

Advanced Design Tools for Ocean Energy Systems Innovation, Development and Deployment

Deliverable D9.10

Data Management Plan – first version

Lead Beneficiary Tecnalia Delivery Date 26/10/2018 Dissemination Level Public Status Released 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 785921



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	785921
Project Acronym	DTOceanPlus
Work Package	WP9
Related Task(s)	Т9.5
Deliverable	D9.10
Title	Data Management Plan – first version
Author(s)	Pablo Ruiz-Minguela (Tecnalia)
File Name	DTOceanPlus_D9.10_Data management plan_Tecnalia_20181026_v1.o.docx

Revision History

Revision	Date	Description	Reviewer
0.1	31/08/2018	Initial outline	FEM
0.2	09/10/2018	Draft contents	PSC members
0.5	19/10/2018	Full draft for peer review	Francisco Fonseca & Marta Silva (WavEC)
1.0	26/10/2018	Final version for EC	EC





EXECUTIVE SUMMARY

This document, D9.10 Data Management Plan (DMP) is a deliverable of the DTOceanPlus project, which is funded by the European Union's H2020 Programme under Grant Agreement #785921.

The overarching objective of the DTOceanPlus project is to develop and demonstrate an open source, integrated suite of 2nd generation design tools for ocean energy technologies that support the entire technology innovation process. This suite of design tools will accelerate the development of the Ocean Energy sector and reduce the technical and financial risks of devices and arrays to achieve the deployment of cost-competitive wave and tidal arrays.

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. DTOceanPlus will develop and demonstrate an open source integrated suite of second generation design tools for ocean energy technologies supporting the entire technology innovation process. The open source licence under which these tools will be distributed will allow free access to all interested parties. As a user progresses through the stages of creating a design in DTOceanPlus, they will require access to reference data to support decision-making. Moreover, a database of long standing reference data will collect all the relevant information produced by the research and demonstration activities in the project.

The DTOceanPlus 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.

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. 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.

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. The next versions of the DMP will get into more detail and describe the practical data management procedures implemented by the DTOceanPlus project.

The expected types of research data that will be collected or generated along the project lie in the following categories: 1) Components; 2) Assessments; 3) Vessels, equipment and ports. Besides, underlying data needed to validate the results presented in scientific publications will be considered insofar possible for open access publication.





TABLE OF CONTENTS

EXECUTIVE SUMMARY
TABLE OF CONTENTS
LIST OF FIGURES
LIST OF TABLES
ABBREVIATIONS AND ACRONYMS
1. INTRODUCTION
1.1 Motivation
1.2 Purpose of the Data Management Plan9
1.3 Research data types in DTOceanPlus
1.4 Roles and responsibilities11
2. DATA COLLECTION, STORAGE AND BACK-UP12
2.1 Component data12
2.2 Assessment data13
2.3 Vessels, equipment and port data13
3. DATA STANDARDS AND METADATA15
4. DATA SHARING AND REUSE
5. DATA ARCHIVING AND PRESERVATION
6. REFERENCES
ANNEX I: KEY PRINCIPLES FOR OPEN ACCESS TO RESEARCH DATA19
ANNEX II: SCIENTIFIC PUBLICATIONS





LIST OF FIGURES

Figure 1.1: Research data life cycle (Adapted from UK Data Archive [2])	. 9
Figure 1.2: Research data options and timing	10





LIST OF TABLES

Table 1.1: DTOceanPlus types of data	10
Table 2.1: Component data	12
Table 2.1: Assessment data	13
Table 2.1: Vessels, equipment and port data	14





ABBREVIATIONS AND ACRONYMS

- DMP Data Management Plan
- DoA Description of Action
- DOF Degree of Freedom
- Dx.y Deliverable numbered, for example D8.6
- EC European Commission
- Tx.y Project task numbered, for example T8.2
- WP Work Package





1. INTRODUCTION

1.1 Motivation

The DTOceanPlus 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, 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. DTOceanPlus will develop and demonstrate an open source integrated suite of second generation design tools for ocean energy technologies supporting the entire technology innovation process. The open source licence under which these tools will be distributed will allow free access to all interested parties.

As a user progresses through the stages of creating a design in DTOceanPlus, they will require access to reference data to support decision-making. Moreover, a database of long standing reference data will collect all the relevant information produced by the research and demonstration activities in the project. Essentially, it will contain a catalogue of components, vessels, ports and equipment, as well as the associated features for assessments of designs such as performance, cost, reliability, environmental or social impact ratings. Actually, user consultation responses highlighted the need for transparent access to this kind of data.

Additionally, the underlying data needed to validate the results presented in scientific publications will be considered insofar possible for open access publication [1].

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. In this sense, the Commission applies the principle of 'as open as possible, as closed as necessary' and allow partial opt out due to intellectual property rights (IPR) concerns, privacy/data protection concerns or for other legitimate reasons. Strategies to limit such restrictions could 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 [2])

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 which will be updated in Month 21 (D9.11). It 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 DTOceanPlus

The data types that will be produced during the project are based on the Description of the Action (DoA) and their results.

According to such consideration, Table 1.1 reports a list of categories of research data that DTOceanPlus will produce. These research data types have been defined, 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 final version of the DMP to take into consideration the





project developments and scientific publications. A detailed description of each dataset is given in the following sections of this document.

#	Dataset category	Lead partner	Related WP(s)
1	Components	TECNALIA	WP5
2	Assessments	WES	WP4, WP6
3	Vessels, equipment and ports	WavEC	WP5

TABLE 1.1: DTOCEANPLUS TYPES OF DATA

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. Datasets can be both collected, unprocessed data as well as analysed, generated data. The policy for open access are summarised in Figure 1.2: Research data options and timingFigure 1.2.



FIGURE 1.2: RESEARCH DATA OPTIONS AND TIMING

Research data directly linked to the proprietary technologies or projects used for the validation of the design tools will not be released in the open domain as they can compromise the commercialisation prospects of industrial partners. 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

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.



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.



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 and be published as soon as they become available.

1.4 Roles and responsibilities

Each DTOceanPlus partner must respect the policies set out in this DMP. Datasets must be created, managed and stored appropriately and in line with applicable legislation.

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

EDF, as WP7 leader, will ensure dataset integrity and compatibility for its use during the validation of the design tools by different partners.

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, which facilitates 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 (particularly WP4-5-6-7), supported by the Project Coordinator.

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 publicly available data.

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





2. DATA COLLECTION, STORAGE AND BACK-UP

The DTOceanPlus project will produce reference data resulting from:

- Supplier datasheets.
- Literature review.
- Model fitting.

A database of long standing reference data will collect all the relevant information produced by the research and demonstration activities in the project. Essentially, it will contain a catalogue of components, vessels, ports and equipment, as well as the associated features for assessments of designs such as performance, cost, reliability, environmental or social impact ratings.

The following sections describe the different datasets that will be produced in the course of the project.

2.1 Component data

The physical characterisation of low level components provides key information to drive the design decisions of ocean energy subsystems, devices and full array projects. Availability of a large family of components will significantly facilitate design optimisation. Default values will be provided insofar they are necessary for completing an ocean energy design but difficult to determine.

Usually components will comprise balance of plant (e.g. mooring lines and shackles, power cables, connectors and switchgear,) and off-the-shelf components (e.g. generators and motors, gearboxes, cylinders, turbines, accumulators).

Among the various component features, it will be captured component materials, mass, sizing and main physical properties.

TABLE 2.1: COMPONENT DATA		
Reference/Name	DS_Component	
Description	Physical features of balance of plant and off-the-shelf components	
Source	 Supplier datasheets and literature review 	
Туре	Derived	
Format	CSV, MS Excel, SQL	
Software	• N/A	
Estimated size	• 1 GB	
Storage	Database	
Back-up	Daily back-ups on both local and cloud-hosted servers	

A short description of the component dataset is given below.





2.2 Assessment data

Assessment data will be required to take global design decisions. There are four types of assessments in DTOceanPlus:

- Performance and energy yield (e.g. efficiency curve, ...)
- Reliability, availability, maintainability and survivability (e.g. failure rate, design limits, ...)
- Lifetime costs (e.g. cost of manufacture, assembly, replace, repair, ...)
- Environmental and social acceptance (e.g. stressors, CO2 emissions, ...)

Quantitative rating of various performance parameters will be required at component level and later on aggregated for subsystems, devices and ultimately the whole array. Benchmarks and thresholds for structured innovation and stage gate design tools will be also considered.

A short description of the assessment datasets is given below.

Reference/Name	DS_Assessment
Description	 Quantitative rating of performance parameters
Source	 Supplier datasheets, literature review and model fitting
Туре	Derived
Format	CSV, MS Excel, SQL
Software	• N/A
Estimated size	• 1 GB
Storage	Database
Back-up	Daily back-ups on both local and cloud-hosted servers

TABLE 2.2: ASSESSMENT DATA

2.3 Vessels, equipment and port data

The suitable design of offshore logistics is a paramount to establish the global design of a particular project. Apart from the physical components and systems, a full characterisation of a wide range of vessels, equipment and port data is required.

Among the various features to be captured, there is:

- Physical description: dock space, loading capacity, storage area, cranes, vessel size & speed, bollard/winch pull, operating limits, crew, divers, ROV, ...
- Quantitative rating: use costs, average fuel consumption, noise level, ...

A short description of the logistics datasets is given below.





Reference/Name	DS_Logisitics
Description	 Range of vessels, equipment and port data
Source	 Supplier datasheets, literature review and model fitting
Туре	Derived
Format	CSV, MS Excel, SQL
Software	• N/A
Estimated size	• 1GB
Storage	Database
Back-up	 Daily back-ups on both local and cloud-hosted servers

TABLE 2.3: VESSELS, EQUIPMENT AND PORT DATA





3. DATA STANDARDS AND METADATA

The following standards should be used for data documentation:

- **DNV-RP-J**301 [3] : Subsea Power Cables in Shallow Water Renewable Energy Applications.
- DNVGL-OS-E301 [4] : it contains criteria, technical requirements and guidelines on design and construction of position mooring systems. The objective of this standard is to give a uniform level of safety for mooring systems, consisting of chain, steel wire ropes and fibre rope.
- IEC TS 62600-10 [5]: technical specification for assessment of mooring system for Marine Energy Converters (MECs).
- ▶ IEC TS 62600-30 [6] : technical specification on electrical power quality requirements for wave, tidal and other water current energy converters
- ► IEC TS 62600-100 [7] : technical specification on power performance assessment of electricity producing wave energy converters
- IEC TS 62600-200 [8] : Electricity producing tidal energy converters Power performance assessment
- ▶ ISO 14224:2006 [9] : collection and exchange of reliability and maintenance data for equipment.

Metadata records will accompany the data files in order to describe, contextualise and facilitate external users to understand and reuse the data.

DTOceanPlus will adopt the DataCite Metadata Schema [10], a domain agnostic metadata schema, as the basis for harvesting and importing metadata about datasets from data archives. The core mission of DataCite is to build and maintain a sustainable framework that makes it possible to cite data through the use of persistent identifiers.

The following metadata should be created to identify datasets:

- Identifier: A unique string that identifies the dataset
- Author/Creator: The main researchers involved in producing the data in priority order
- Title: A name or title by which a data is known
- Publisher: The name of the entity that holds, archives, publishes prints, distributes, releases, issues, or produces the data.
- Publication Year: The year when the data was or will be made publicly available
- Subject: Subject, keyword, classification code, or key phrase describing the resource.
- Contributor: Name of the funding entity (i.e. "European Union" & "Horizon 2020")
- Size: Unstructured size information about the dataset (in GBs)
- Format: Technical format of the dataset (e.g. csv, txt, xml, ...)
- Version: The version number of the dataset
- Access rights: Provide a rights management statement for the dataset. Include embargo information if applicable
- Geo-location: Spatial region or named place where the data was gathered





4. DATA SHARING AND REUSE

During the life cycle of the DTOceanPlus project datasets will be stored and systematically organised in a relational database tailored to comply with the requirements of WP7. 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 [11], which is the open access repository of the Open Access Infrastructure for Research in Europe, OpenAIRE [12].

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.

By default, data access policy will be unrestricted unless otherwise specified. The generic Creative Commons CC-BY licences will be used. This licence allows:

- Sharing copy and redistribute the material in any medium or format
- Adapting remix, transform, and build upon the material for any purpose, even commercially.





5. DATA ARCHIVING AND PRESERVATION

The DTOceanPlus 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.





6. REFERENCES

- [1] D2.1 Results from user-groups consultation (2018). DTOceanPlus Project.
- [2] UK Data Archive. Research Data Lifecycle. Available at: <u>http://data-archive.ac.uk/create-manage/life-cycle</u>
- [3] DNVGL-RP-J301 Subsea Power Cables in Shallow Water Renewable Energy Applications (2014). DNVGL.
- [4] DNVGL-OS-E301 Position mooring (2015). DNVGL.
- [5] 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.
- [6] IEC TS 62600-30:2018 Marine energy Wave, tidal and other water current converters Part 30: Electrical power quality requirements (2018). IEC.
- [7] 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.
- [8] IEC TS 62600-200:2013 Marine energy Wave, tidal and other water current converters Part 200: Electricity producing tidal energy converters - Power performance assessment (2013). IEC.
- [9] ISO 14224:2006, Petroleum, petrochemical and natural gas industries -- Collection and exchange of reliability and maintenance data for equipment. (2010).
- [10] DataCite. Available at: <u>https://www.datacite.org/</u>
- [11] ZENODO. Available at: <u>http://www.zenodo.org/</u>
- [12] OpenAIRE H2020 Project. Available at: <u>https://www.openaire.eu/</u>
- [13] Guidelines on Data Management in Horizon 2020, European Commission, Research & Innovation, 2013.
- [14] D9.2 Dissemination and communication plan (2018). DTOceanPlus Project.
- [15] ORCID. Available at: <u>http://orcid.org/</u>
- [16] 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 [13], 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 DTOceanPlus DoA, at least 6 indexed manuscripts will be produced. The list of scientific publications is available in Deliverable D9.2 Dissemination and communication plan [14].

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):
 - o End date.
 - Access mode.
- Project Information:

- o Grant number: "785921"
- Name of the action: "Demonstration action"
- Project Acronym: "DTOceanPlus"
- Project Name: "Advanced Design Tools for Ocean Energy Systems Innovation, Development and Deployment"
- Publication Date.
- Persistent Identifier:
 - Authors and Contributors. Wherever possible identifiers should be unique, nonproprietary, open and interoperable (e.g. through leveraging existing sustainable initiatives such as ORCID [15] for contributor identifiers and DataCite [10] for data identifiers).
 - o Research Outcome
- License. The Commission encourages authors to retain their copyright and grant adequate licences to publishers. Creative Commons offers useful licensing solutions.





DTOceanPlus 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 [16] (a platform which gathers all scientific repositories at Spanish national level) and OpenAIRE [12] (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 DTOceanPlus 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 DTOceanPlus Grant Agreement. Similarly, as with the scientific publications, ZENODO [4], which is the open access repository of OpenAIRE [12], will be used principally.





CONTACT DETAILS

Mr. Pablo Ruiz-Minguela Project Coordinator, TECNALIA www.dtoceanplus.eu





This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 785921