atlantis

The Atlantic Testing Platform for Maritime Robotics

Topic	ICT-09-2019-2020 (H2020)
Acronym	ATLANTIS
Title	The Atlantic Testing Platform for Maritime Robotics: New Frontiers for Inspection and Maintenance of Offshore Energy Infrastructures.
Project number	871571
Delivery date	30.06.2020
Deliverable number	D7.7
Dissemination level	PUBLIC
Lead Beneficiary	INESC TEC

Data management manual

Written by INESC TEC

Actions

	Action	Organisation	Date
Technical Manager	Requested deliverable from the Deliverable Responsible.	VTT	20.05.2020
Deliverable Responsible	Prepared draft of the deliverable.	INESC TEC	16.06.2020
Technical Manager	Approved the updated draft as the first version.	VTT	24.06.2020
Quality Manager	Approved the updated first version as the second version.	UDG	29.06.2020
Project Coordinator	Approved the updated second version as the final version and sent to the European Commission.	INESC TEC	30.06.2020

Disclaimer

This project has received funding from the European Union's Horizon 2020 research and innovation programme, under the Grant Agreement no. 871571.

Table of Contents

This project has received funding from the European Union's Horizon 2020 research and innovation programme, under the Grant Agreement no. 871571.

List of Abbreviations

PMM	Project Management Manual
GA	Grant Agreement
DoA	Description of Action
O&M	Operation and Maintenance
UAV	Unmanned Aerial Vehicle
SMEs	Small and Medium Enterprises
AD	Advisory Board
PGA	Project General Assembly
PCB	Project Coordination Board
PMT	Project Management Team
FUA	Follow Up Actions
EC	European Commission
EU	European Union
DIHs	Digital Innovation Hubs
QAP	Quality Assurance Procedure
PDF	Portable Document Format
IMR	Inspection, Maintenance and Repair
IPR	Intellectual Property Rights
WFA	WindFloat Atlantic

This project has received funding from the European Union's Horizon 2020 research and innovation programme, under the Grant Agreement no. 871571.

1. Introduction

This document is entitled "*Data Management Plan"* and it is developed as part of the ATLANTIS project that received funding from the European Union's Horizon 2020 Research and Innovation program under the Grant Agreement number 871571.

The main objective of the Data Management Plan (DMP) is to provide a full description of the data management life cycle for all datasets to be collected, processed, or generated during the execution of the project.

This document is framed by the work package 7 (*Impact Analysis, Dissemination, Communication and Exploitation*) to act as deliverable D7.7, and it will be used to disclose the data management policies to apply throughout the project to aid the Consortium to meet their responsibilities as regards research data quality, sharing and security.

1.1. Purpose of the document

This document provides a full description of the data management policy for the complete life cycle of all datasets to be collected, processed, or generated during the execution of the project. It will cover what data will be collected, the methodologies and standards used to handle the datasets, the data accessibility to comply with H2020 regulations without damage to the EU privacy laws and the consortium intellectual property, and the frameworks to preserve the data.

Initially, the DMP will provide an overview that will describe what data the project will use and produce, whether and how data produced will be exploited or made (openly) accessible for verification and re-use, and how these data will be curated and preserved after the end of the project.

Furthermore, it will define the guiding principles, the roles, and responsibilities for the ATLANTIS consortium. The data accessibility and sharing principles will be provided to ensure open access datasets without compromising the confidential information of the consortium. The respect of the policies and responsibilities will be assured by the data management team here depicted with the respective role definitions for each member.

Moreover, the ATLANTIS Data Management Policy will be described in accordance with the guidelines for effective data management during the execution of Horizon 2020 projects, established by the EC [1], addressing aspects related to the collected, processed, and generated datasets, such as the dataset identification, the standards and metadata, the data sharing and ownership, and archiving and preservation.

The DMP will also provide a catalogue of all open and non-open data used and produced by the ATLANTIS project. Thus, depicting the copyright and permissions for reusing third-party data sets, personal data treatment and confidentiality issues and other restrictions. Also, a workflow will be defined with the description of the DMP review process and timetable to create datasets, add resources, and evaluate ethics and confidentiality compliance.

Finally, this document produces a summary describing the tentative datasets aimed by ATLANTIS in relation to all tasks in the proposal identifying the expected data to retrieve, the standards to be applied and the nature of the accessibility (Open or Non-Open).

The term of data that is used in this document shall also include images, 3D models, videos or any other media form in any image format.

This project has received funding from the European Union's Horizon 2020 research and innovation programme, under the Grant Agreement no. 871571.

1.2. Key documents

This is the list of key documents that will be addressed for the execution of the ATLANTIS project:

- Grant Agreement number 871571, and the DoA (part A of Annex 1) is also a key document to be taken into account;
- Consortium Agreement;
- Quality plan, risk and peer review (deliverable D8.5).

1.3. Structure of the document

The structure of the deliverable is organized in the following sections:

- Section 1 describes the context of this document.
- Section 2 presents the data management procedures to be followed during and after the project duration describing the data management life cycle for the data to be collected, processed, and generated by the ATLANTIS project.
- Section 3 summarizes the datasets that ATLANTIS aims to use and produce throughout the project duration;
- Section 4 presents a brief conclusion to all the procedures and datasets proposals throughout this document.

This project has received funding from the European Union's Horizon 2020 research and innovation programme, under the Grant Agreement no. 871571.

2. Data management

The ATLANTIS Data Management Plan will describe the data management life cycle for the data to be collected, processed and generated by the ATLANTIS project. The initial version of the Data Management Plan will be conducted by M6 and it will evolve during the lifetime of the project to present the status of the project's reflections on data management. Information about the handling of research data during and after the end of the project, what data will be collected, processed and generated, which standards or methodologies will be applied. Moreover, the D7.7 Data Management Plan due in M6 will describe procedures and rules for data that will be shared or made open access which includes how this data will be preserved and curated after the end of the project.

2.1. Overview

Here is established and detailed a project data management policy that handles the above issues on both an administrative and technical level. This includes for example topics like the definition of roles and responsibilities of individuals involved in data management, data identification guidelines, standardization in the context of metadata definition, data ownership and confidentiality and data archiving and indexing solutions. Furthermore, and perhaps more importantly, the ATLANTIS data catalogue is described, considering the several types of data that are used and produced within the project and the access restrictions that may apply. Finally, a workflow is provided to guide the project partners in uploading and indexing a dataset and its metadata on the chosen platforms. In particular, ATLANTIS will use Zenodo to share open data generated during the execution of the project. Zenodo, a multi-disciplinary open repository maintained by CERN, is compliant with the data management requirements of Horizon 2020 and Horizon Europe, the EU's research and innovation funding programmes. The ATLANTIS DMP has been designed to be compliant with OpenAIRE - the Open Access Infrastructure for Research in Europe.

2.2.Guiding Principles, Roles and Responsibilities

Data management activities concern the whole project and need to be coordinated and monitored both at project and work package level. Data management is also linked to publication of project results and thus dissemination activities. This said, a set of relevant roles and responsibilities have been identified.

The Project Data Manager, INESC TEC, is responsible for:

- developing the data management plan and policy in cooperation with the WP7 leader and the technical partners;
- coordinating the technical realisation of ATLANTIS 's OpenAIRE-compliant public data repository;
- developing a user guide for the usage of ATLANTIS ’s data repository;
- monitoring data management activities (both collection and publication) and deadlines and sending reminders to WP data managers;
- providing support to WP data managers;
- coordinating the writing of the DMP deliverable documents (D7.7);
- providing solutions for data management related issues in accordance with this plan.

Each Work Package Leader is responsible for:

This project has received funding from the European Union’s Horizon 2020 research and innovation programme, under the Grant Agreement no. 871571.

- the implementation of the data management policy in their respective WPs;
- monitoring data management activities and deadlines and sending reminders to partners;
- providing assistance in using and managing the ATLANTIS 's open data repository;
- asking partners for missing information or clarifications;
- providing input to the DMP deliverable documents (D7.7) by analysing and summarizing the WP specific datasets listed in ATLANTIS 's open data repository;
- ensuring that the published open data stored in Zenodo is properly linked to the ATLANTIS OpenAIRE initiative;
- contacting the quality assurance and ethics committee in case of questions and ethical and privacy issues that may forbid a publication of the data.

The Dissemination Manager (VTT) is responsible for:

- offering assistance in choosing the right publication path;
- offering help and guidance for publishing scientific publications;
- ensuring that the open access policy of the journal complies with the H2020 open data requirements [2] before the researcher submits a manuscript;
- guaranteeing that publications are deposited in the repositories and sending reminders to partners;
- monitoring that metadata about publications is made available in the R&I Participant Portal (preferably automatically through OpenAIRE) and on the ATLANTIS website (https://www.atlantis-h2020.eu/);
- monitoring that research data related to a publication is made available in the Zenodo repository and linked to respective publication;
- assessing possible embargo periods and sending reminders to partners;
- guarantee that publications available in OpenAIRE/Zenodo are properly linked with ATLANTIS.

2.3.The ATLANTIS Data Management Policy

The DMP's Data Management Policy will address the points below and will describe the perspective within the consortium regarding the data that will be produced, covering the complete research data life cycle.

This section is in accordance with the guidelines for effective data management during the execution of Horizon 2020 projects, established by the EC [1]. Their primary objective is to help H2020 beneficiaries make their research data findable, accessible, interoperable and re-usable (FAIR), to ensure it is soundly managed. By adopting those guidelines, the ATLANTIS commits to implement FAIR Data Management Protocols. In particular, the following aspects related to the collected, processed and generated data will be addressed:

- Dataset identification;
- Standards and metadata;
- Data access, sharing and ownership;
- Archiving and preservation.

This project has received funding from the European Union's Horizon 2020 research and innovation programme, under the Grant Agreement no. 871571.

2.3.1. Dataset identification

In addition to the standardized dataset identifier that will be used to share the generated data (see below), an internal project dataset identifier is required to keep track of all the data collected or created within the scope of each task. This will follow the format:

WPNumber_TaskNumber_PartnerName_DataSubset_DatasetName_Version_DateOfStorage

where the PartnerName represents the name of the data custodian (the one that acquired the data). An example of this naming format would be: WP3_T3.5_INESCTEC_Subset a_ArrayCableSonarScan_V1_08.05.21.

2.3.2. Standards and metadata

The OpenAIRE project, in the vanguard of the open access and open data movements in Europe was commissioned by the EC to support their nascent Open Data policy by providing a catch-all repository for EC funded research. OpenAIRE adopted the DataCite Metadata Schema as the core format to gather metadata about datasets. The mission of DataCite is to build and maintain a framework that enables the possibility of citing data through the use of persistent identifiers. OpenAIRE shares the goal of the DataCite Metadata Schema:

- to provide a domain agnostic metadata schema;
- provide interoperability through a small number of properties.

This methodology makes interoperability possible without adding complexity, therefore reducing the technical barriers for implementation.

All data that is generated and made publicly available by ATLANTIS will be preserved together with a common set of metadata following OpenAIRE guidelines [3], whether it is for public dissemination or just for internal use. Additional metadata can be included to accommodate user requirements. The common set of metadata is:

- **Identifier**: unique string that identifies a resource to be shared; allowed values include ARK, DOI, Handle, PURL, URN and URL.
- **Alternate identifier**: identifier other than the primary Identifier applied to the resource; may be the internal identifier defined previously (optional).
- **Creator**: main researchers involved in producing the data; may be a corporate/institutional or personal name.
- **Title**: name or title by which a resource is known.
- **Publisher**: the name of the entity that holds, archives, publishes prints, distributes, releases, issues, or produces the resource; this property will be used to formulate the citation.
- **Publication year**: year when the data was or will be made publicly available; in case an embargo period has been in effect, use the date when the embargo period ends.
- **Subject**: subject, keyword, classification code, or key phrase describing the resource.

This project has received funding from the European Union's Horizon 2020 research and innovation programme, under the Grant Agreement no. 871571.

- **Contributor**: institution responsible for collecting, managing, distributing, or otherwise contributing to the development of the resource. This property to allow identification of the funder who has funded wholly or partly the dataset described.
- **Date**: dates relevant to the work (accepted, collected, created, updated, etc), in the format YYYY-MM-DD.
- **Language**: primary language of the resource.
- **Resource** Type: description of the resource and the type of data it contains (ex: Audiovisual, dataset, image, model, text, etc).
- **Format**: technical format of the resource (ex: PDF, MPG, etc).
- **Rights**: provide a rights management statement for the resource; include embargo information if applicable.
- **Description**: additional information that does not fit in any of the other categories, including its origin (in cases where data is collected), parameters, units/scale, to whom it could be useful, and whether it underpins a scientific publication; may be used for specifying technical information related to methods and sensors used to produce the data.

The standard output formats that will be produced in ATLANTIS are:

- .png, .jpg or .pdf (figure) for maps, plots;
- .mp4 or .mpg for audiovisual content;
- .csv or .xslx for point and time series data;
- .bag for raw data from ROS-based platforms.

As for sonar data, the format depends on the sonar vendor, which typically is proprietary.

2.3.3. Data access, sharing and ownership

Most of the data produced in ATLANTIS will be retrieved from the robotic platforms owned and operated by the technological partners during the IMR activities, or during the platform development and adaptation phase.

According to EU guidelines, all data will be made publicly available for verification and reuse, unless a justification is given on why the data cannot be made openly accessible. The section 3.1.1 presents a categorization of data that will not be made public nor accessible outside the ATLANTIS Consortium. Outside of the scope of section 3.1.1. ATLANTIS implements a three-step decision making process for publishing data publicly:

- First, the required publication shall be submitted to the leader of the task whose activities involved the generation of the data.
- Second, if the task leader understands that such publication shall not be allowed he must provide a reason why the data cannot be made openly accessible;
- Thirdly, the Project Coordination Board (PCB) will assess such justifications and make the final judgement on whether to publish the data openly or not.

This decision making process is based on close examination of the following criteria regarding the confidentiality of the datasets:
(i) Commercial and technical sensitivity of datasets;

This project has received funding from the European Union's Horizon 2020 research and innovation programme, under the Grant Agreement no. 871571.

(ii) Data confidentiality for security reasons;
(iii) Conflicts between open-access rules and national and European legislation (e.g. data protection regulations);
(iv) Sharing data would jeopardise the aims of the project;
(v) WindFloat Technology IP protection. The catalogue of data describing the WindFloat Technology is provided by PPF;
(vi) Other legitimate reasons, to be validated by the PCB.

Further details on the workflow for sharing data generated under the ATLANTIS project are specified in section 3.2.

The GDPR imposes an important restriction for sharing data generated during the ATLANTIS project. Under this project, the following personal data from partners will be collected and stored:

- Contact names;
- Employing institutions;
- Email addresses;
- Photos taken at meetings and workshops.

This information will be kept in an Excel file stored at the INESC TEC's private storage system (https://drive.inesctec.pt/apps/files/?dir=/ATLS_Management/ATLS_IPR), only reachable for selected individuals related to the project. This data will therefore not be made public, only shared with authorized personnel from the consortium and won't be used for purposes other than the ATLANTIS Stakeholders activities. Whenever this information is requested, explicit consent will be required from each involved individual. At the end of the project, the contact file will be destroyed.

The selected data from several tests to be performed during the project will be submitted to Zenodo OpenAIRE repository (see below) once approved by the PCB. The consortium will also explore the possibility of making any earth observation (EO) (meta)data acquired by the vehicles available via the European Marine Observation and Data Network (EMODnet). The data owner is the respective data provider, namely the partner operating the vehicle. All processed and higher-level data generated in ATLANTIS are considered as ATLANTIS products. Partners that have contributed to each product (by developing methods and procedures, data processing or validation) will be considered as shared owners of the products.

For avoidance of the doubt, PPF owns all data collected concerning WindFloat Platform and such data cannot be used outside of the ATLANTIS project, or published, without PPF prior written consent.

A PMB meeting at the end of each reporting period will be used to agree on the owners of each of the products developed in this period. All partners have access to all products. They will be shared internally via the INESC TEC's cloud storage system.

2.3.4. Archiving and preservation

In addition to the internal project database for IPR-restricted data (located at the INESC TEC's cloud storage system at https://drive.inesctec.pt/apps/files/?dir=/ATLS_Management/ATLS_IPR), public datasets will be stored in Zenodo, which is the open access repository of the Open Access Infrastructure for Research in Europe, OpenAIRE. Datasets uploaded to Zenodo will be held for the lifetime of the repository, which is currently the lifetime of the host institution CERN. Data files and related metadata

This project has received funding from the European Union's Horizon 2020 research and innovation programme, under the Grant Agreement no. 871571.

are backed up on a nightly basis, as well as replicated in redundant copies in the online platform. Zenodo was built and developed by researchers, to ensure that everyone can join in Open Science. At the moment, data uploading to Zenodo is free of charge (for datasets up to 50GB), but donations towards sustainability are encouraged. The ATLANTIS features a community page at Zenodo (https://zenodo.org/communities/atlantis/) that will aggregate all data related to the ATLANTIS project and data gathered by robotic assets at the infrastructure site after the project ends.

In terms of the information regarding the procedures to archive and preserve each dataset, in some cases it may be required to define how long the data should be preserved.

3. Data catalogue

The overall role of the OpenAIRE catalogue in ATLANTIS data management activities is that of a metadata catalogue that indexes and aggregates data that is stored in separate data repositories. The open data produced by the project will be stored in a public research data repository (Zenodo) and the respective metadata will be automatically replicated in OpenAIRE (https://explore.openaire.eu/search/project?projectId=corda__h2020::b575ec66952093a175c6e26e9cf86713) to enhance the "findability" and the long-term preservation of project data and findings. As for the private data produced within the project, it will be securely stored in the (access restricted) INESC TEC cloud storage system (https://drive.inesctec.pt/apps/files/?dir=/ATLS_Management/ATLS_IPR).

3.1.1. Non-open data used by ATLANTIS

This data is made available or sold by a data provider under a restricted license. Regarding ATLANTIS, only the respective meta-data can be published in OpenAIRE and the actual data is stored on the INESC TEC's cloud storage system with access restrictions since it is not meant to be shared outside of the ATLANTIS consortium. In the OpenAIRE, such information entries are characterized with restricted or closed access modes.

Thus, in this case, the respective Resources (datasets, images, videos, documentation, etc.) associated with the meta-data in OpenAIRE are just simple links to data files on INESC TEC 's access-controlled cloud storage system.

In particular, we have the types of data discussed below.

Imaging data related to the WFA

Type	Format	Size
Optical Images	JPEG, PNG, PDF	10's MB
Data related to the WindFloat	any format	1's MB to 10's GB

Purpose of the data usage: This data was provided to help accomplish part of WP1 tasks, namely defining the showcases, identifying the components of the floating structure and typical IMR procedures.

This project has received funding from the European Union's Horizon 2020 research and innovation programme, under the Grant Agreement no. 871571.

Relation to the objectives of the project: Helps the technological partners to provide tailored solutions to optimise the I&M activities using its unmanned systems.

Re-use of existing data: None.

Origin of data: The data was gathered following the execution of several activities involving the WFA. WFA is owned by WindPlus consortium, which is jointly owned by EDP Renováveis (54.4%), ENGIE (25%), Repsol (19.4%) and Principle Power Inc. (1.2%).

Data utility: This data is useful for robotic platform manufacturers and operators of the ATLANTIS consortium that will perform IMR operations throughout the project. Such data enables a better understanding of the components of the structure to be inspected and its IMR requirements, which allow the future development of customized algorithms and procedures to perform the operations defined in the identified scenarios.

Wind farm monitoring data

Purpose of the data usage: Understand the environmental conditions in the target area with greater detail. The data will be used to prepare for the offshore activities in the wind farm that take place in WP5.

Relation to the objectives of the project: Helps the technological partners to provide tailored solutions to optimise the I&M activities using its unmanned systems.

Re-use of existing data: None.

Origin of data: The data will be made available following the development of the Supervisory Control Center (SCC), which could be integrated with the wind farm monitoring infrastructure.

Data utility: This data is useful for robotic platform manufacturers and operators of the ATLANTIS consortium that will perform IMR operations throughout the project. Such data enables a better understanding of the conditions in the target area and allows for: (1) an assessment on the executability of the IMR operations at any given time and (2) an appropriate deployment and mission plan to be created.

3.1.2. Open data used by ATLANTIS

These datasets may encompass data that are collected from public authorities and institutions like Instituto Hidrográfico (https://www.hidrografico.pt/) or Copernicus (https://www.copernicus.eu/en), that are released under an open license that allows using the data for research and non-commercial purposes. While ATLANTIS 's OpenAIRE Catalogue provides some meta-data related to the usage of this data in ATLANTIS, the actual data can be downloaded directly from the websites of the respective organisations that produced it. Although some of the data is also stored on the INESC TEC cloud storage system for the purpose of processing, analysis and visualisation, the respective OpenAIRE resources associated with the OpenAIRE dataset entry (meta-data) links to the original source of the dataset.

This project has received funding from the European Union's Horizon 2020 research and innovation programme, under the Grant Agreement no. 871571.

In particular, we consider the types of data discussed below.

Environmental data

Type	Format	Size
Bathymetry maps / contour lines	ArcGIS Shapefile	10's MB
Tide times and charts	Text	MB
Weather forecasts	Text	MB

Purpose of the data usage: Identify environmental factors that may limit the execution of IMR operations in the field trials that will take place in WP3 and WP5.

Relation to the objectives of the project: (1) Helps the technological partners to provide tailored solutions to optimise the I&M activities using its unmanned systems, including methodologies for the operation of such systems. (2) Promotes safety in the context of those activities.

Re-use of existing data: External analysis of any information acquired by the robots during the execution of the IMR activities may require access to this data to contextualize and provide a better understanding of the behaviour of the robot and the performance of the procedures.

Origin of data: The data is gathered as required, specifically when real world experiments need to take place. The bathymetric maps are downloaded from the Instituto Hidrográfico web page (https://www.hidrografico.pt/) and the tide and weather forecasts are published in Instituto Português do Mar e Atmosfera web page (https://www.ipma.pt).

Data utility: This data is useful for robotic platform operators that will have to perform field experiments. Such data gives a more detailed picture of the environment and the environmental conditions available at wind farm and coastal testbed locations, which enables the specification of an informed operations schedule maximizing the efficiency of the activities to be performed.

Documentation

Type	Format	Size
Scientific publications	PDF	10's MB
Device datasheets	PDF	10's MB

Purpose of the data usage: Gather information about relevant procedures and techniques to be used by the robotic platforms when used in the context of IMR operations.

Relation to the objectives of the project: (1) Helps the technological partners to develop adequate and efficient solutions to optimise the I&M activities using its unmanned systems. (2) Ensures any solutions developed under the ATLANTIS project are in line, if not beyond, with the current state of the art concerning these type of activities.

This project has received funding from the European Union's Horizon 2020 research and innovation programme, under the Grant Agreement no. 871571.

Re-use of existing data: May be used for analysis and comparison purposes, considering the developments made during this project.

Origin of data: There are two main routes for accessing open access publications which differ in the way in the way it was published: Green and Gold access. Gold open access means the article is available as open access by the scientific publisher. Green open access or self-archiving and means that the published article or the final peer-reviewed manuscript is archived by the researcher in an online repository (e.g. the project website and other open research repositories), in most cases after its publication. Information about systems and sensors is typically available at their manufacturer's website.

Data utility: This data is useful for establishing a baseline on expected behaviour of the robotic platforms in the context of IMR activities. It also allows for comparison of adopted approaches considering the individual performance of each robot subsystem and, in general, of the robotic platform.

3.1.3. Non-open data produced by ATLANTIS

This part of the ATLANTIS DMP reports non-open data produced by the ATLANTIS project. According H2020 data management obligations, data produced by the project should be open by default. However, not all datasets produced within the project will be publicly available, and in the cases were a dataset is public, there might still be parts of the dataset that remain non-public. The main reasons that justify this approach are:

- Some of the data shared or acquired by the partners *reveals details of critical infrastructure design or operations that are to be considered technical, business and organisational sensitive information* and we do not have permission from the concerned organisations to make this data available. Under ATLANTIS, this exception will apply to data regarding the WFA's WindFloat structures as well as to a variety of gathered technical information about the wind power turbine and the various robotic platforms (and their systems). Related information is only to be made publicly available with explicit permission from the copyright holder and if it does not lead to a potential misuse of research findings.
- Copyright and permissions for reusing third-party data sets - Processing and integrating input data from different sources may lead to unclear IPR situations regarding the generated output data. Such data can only be made open if any of the underlying data, on which it is based, is open, too. Under ATLANTIS, this exception will apply to the 3D models of the structures to be inspected (including similar models of the WindFloat, build up with information provided by the WindFloat designers (PPF)), to the data acquired by the various robotic platforms (and their systems) during the experiments and to the wind farm system data.
- The data acquired in this project is context specific, and the publicly available part of this data is at the highest level of detail (as in the form of "raw data"), which could lead to misinterpretations and erroneous assessments when analysed and processed by external entities to the project.
- Personal data treatment and confidentiality issues - Datasets referring to the quality and quantity of certain elements at risk, such as people and critical infrastructures, are not open by default as their publication may pose privacy, ethical or security risks.
- Other restrictions that can be identified in the future by the PCB.

This project has received funding from the European Union's Horizon 2020 research and innovation programme, under the Grant Agreement no. 871571.

In particular, we have the types of data discussed below.

Imaging data including the WFA

Type	Format	Size
Optical Images	JPEG, PNG, PDF	10's MB
Optical videos	MP4, AVI, or other	100's MB - GB
Acoustic	Proprietary formats tied to sonar model, XTF	100's MB - 10's GB

Purpose of the data collection/generation: (1) Assessing the operational status and condition of various components of the WindFloat platforms at WFA site, and predict their expected remaining service life, (2) identify regions for further maintenance operations, (3) obtain a performance indicator of the robotic-based IMR and demonstrate its potential. This data will be produced in offshore field trials that take place in WP5.

Relation to the objectives of the project: This data contributes to one of the main project objectives, which is to establish robotic assets as an indispensable tool when performing IMR operations. It helps to maximize the integrity of the target structure and the safety and efficiency of the IMR process.

Re-use of existing data: This data can be used for simulation purposes to improve perception and navigation systems, for example, and overall mission procedures.

Origin of data: This data will be gathered by all the technological partners that will deploy robotic assets with the corresponding sensing capabilities during offshore testing. Other entities that may be involved during this project will have specific access conditions to the WFA.

Data utility: The results of equipment performance will be used and treated in order to assess the impact of such technologies in the global projects/ technology LCOE. This data can be useful for actors that are interested in deploying robotic assets for performing similar inspection operations. In particular, sonar data acquisition allows the real-time mapping of the environment during mission execution, enabling online planning and computation of obstacle-free trajectories. The data gathered here can be used by the research and industrial partners only for the purpose of the ATLANTIS project to foster the development of algorithms and new systems (e. g. through machine learning techniques and new sensing technologies) to increase autonomy and perception capability of robots during IMR operations. Furthermore, all the data acquired here is useful for reporting purposes and for further diagnosis by qualified personnel. One expected outcome, if deemed required, may be delivering an operational forecast to the decision makers for planning of inspection and maintenance operations.

Documentation involving data that concerns the WFA

Type	Format	Size
Internal IMR reports	PDF	10's MB

This project has received funding from the European Union's Horizon 2020 research and innovation programme, under the Grant Agreement no. 871571.

Internal notes	PDF	10's MB

Purpose of the data collection/generation: Documenting all the information gathered during IMR operations, as well as a preliminary analysis of the data. This data will be produced during the definition of the showcases in WP1 and during the elaboration of the reports related to the offshore field trials in WP5.

Relation to the objectives of the project: This data contributes to one of the main project objectives, which is to establish robotic assets as an indispensable tool when performing IMR operations in challenging environments.

Re-use of existing data: May be used for defining an optimal maintenance schedule in the future. Historical data and knowledge can be used to adjust or redesign components/equipment/systems/ procedures to improve reliability and efficiency.

Origin of data: This data will be produced with the input of all partners. The data produced here is dependent on the imaging data acquired during the field experiments and IMR operations.

Data utility: Same as above, but more focused on the end user needs.

3.1.4. Open data produced by ATLANTIS

Unless otherwise specified in this document, all data produced by ATLANTIS is open by default. As mentioned previously, the exceptions are data that explicitly involves the WFA's project assets (namely imaging data) and the wind farm operator. Therefore, all the data which does not fit this cases will be publicly available. This includes imaging data (images, videos and sonar data) from the coastal testbed that will be deployed in WP3, as well as other generic data collected by onboard sensors and estimated by its onboard systems. All public data will be released after the completion of the experimental demonstrations to ensure high quality through validation.

In particular, we have the types of data discussed below.

Robotic platforms telemetry data

Type	Format	Size
Navigation data	CSV, binary	100's MB
Sensors output data	CSV, binary, proprietary, MP4, JPG	GBs
Generic data from ROS-based platforms	BAG	GBs

Purpose of the data collection/generation: (1) To provide operators at the SCC with situational awareness of the operations in near real time, (2) to enable future analysis of data for improving behaviour of each robot subsystem and the robot as a whole, (3) to allow the development of new data fusion or machine learning techniques. This data will be produced in all simulations and field trials that take place during the project.

This project has received funding from the European Union's Horizon 2020 research and innovation programme, under the Grant Agreement no. 871571.

Relation to the objectives of the project: This data contributes to one of the main project objectives, which is to establish robotic assets as an indispensable tool when performing IMR operations. It helps maximize the integrity of the target structure and the safety and efficiency of the IMR process.

Re-use of existing data: This data can be used for simulation purposes in order to improve perception and navigation systems, for example, and overall mission procedures.

Origin of data: This data will be gathered by all the technological partners that will deploy robotic assets with the corresponding sensing capabilities during offshore testing. All the data related to the robotic platforms and its sensors may be generated through simulation, in addition to the real field experiments. Simulation data, if not violating any IPR, is open by default.

Data utility: The results of equipment performance will be used and treated to assess the impact of such technologies to the LCOE of the offshore wind production. This data can also be useful for entities that are interested in deploying robotic assets for performing similar inspection operations, namely in the offshore wind farm industry. The data gathered here can be used by the research and industrial community to foster the development of algorithms and new systems (e. g. new data processing techniques and sensing technologies) to increase autonomy and performance of the robots during IMR operations.

Documentation

Type	Format	Size
Project deliverables	PDF, DOCX	10's MB
Scientific publications	PDF	10's MB
Software library	.LIB/.DLL	10's MB

Purpose of the data collection/generation: Documenting all the information gathered during the project, including from field experiments, as well as the analysis of the open data and all the WP deliverables. This information is produced throughout the duration of the project. The provided software library will be developed to allow interoperability between the existing robots.

Relation to the objectives of the project: This data contributes to establishing robotic assets as an indispensable tool when performing IMR operations in challenging environments and to assure that scientific and technical progress takes place within the scope of the project.

Re-use of existing data: May be used for defining an optimal maintenance schedule in the future. Historical data and knowledge can be used to adjust or redesign components/equipment/systems/ procedures to improve reliability and efficiency.

This project has received funding from the European Union's Horizon 2020 research and innovation programme, under the Grant Agreement no. 871571.

Origin of data: This data will be produced with the input of all partners. The data produced here is dependent on the imaging data acquired during the field experiments and IMR operations.

Data utility: Same as above.

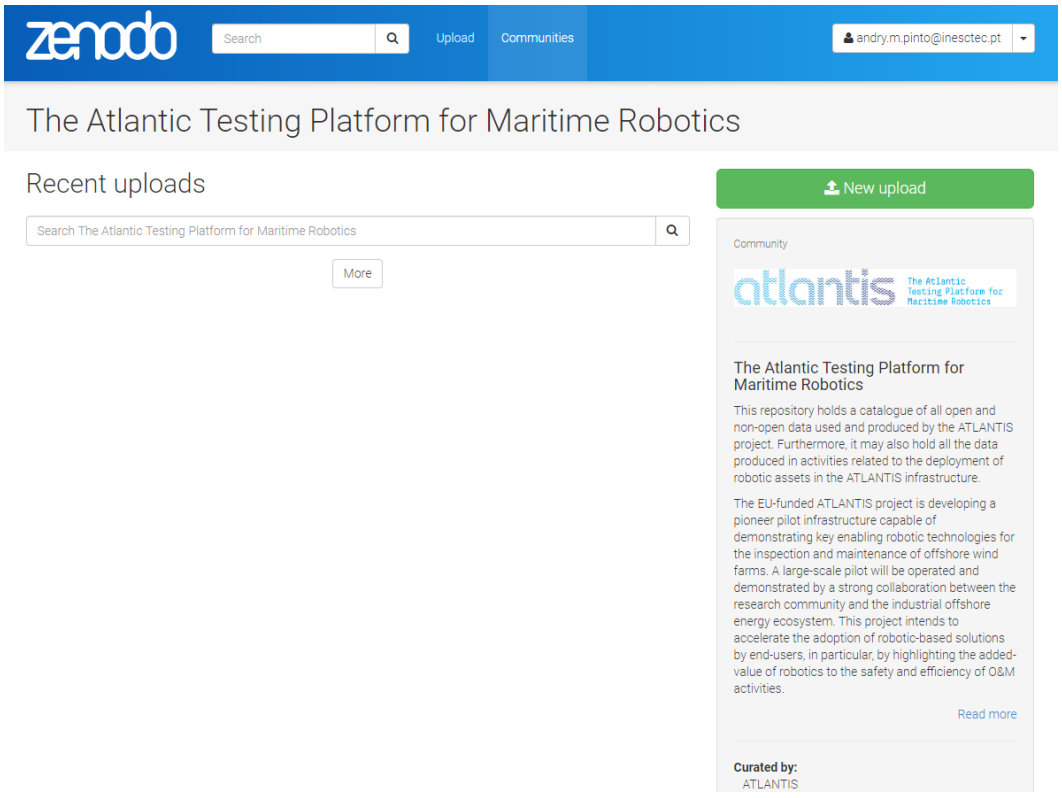
3.2.Workflow

The following diagrams describe the general data management workflow for uploading datasets to ATLANTIS's community page at Zenodo and linking the corresponding meta-data in ATLANTIS 's OpenAIRE project page (which was created when The ATLANTIS project entry was added in CORDIS).

3.2.1. Uploading a dataset

New users that want to register datasets in ATLANTIS 's Zenodo community page have to create an account (https://zenodo.org/signup/). The authorization for adding datasets to the ATLANTIS community page will then be granted by the page administrator. Users just interested in viewing the meta-data stored in the catalogue and downloading the linked and public datasets from the public data repositories do not need an account. As displayed in figure 1, in order to upload data to Zenodo, just access the ATLANTIS community page (https://zenodo.org/communities/atlantis/).

Figure 2 shows the menu presented to the user. After selecting the data to upload, filling in all the information (described previously in 2.3.2) and publishing it, the data will be visible at the ATLANTIS community page and available for everyone to check (depending obviously on the specified access restrictions). When data is uploaded to Zenodo, the meta-data is automatically and immediately indexed at OpenAIRE (see figure 3).



This project has received funding from the European Union's Horizon 2020 research and innovation programme, under the Grant Agreement no. 871571.

Figure 1 - ATLANTIS community page at Zenodo (*https://zenodo.org/communities/atlantis/*). To add a dataset, click the green button that states "New upload".

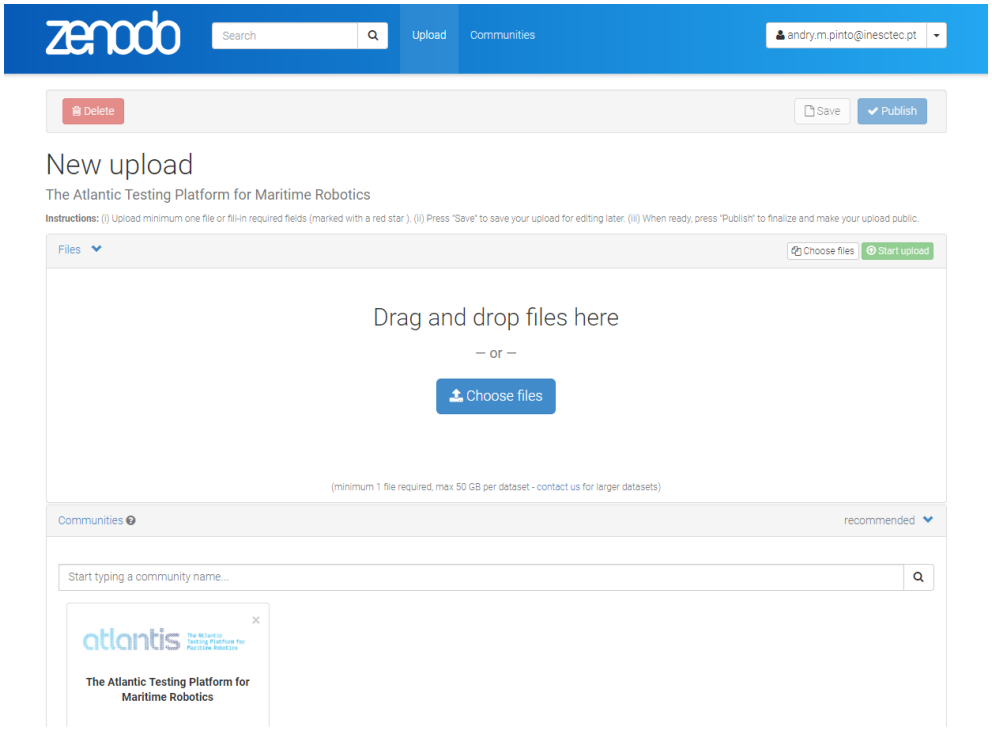
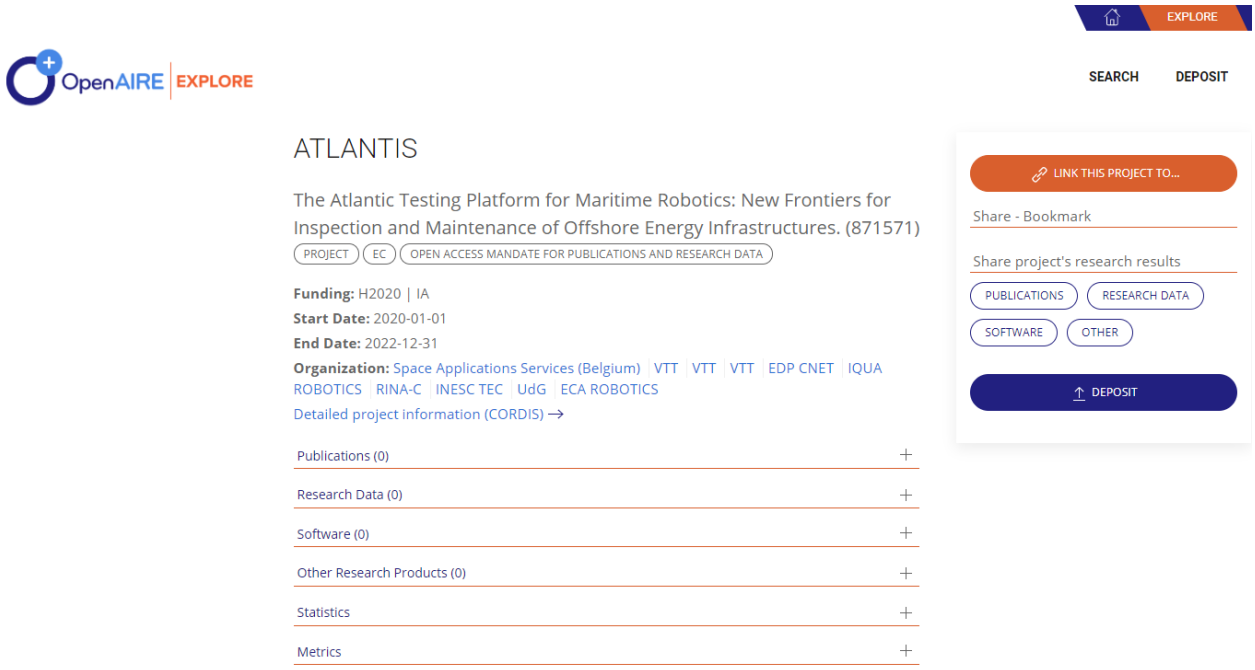


Figure 2 - Choose the files to upload and type all the information describing the data (meta-data), according to OpenAIRE guidelines (described in 2.3.2). Note that it is possible to define access rights to this data: it can be open, embargoed, restricted or closed access. When all the information is typed just press the "Publish" button.



This project has received funding from the European Union's Horizon 2020 research and innovation programme, under the Grant Agreement no. 871571.

Figure 3 - ATLANTIS project page at OpenAIRE (https://explore.openaire.eu/search/project?projectId=corda__h2020::b575ec66952093a175c6e26e9cf86713project). It aggregates all information related to the project. The meta-data from the data uploaded to Zenodo will be visible here.

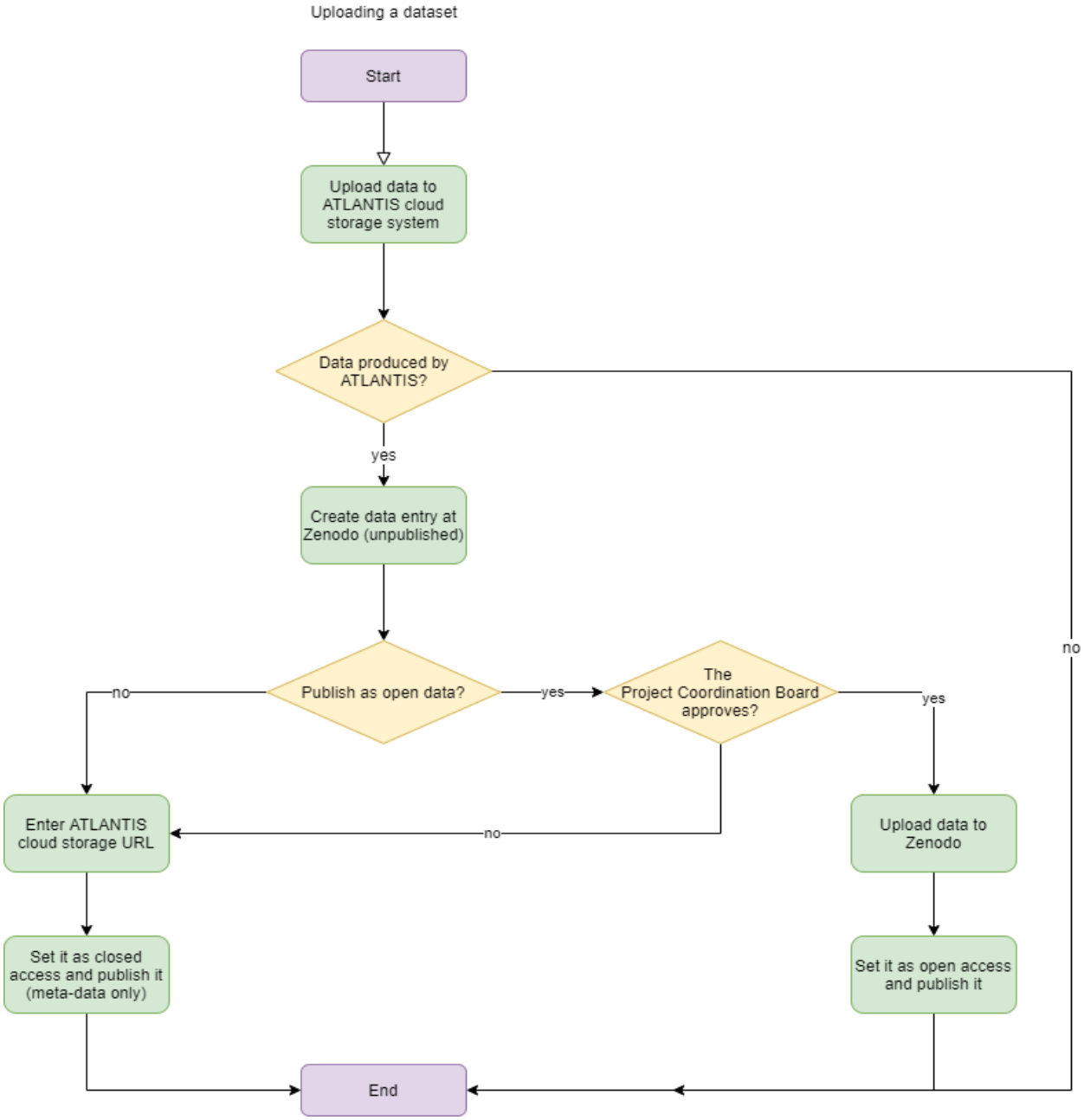


Figure 4 - Flowchart describing the process of uploading a dataset to Zenodo. The task leader is responsible for executing this workflow with the intervention of the PCB.

This project has received funding from the European Union's Horizon 2020 research and innovation programme, under the Grant Agreement no. 871571.

3.2.2. Linking meta-data

If, for data management purposes, it is desired to create an entry to a closed or restricted access dataset at Zenodo (although the actual data would be stored in the local cloud storage system), then just upload an empty text file and type the URL to the dataset stored in the restricted access storage location in "Related/alternate identifiers" field and publish it.

If for some reason (e. g. data not added though Zenodo or scientific publication stored at the publisher servers) we want to link available meta-data to ATLANTIS OpenAIRE project page, then just go to this page and click on the "Link this project to..." button, visible in figure 5.

The user will be presented with a search input field and, after typing the desired keywords, the platform will compile search results from four databases (OpenAIRE, Crossref, DataCite and ORCID). To add any search result to the project, just click the "+" sign.

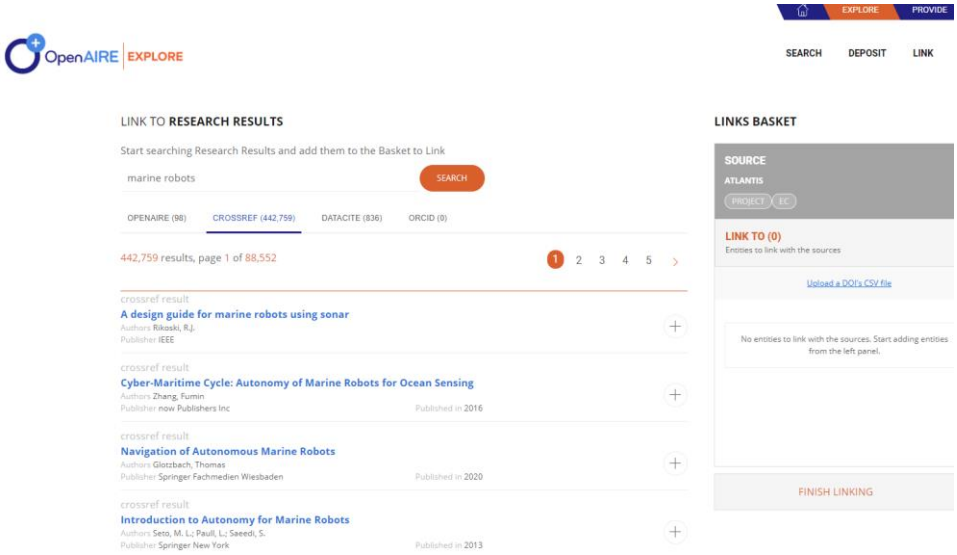


Figure 5 - Linking search results to the project in the OpenAIRE platform.

This project has received funding from the European Union's Horizon 2020 research and innovation programme, under the Grant Agreement no. 871571.

4. Conclusion

The ATLANTIS project makes use of a diverse amount of open data offered by the scientific community, European agencies and EU-funded projects and initiatives, as well as data provided by the partners involved in the project.

The ATLANTIS project will provide a valuable data repository, including detailed documentation related to wind farm IMR operations and raw and processed data from the robotic assets (and its sensors) available in the project. Data from field trials using the robotic assets in the ATLANTIS coastal testbed will be made available to interested parties that are involved in IMR operations or that are interested in testing their robots in this testbed. Sharing of data related to trials in the wind farm will only take place with explicit consent from the PCB and any external entities responsible for or connected to the generation of such information, under an NDA agreement. Industrial partners participating on ATLANTIS project may offer restrictions on the data availability for sharing and use due to the confidential nature of their business.

One of the main aims of data management procedures established in ATLANTIS is to ensure that data that is produced in the context of the project and that is not subject to IPR, commercial exploitation or access restrictions can be openly published following the FAIR (findable, accessible, interoperable and reusable) principles. In particular, open data that can be used by third parties, as entities working in the offshore wind industry or companies that perform IMR operations using robotic assets, to generate new beneficial results, including new open data. Accordingly, such open data is one of the sustainable results of the project and is/will be deposited for long term preservation in the Zenodo research data repository (https://zenodo.org/communities/atlantis/). The data will be stored, archived and secured on local and cloud based services for the duration of the project and following the project. Due to the large data volume and complex scenarios involved, the retrieval and use of the data by those interested will require some support and training provided by the ATLANTIS project members.

The ATLANTIS DMP reports on data used and produced within the ATLANTIS project, including open and non-open data. This DMP can be subject to an update to account for other types of data that may be used or produced throughout the entire course of the project.

5. Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement ID 871571.

6. References

[1] "Data management", HORIZON H2020 online manual, European Commission, 2020.

https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm

This project has received funding from the European Union's Horizon 2020 research and innovation programme, under the Grant Agreement no. 871571.

[2] “Open access & Data management”, HORIZON H2020 online manual, European Commission, 2020.

https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-dissemination_en.htm

[3] “OpenAIRE guidelines”, OpenAIRE, 2020.

https://guidelines.openaire.eu/en/latest/

This project has received funding from the European Union’s Horizon 2020 research and innovation programme, under the Grant Agreement no. 871571.