreeing on a limited embargo period or publishing selected datasets.



1.2 PURPOSE OF THE DATA MANAGEMENT PLAN

The purpose of the DMP is to provide an analysis of the main elements of the data management policy that will be used by the Consortium with regard to the project research data.

The DMP covers the complete research data life cycle. It describes the types of research data that will be generated or collected during the project, the standards that will be used, how the research data will be preserved and what parts of the datasets will be shared for verification or reuse. It also reflects the current state of the Consortium agreements on data management and must be consistent with exploitation and IPR requirements.

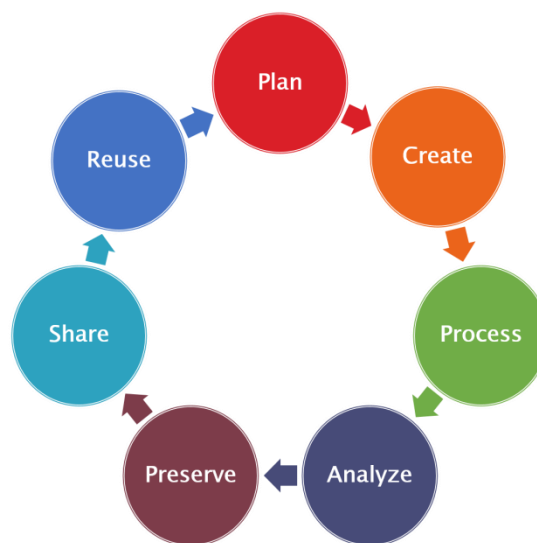


FIGURE 1.1: RESEARCH DATA LIFE CYCLE (ADAPTED FROM UK DATA ARCHIVE [1])

The DMP is not a fixed document, but will evolve during the lifespan of the project, particularly whenever significant changes arise such as dataset updates or changes in Consortium policies.

This document is the first version of the DMP, delivered in Month 6 of the project. It includes an overview of the datasets to be produced by the project, and the specific conditions that are attached to them. The next versions of the DMP will get into more detail and describe the practical data management procedures implemented by the OPERA project with reference with the IT tools developed in WP1. At a minimum, the DMP will be updated in Month 18 (D8.6) and Month 30 (D8.7) respectively.

This document has been produced following the EC guidelines for project participating in this pilot and additional consideration described in ANNEX I: KEY PRINCIPLES FOR OPEN ACCESS TO RESEARCH DATA.

1.3 RESEARCH DATA TYPES IN OPERA

For this first release of DMP, the data types that will be produced during the project are focused on the Description of the Action (DoA) and on the results obtained in the first months of the project.

According to such consideration, Table 1.1 reports a list of indicative types of research data that OPERA will produce. These research data types have been mainly defined in WP1, including data structures, sampling and processing requirements, as well as relevant standards. This list may be adapted with the addition or removal of datasets in the next versions of the DMP to take into consideration the project developments. A detailed description of each dataset is given in the following sections of this document.

TABLE 1.1: OPERA TYPES OF DATA

#	Dataset	Lead partner	Related WP(s)
1	Environmental monitoring	TECNALIA	WP1
2	Mooring performance	UNEXE	WP1, WP2, WP5
3	Biradial performance	IST	WP1, WP3
4	Power output	OCEANTEC	WP1, WP4, WP5
5	Power quality	UCC	WP1, WP5
6	Offshore operations	TECNALIA	WP6

Specific datasets may be associated to scientific publications (i.e. underlying data), public project reports and other raw data or curated data not directly attributable to a publication. The policy for open access are summarised in the following picture.

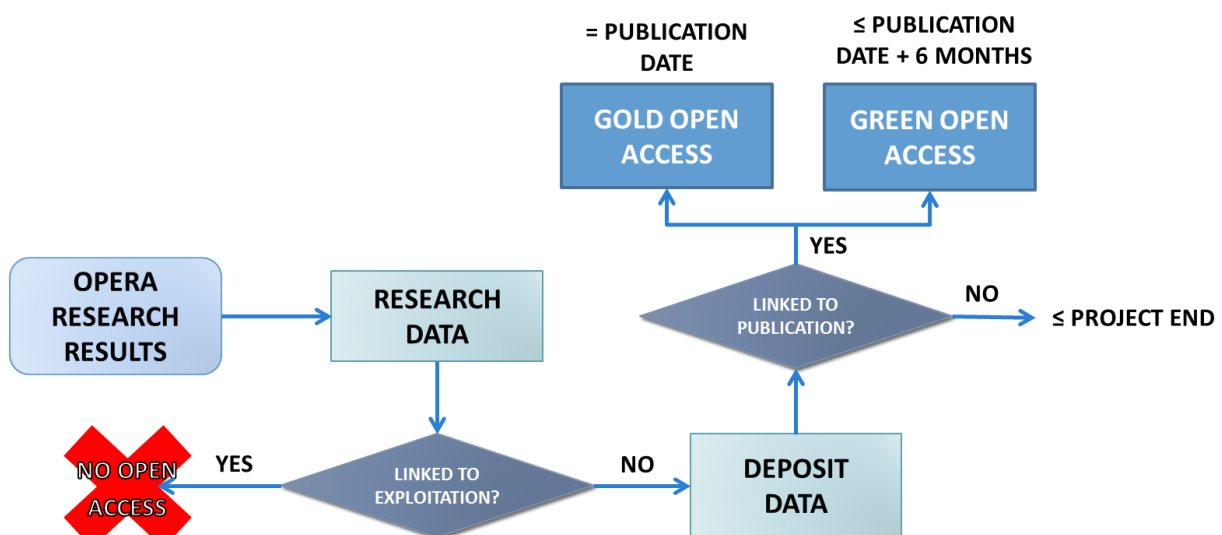


FIGURE 1.2: RESEARCH DATA OPTIONS AND TIMING

Research data linked to exploitable results will not be put into the open domain if they compromise its commercialisation prospects or have inadequate protection, which is a H2020 obligation. The rest of research data will be deposited in an open access repository.

When the research data is linked to a scientific publication, the provisions described in ANNEX II: SCIENTIFIC PUBLICATIONS will be followed. Research data needed to validate the results presented in the publication should be deposited at the same time for “Gold” Open Access¹ or before the end of the embargo period for “Green” Open Access². Underlying research data will consist of selected parts of the general datasets generated, and for which the decision of making that part public has been made.

Other datasets will be related to any public report or be useful for the research community. They will be selected parts of the general datasets generated or full datasets (i.e. up to 2 years of key operating data), and be published as soon as possible.

1.4 RESPONSIBILITIES

Each OPERA partner has to respect the policies set out in this DMP. Datasets have to be created, managed and stored appropriately and in line with applicable legislation.

The Project Coordinator has a particular responsibility to ensure that data shared through the OPERA website are easily available, but also that backups are performed and that proprietary data are secured.

OCEANTEC, as WP1 leader, will ensure dataset integrity and compatibility for its use during the project lifetime by different partners.

Validation and registration of datasets and metadata is the responsibility of the partner that generates the data in the WP. Metadata constitutes an underlying definition or description of the datasets, and facilitate finding and working with particular instances of data.

Backing up data for sharing through open access repositories is the responsibility of the partner possessing the data.

Quality control of these data is the responsibility of the relevant WP leader, supported by the Project Coordinator.

¹ **“Gold” Open Access:** Authors make a one-off payment to the publisher so that the scientific publication is immediately published in open access mode.

² **“Green” Open Access:** Due to the contractual conditions of the publisher, the scientific publication can undergo an embargo period up to six months since publication date before the author can deposit the published article or the final peer-reviewed manuscript in open access mode.



If datasets are updated, the partner that possesses the data has the responsibility to manage the different versions and to make sure that the latest version is available in the case of publically available data. WP1 will provide naming and version conventions.

Last but not least, all partners must consult the concerned partner(s) before publishing data in the open domain that can be associated to an exploitable result.

2. ENVIRONMENTAL MONITORING

2.1 DATASET REFERENCE AND NAME

DS_Environmental_Monitoring

2.2 DATASET DESCRIPTION

The DS_Environmental_Monitoring datasets mainly consists of several wave parameters (such as wave Hs, Tp, direction and spreading), but may also include wind, tide, current and temperature parameters. These numeric datasets will be directly obtained through observations, and derived using statistical parameters and models. In the latter, specialist software may be needed for further processing of data.

Environmental monitoring data will be collected at two locations, namely the Mutriku shoreline plant and the open sea test site BiMEP.

Currently at Mutriku, the single environmental parameter collected is the wave elevation in the Oscillating Water Column. In order to increase the characterisation of the wave resource, a new wave instrument will be installed about 300 m off the shoreline for approximately 6 months. It will be of the bottom mounted pressure gauge type.

The BiMEP reference buoy is the main wave instrument for the test site Fugro-OCEANOR Wavescan buoy, deployed in March 2009, recording almost continuously except for a significant gap from January 2013 to April 2014. This wave instrument records 17 minutes heave, pitch and roll time series, from which omnidirectional and directional spectra can be estimated, as well as standard sea-state parameters. It is located at about 1.1 km to the WSW of the prototype deployment. Additionally, a surface following buoy will be installed close to the prototype for the research activities of the project at BiMEP for approximately two years.

In general, environmental monitoring datasets will be useful for further research activities beyond the scope of OPERA objectives. Metocean observations are common practice for different uses. Dataset could be integrated and reused, particularly for the characterisation of wave resource and the estimation of device performance. They will be also valuable for technology developers who plan to test their devices at either Mutriku or BiMEP.

Although the raw datasets are useful by themselves, it is the objective of the OPERA project to use the dataset as a basis for at least one scientific publication.



2.3 STANDARDS AND METADATA

There have been many discussions for processing data and information in oceanography. Many useful ideas have been developed and put into practice, but there have been few successful attempts to develop and implement international standards in managing data.

The Ocean Data Standards Project [2] contains an extensive number of references on Oceanographic Data Management and Exchange Standards. It includes references on Metadata, Date and Time, Lat/Lon/Alt, Country names, Platform instances, Platform types, Science Words, Instruments, Units, Projects, Institutions, Parameters, Quality Assurance and Quality Control.

The ISO 19156:2011 defines a conceptual schema for observations, and for features involved in sampling when making observations. These provide models for the exchange of information describing observation acts and their results, both within and between different scientific and technical communities.

Additionally, regarding the wave energy application, the relevant standard is the technical specification for wave energy resource assessment and characterization IEC TS 62600-101 [4].

The environmental monitoring system will be integrated in the existing IT infrastructure at Mutriku and BiMEP. A SCADA system will be developed that allows partners to access monitoring information locally and remotely. TECNALIA will be responsible for version control and validation dataset of datasets to be shared open access.

2.4 DATA SHARING

During the lifecycle of the OPERA project datasets will be stored and systematically organised in a database tailored to comply with the requirements of WP1 (for more details on the database architecture, please see D1.1 Process instrumentation definition). An online data query tool will be operational by Month 12 and for open dissemination by Month 18. The database schema and the queryable fields, will be also publicly available to the database users as a way to better understand the database itself.

In addition to the project database, relevant datasets will be also stored in ZENODO [5], which is the open access repository of the Open Access Infrastructure for Research in Europe, OpenAIRE [6].

Data access policy will be unrestricted since no confidentiality or IPR issues are expected regarding the environmental monitoring datasets. All collected datasets will be disseminated



without an embargo period unless linked to a green open access publication. Data objects will be deposited in ZENODO under:

- Open access to data files and metadata and data files provided over standard protocols such as HTTP and OAI-PMH.
- Use and reuse of data permitted.
- Privacy of its users protected.

2.5 ARCHIVING AND PRESERVATION

The OPERA project database will be designed to remain operational for 5 years after project end. By the end of the project, the final dataset will be transferred to the ZENODO repository, which ensures sustainable archiving of the final research data.

Items deposited in ZENODO will be retained for the lifetime of the repository, which is currently the lifetime of the host laboratory CERN and has an experimental programme defined for the at least next 20 years. Data files and metadata are backed up on a nightly basis, as well as replicated in multiple copies in the online system. All data files are stored along with a MD5 checksum of the file content. Regular checks of files against their checksums are made.

3. MOORING PERFORMANCE

3.1 DATASET REFERENCE AND NAME

DS_Mooring_Performance

3.2 DATASET DESCRIPTION

The DS_Mooring_Performance datasets consists of extreme loads and motion response (6 DOF) to different sea states for the mooring lines. Mooring performance data will be both experimental and observational, raw and derived (using statistical parameters and models).

Experimental data will be collected at the DMAC facility in UNEXE [7] . These tests will be focused on the characterisation of design load behaviour, fatigue and durability of several elastomeric tether specimens. Raw data will be used. Selected parts of the experimental datasets generated will be made public.

Field tests will be conducted at the open sea test site at BiMEP. A mooring condition monitoring will be implemented for the project consisting of 4 load shackles deployed in two mooring nodes of the prototype. Extreme loads and motion response to different sea states will be monitored. The loading data will be combined with the environmental monitoring dataset to derive the final mooring performance dataset. Selected parts of the field test datasets generated will be made public.

The mooring performance dataset will be useful to inform technology compliance, survivability and reliability as well as economical improvements. They will be also valuable for the certification processes of other technology developers.

This dataset will be the basis for at least one scientific publication.

3.3 STANDARDS AND METADATA

In order to ensure the required compatibility, this dataset will use the same ocean data standards than the previous environmental monitoring dataset for data and metadata capture/creation.

The offshore standard DNVGL-OS-E301 [8] contains criteria, technical requirements and guidelines on design and construction of position mooring systems. The objective of this standard shall give a uniform level of safety for mooring systems, consisting of chain, steel wire ropes and fibre rope. Besides, regarding the wave energy application, the relevant standard is the technical specification for assessment of mooring system for Marine Energy Converters (MECs) IEC TS 62600-10 [9] .



During the OPERA project, a SCADA system will be developed that allows partners to access monitoring information locally and remotely. UNEXE will be responsible for version control and validation dataset of datasets to be shared open access.

3.4 DATA SHARING

As it has been described before, during the lifecycle of the OPERA project datasets will be stored and systematically organised in a database tailored to comply with the requirements of WP1. An online data query tool will be operational by Month 12 and for open dissemination by Month 18. The database schema and the queryable fields will be also publicly available to the database users as a way to better understand the database itself.

Full data access policy will be restricted to WP2 participants, in order to protect the commercial and industrial prospects of exploitable results (ER1 and ER3). However, aggregated data will be used in order to limit this restriction.

The aggregated dataset will be disseminated as soon as possible. In the case of the underlying data of a publication this might imply an embargo period for green open access publications.

Data objects will be deposited in ZENODO under open access to data files and metadata, permitting its use and reuse, as well as protecting privacy of its users.

3.5 ARCHIVING AND PRESERVATION

As it has been described before, the OPERA project database will be designed to remain operational for 5 years after project end. By the end of the project, the final dataset will be transferred to the ZENODO repository, which ensures sustainable archiving of the final research data.

4. BIRADIAL TURBINE PERFORMANCE

4.1 DATASET REFERENCE AND NAME

DS_Biradial_Turbine_Performance

4.2 DATASET DESCRIPTION

The DS_Biradial_Turbine_Performance datasets mainly consists of internal water level, chamber pressure/temperature/humidity, rotation speed and torque to assess turbine efficiency in response to different sea states. Biradial turbine performance data will be both experimental and observational, raw and derived (using statistical parameters and models).

Experimental data will be collected at existing rig in IST Turbomachinery Laboratory for tests in varying unidirectional flow. Built-in sensors will measure rpm, pressure differential across rotor, vibration and generator temperature, voltage and current. Raw data will be used. Selected parts of the experimental datasets generated will be made public.

Field tests will be conducted both at Mutriku shoreline plant and the BiMEP open sea test site.

Testing at Mutriku will assess turbine performance and collect extensive data on drivers of components fatigue such as high rpm and accelerations; electrical, temperature and pressure load cycles; humidity in the cabinet (which exacerbates electrical stress damages); rate of salt accumulation and corrosion.

Similar data will be collected at BiMEP and results will be compared. Additionally, low-frequency accelerometers will assess loads on the rotor and bearings.

The loading data will be combined with the environmental monitoring dataset to derive the final biradial turbine performance dataset. Non-dimensional values, aggregated data and selected parts of the field test datasets generated will be made public.

The biradial turbine performance dataset will be useful to assess turbine efficiency and reliability.

This dataset will be the basis for at least one scientific publication.



4.3 STANDARDS AND METADATA

In order to ensure the required compatibility, this dataset will use the same ocean data standards than the previous environmental monitoring dataset for data and metadata capture/creation.

DNV GL will advise on applicable rules and standards to ensure appropriate design and data capture for open ocean operating conditions.

During the OPERA project, a SCADA system will be developed that allows partners to access monitoring information locally and remotely. IST will be responsible for version control and validation dataset of datasets to be shared open access.

4.4 DATA SHARING

OPERA project datasets will be stored and systematically organised in a database tailored to comply with the requirements of WP1. An online data query tool will be operational by Month 12 and for open dissemination by Month 18. The database schema and the queryable fields, will be also publicly available to the database users as a way to better understand the database itself.

Full data access policy will be restricted to WP3 participants, in order to protect the commercial and industrial prospects of exploitable results (ER1 and ER2). However, aggregated data will be used in order to limit this restriction.

The aggregated dataset will be disseminated as soon as possible. In the case of the underlying data of a publication this might imply an embargo period for green open access publications.

Data objects will be deposited in ZENODO under open access to data files and metadata, permitting its use and reuse, as well as protecting privacy of its users.

4.5 ARCHIVING AND PRESERVATION

As it has been described before, the OPERA project database will be designed to remain operational for 5 years after project end. By the end of the project, the final dataset will be transferred to the ZENODO repository, which ensures sustainable archiving of the final research data.



5. POWER OUTPUT

5.1 DATASET REFERENCE AND NAME

DS_Power_Output

5.2 DATASET DESCRIPTION

The DS_Power_Output datasets mainly consists of generator speed, voltage, frequency and electric power. Power output data will be both experimental and observational, raw and derived such as mean, standard deviation, minimum and maximum values.

Experimental data will be collected at electrical test rigs of UCC [10] and TECNALIA [11] . Field tests data will be collected at Mutriku shoreline plant and at the BiMEP open sea test site. Numerical models will be also used to extend the dataset beyond sea-trials data. In the latter, specialist software may be needed for further processing the data. Selected parts of the generated datasets generated will be made public.

Power output data will be useful to identify sources of uncertainty in power performance prediction. They will be also valuable for the certification processes of other technology developers.

This dataset will be the basis for at least one scientific publication.

5.3 STANDARDS AND METADATA

In order to ensure the required compatibility, this dataset will use the same ocean data standards than the previous environmental monitoring dataset for data and metadata capture/creation.

Additionally, regarding the wave energy application, the relevant standards are the technical specification on power performance assessment of electricity producing wave energy converters IEC TS 62600-100 [12] , and the technical specification on wave energy converter power performance assessment at a second location using measured assessment data IEC TS 62600-102 [13] .

As indicated in the technical specifications, the datasets shall provide a record of sea state and electrical power production over time. Each aggregated data record shall be date and time stamped using ISO 8601. The records shall be annotated with quality control flags giving the results of the quality control checks carried out during recording and analysis.



A SCADA system will be developed that allows partners to access monitoring information locally and remotely. OCEANTEC will be responsible for version control and validation dataset of datasets to be shared open access.

5.4 DATA SHARING

During the lifecycle of the OPERA project datasets will be stored and systematically organised in a database tailored to comply with the requirements of WP1. An online data query tool will be operational by Month 12 and for open dissemination by Month 18. The database schema and the queryable fields, will be also publicly available to the database users as a way to better understand the database itself.

Full data access policy will be restricted to WP4 and WP5 participants, in order to protect the commercial and industrial prospects of exploitable results (ER1, ER4 and ER6). However, aggregated data will be used in order to limit this restriction.

The aggregated dataset will be disseminated as soon as possible. In the case of the underlying data of a publication this might imply an embargo period for green open access publications.

Data objects will be deposited in ZENODO under open access to data files and metadata, permitting its use and reuse, as well as protecting privacy of its users.

5.5 ARCHIVING AND PRESERVATION

As it has been described before, the OPERA project database will be designed to remain operational for 5 years after project end. By the end of the project, the final dataset will be transferred to the ZENODO repository, which ensures sustainable archiving of the final research data.

6. POWER QUALITY

6.1 DATASET REFERENCE AND NAME

DS_Power_Quality

6.2 DATASET DESCRIPTION

The DS_Power_Quality datasets consists of current, voltage, power quality characteristic parameters (such as voltage fluctuations, harmonics, inter-harmonics, active/reactive power, and flicker). Power quality data will be both experimental and observational, raw and derived (using statistical parameters and models).

Experimental data will be collected at electrical test rig of UCC [10]. Field tests data will be collected at the Mutriku shoreline plant. Simulated models may be used to assess the power quality for other operating conditions, such as varying control algorithms, resource conditions, grid strengths, and control using a dry-lab to create a wider profile for the WEC. In the latter, specialist software may be needed for further processing the data. Selected parts of the experimental datasets generated will be made public.

In Mutriku, data will be collected from both a single turbine and the plant as a whole, obtaining valuable conclusions about how aggregation of multiple turbines affects the power quality. Non-dimensional values, aggregated data and selected parts of the Mutriku field test datasets generated will be made public.

Power quality data will be useful to identify sources of uncertainty in power performance prediction. They will be also valuable for the certification processes of other technology developers.

This dataset will be the basis for at least one scientific publication.

6.3 STANDARDS AND METADATA

In order to ensure the required compatibility, this dataset will use the same ocean data standards than the previous environmental monitoring dataset for data and metadata capture/creation.

Additionally, regarding the wave energy application, the relevant standards are the technical specification on electrical power quality requirements for wave, tidal and other water current energy converters IEC TS 62600-30 [14].



Further instructions on processing harmonic current components are given in IEC 61000-4-7:2002 [15], for power supply systems and equipment connected thereto.

A SCADA system will be developed that allows partners to access monitoring information locally and remotely. UCC will be responsible for version control and validation dataset of datasets to be shared open access. Additional attention must be given to integrating and combining the power quality datasets with others due to the largely varying timescales.

6.4 DATA SHARING

Datasets will be stored and systematically organised in a database tailored to comply with the requirements of WP1. An online data query tool will be operational by Month 12 and for open dissemination by Month 18. The database schema and the queryable fields, will be also publicly available to the database users as a way to better understand the database itself.

Full data access policy will be restricted to WP5 participants, in order to protect the commercial and industrial prospects of exploitable results (ER1, ER4 and ER6). However, aggregated data will be used in order to limit this restriction.

The aggregated dataset will be disseminated as soon as possible. In the case of the underlying data of a publication this might imply an embargo period for green open access publications.

Data objects will be deposited in ZENODO under open access to data files and metadata, permitting its use and reuse, as well as protecting privacy of its users.

6.5 ARCHIVING AND PRESERVATION

The OPERA project database will be designed to remain operational for 5 years after project end. By the end of the project, the final dataset will be transferred to the ZENODO repository, which ensures sustainable archiving of the final research data.

7. OFFSHORE OPERATIONS

7.1 DATASET REFERENCE AND NAME

DS_Offshore_Operations

7.2 DATASET DESCRIPTION

The DS_Offshore_Operations datasets consists of failures, type of maintenance, offshore resources (such as vessels, equipment, personnel, parts and consumables), health & safety, and activity log.

Offshore Operations data will be observational and derived.

Field tests will be conducted at the BiMEP open sea test site. The offshore operations data will be combined with the environmental monitoring dataset to derive the final dataset. Full datasets will be made public.

Offshore operations data will be useful to reduce the uncertainty on the determination of risk and cost of offshore operations, and to optimise these activities. The offshore logistics experience can be extrapolated to different scenarios of larger deployment with a view to more accurately assess the economies of scale and identify logistics bottlenecks when deployed in large arrays.

Although the raw datasets are useful by themselves, it is the objective of the OPERA project to use the dataset as a basis for at least one scientific publication.

7.3 STANDARDS AND METADATA

Unlike the previous datasets, these are not based on process instrumentation and therefore will not be stored in the WP1 database. This dataset can be imported from, and exported to a CSV, TXT or Excel file.

Failure data will be reported according to Failure the Reporting, Analysis and Corrective Action System (FRACAS) [16] and the ISO 14224:2006 Collection and exchange of reliability and maintenance data for equipment [17] .

The DataCite Metadata Schema [18] will be used for publication of the offshore operations datasets. DataCite is a domain-agnostic list of core metadata properties chosen for the accurate and consistent identification of data for citation and retrieval purposes.



TECNALIA will be responsible for version control and validation dataset of datasets to be shared open access.

7.4 DATA SHARING

As it has been described before, the datasets will be organised in files tailored to comply with the requirements of WP6. The file structure will be also publicly available to the data users as a way to better understand the file itself.

The aggregated dataset will be disseminated in order to protect the commercial and industrial prospects of exploitable results (ER1 and ER8). In the case of the underlying data of a publication this might imply an embargo period for green open access publications.

Data objects will be deposited in ZENODO under open access to data files and metadata, permitting its use and reuse, as well as protecting privacy of its users.

7.5 ARCHIVING AND PRESERVATION

As it has been described before, the OPERA project database will be designed to remain operational for 5 years after project end. By the end of the project, the final dataset will be transferred to the ZENODO repository, which ensures sustainable archiving of the final research data.

8. REFERENCES

- [1] UK Data Archive. Research Data Lifecycle. Available at: <http://data-archive.ac.uk/create-manage/life-cycle>
- [2] Ocean Data Standards Project. Available at: <http://www.oceandatastandards.org/>
- [3] ISO 19156:2011 Geographic information -- Observations and measurements (2011). ISO.
- [4] IEC TS 62600-101:2015 Marine energy - Wave, tidal and other water current converters - Part 101: Wave energy resource assessment and characterization (2015). IEC.
- [5] ZENODO. Available at: <http://www.zenodo.org/>
- [6] OpenAIRE H2020 Project. Available at: <https://www.openaire.eu/>
- [7] DMAC: Dynamic Marine Component Test Facility, University of Exeter. Available at: <http://emps.exeter.ac.uk/renewable-energy/research/>
- [8] DNVGL-OS-E301 Position mooring (2015). DNVGL.
- [9] IEC TS 62600-10:2015 Marine energy - Wave, tidal and other water current converters - Part 10: Assessment of mooring system for marine energy converters (MECs) (2015). IEC.
- [10] UCC Beaufort Rotating test rig. Available at: http://www.fp7-marinet.eu/UCC_beaufort-rotating-test-rig.html
- [11] TECNALIA electrical PTO Lab. Available at: <http://www.tecnalia.com/en/energy-environment/news/electrical-ptolab.htm>
- [12] IEC TS 62600-100:2012 Marine energy - Wave, tidal and other water current converters - Part 100: Electricity producing wave energy converters - Power performance assessment (2012). IEC.
- [13] IEC TS 62600-102 Wave energy converter power performance assessment at a second location using measured assessment data (2016, draft). IEC.
- [14] IEC TS 62600-30 Marine Energy - Wave, tidal and other water current converters - Part 30: Electrical power quality requirements for wave, tidal and other water current energy converters (2016, draft). IEC.
- [15] IEC 61000-4-7:2002 Electromagnetic compatibility (EMC) – Part 4-7: Testing and measurement techniques – General guide on harmonics and interharmonics measurements and instrumentation (ed.2.0, 2002). IEC
- [16] Failure Reporting, Analysis and Corrective Action System (FRACAS) Application Guidelines, Product code Fracas. (1999). Reliability and maintenance data for equipment.
- [17] ISO 14224:2006, Petroleum, petrochemical and natural gas industries -- Collection and exchange of reliability and maintenance data for equipment. (2010).
- [18] DataCite. Available at: <https://www.datacite.org/>



- [19] Guidelines on Data Management in Horizon 2020, European Commission, Research & Innovation, 2013.
- [20] D8.2 Plan for dissemination and communication (2016). OPERA Project.
- [21] ORCID. Available at: <http://orcid.org/>
- [22] RECOLECTA by FECYT. Available at: <http://recolecta.fecyt.es/portada?language=es>



ANNEX I: KEY PRINCIPLES FOR OPEN ACCESS TO RESEARCH DATA

These principles can be applied to any project that produces, collects or processes research data. As indicated in Guidelines on Data Management in H2020 [19] , scientific research data should be easily:

1. Discoverable

The data and associated software produced and/or used in the project should be discoverable (and readily located), identifiable by means of a standard identification mechanism (e.g. Digital Object Identifier)

2. Accessible

Information about the modalities, scope and licenses (e.g. licencing framework for research and education, embargo periods, commercial exploitation, etc.) in which the data and associated software produced and/or used in the project is accessible should be provided.

3. Assessable and intelligible

The data and associated software produced and/or used in the project should be assessable for and intelligible to third parties in contexts such as scientific scrutiny and peer review (e.g. the minimal datasets are handled together with scientific papers for the purpose of peer review, data are provided in a way that judgments can be made about their reliability and the competence of those who created them).

4. Useable beyond the original purpose for which it was collected

The data and associated software produced and/or used in the project should be useable by third parties even long time after the collection of the data (e.g. data are safely stored in certified repositories for long term preservation and curation; they are stored together with the minimum software, metadata and documentation to make it useful; the data are useful for the wider public needs and usable for the likely purposes of non-specialists).

5. Interoperable to specific quality standards

The data and associated software produced and/or used in the project should be interoperable allowing data exchange between researchers, institutions, organisations, countries, etc. (e.g. adhering to standards for data annotation, data exchange, compliant with available software applications, and allowing re-combinations with different datasets from different origins).



ANNEX II: SCIENTIFIC PUBLICATIONS

Project Partners are responsible for the publication of relevant results to scientific community by Scientific Publications. According to OPERA DoA, at least 8 indexed manuscripts will be produced. An initial list of Scientific Publications is available in Deliverable D8.2 Plan for dissemination and communication [19] .

The data (including associated bibliographic metadata) needed to validate the results presented in scientific publications will be deposited in a research data repository. This requirement is based on the fact that the concept of 'publication' has rapidly evolved over the past years and in the context of the digital era. Therefore, the notion of 'publication' increasingly includes the data underpinning the publication and results presented, also referred to as 'underlying' data. This data is needed to validate the results presented in the deposited scientific publication and is therefore seen as a crucial part of the publication and an important ingredient enabling scientific best practice.

Metadata will maximise the discoverability of publications and ensure the acknowledgment of EU funding. Bibliographic data mining is more efficient than mining of full text versions. The inclusion of metadata is necessary for adequate monitoring, production of statistics, and assessment of the impact of H2020. In addition to basic bibliographic information about deposited publications the following metadata information is expected:

- EU funding acknowledgement:
 - Contributor: "European Union (EU)" & "Horizon 2020"
- Peer Reviewed type (e.g. accepted manuscript; published version).
- Embargo Period (if applicable):
 - End date.
 - Access mode.
- Project Information:
 - Grant number: "654444"
 - Name of the action: "Research and Innovation action"
 - Project Acronym: "OPERA"
 - Project Name: "Open Sea Operating Experience to Reduce Wave Energy Costs"
- Publication Date.
- Persistent Identifier:
 - Authors and Contributors. Wherever possible identifiers should be unique, non-proprietary, open and interoperable (e.g. through leveraging existing sustainable initiatives such as ORCID [21] for contributor identifiers and DataCite [18] for data identifiers).
 - Research Outcome



- License. The Commission encourages authors to retain their copyright and grant adequate licences to publishers. Creative Commons offers useful licensing solutions.

OPERA project will support the open access approach to Scientific Publications (as defined in article 29.2 of the Grant Agreement). Scientific Publications covered by an editorial copyright will be made available internally to the partners and shared publicly through references to the copyright owners web sites.

Whenever possible, a Scientific Publication, as soon as possible and at the latest six months after the publication time, will be deposited in a machine-readable electronic copy of the published version or final peer-reviewed manuscript accepted for publication in a repository for scientific publications. Moreover, the beneficiary should aim at depositing at the same time the research data needed to validate the results presented in the deposited scientific publications.

TECNALIA has just finalised the development of the TECNALIA Publications repository which is an open access repository accessible by RECOLECTA [22] (a platform which gathers all scientific repositories at Spanish national level) and OpenAire [6] (a new platform aimed at gathering a H2020 EU funded-projects' scientific publications). The repository is indexed by Google and fulfils international interoperability standards and protocols to gain long-term sustainability.

All scientific publications of the OPERA project will be available through OpenAire repository which allows searching publications per project. The potential delayed access ('embargo periods') required by specific publishers and magazines will be negotiated in a case-by-case basis.

All research data underpinning a publication will be openly accessible as defined in article 29.3 of OPERA Grant Agreement. Similarly as with the scientific publications, ZENODO [5] , which is the open access repository of OpenAIRE [6] , will be used principally.

