



**Water-tightness Airborne Detection Implementation**

## **Deliverable 10.2: Data Management Plan**

---

**Authors: Anabela Oliveira (LNEC); Elsa Alves (LNEC); Jorge Costa (LNEC); Joaquim Neto (LNEC)**



*This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 689239.*

## Deliverable 10.2: Data Management Plan

### Technical references

Project Acronym	WADI
Project Title	Water-tightness Airborne Detection Implementation
Project Coordinator	Elena Gaboardi, youris.com (YOURIS) elena.gaboardi@youris.com
Project Duration	October 2016 – March 2020 (42 months)
Deliverable No.	D10.2
Dissemination level*	PU
Work Package	WP 10 – Communication and Dissemination
Task	T10.1 - Communication, Dissemination and Data Management Plan
Lead beneficiary	LNEC
Contributing beneficiary/ies	1 - youris.com, 2 - ONERA, 3 - Air Marine, 4 – LNEC, 5 – EDIA, 6 - Société du Canal de Provence (SCP), 7 – NTGS, 8 – CIRCE, 10 - SST Consult, 11 - Galileo Geosystems, 12 - SGI Studio Galli Ingegneria
Due date of deliverable	31 march 2017
Actual submission date	V1.0 31 March 2017; V1.1 31 October 2017

*PU = Public*

*PP = Restricted to other programme participants (including the Commission Services)*

*RE = Restricted to a group specified by the consortium (including the Commission Services)*

*CO = Confidential, only for members of the consortium (including the Commission Services)*

## Deliverable 10.2: Data Management Plan

v	Date	Beneficiary	Author
1.0	31/03/2017	youris.com	Elena Gaboardi; Alice De Ferrari
1.0	31/03/2017	ONERA	Christian Chatelard; Yves Michel Frederic; Yolande Louvet
1.0	31/03/2017	Air Marine	Vincent Olichon; Christophe Mazel
1.0	31/03/2017	LNEC	Anabela Oliveira; Elsa Alves; Jorge Costa; Joaquim Neto; Ana Mendes
1.0	31/03/2017	EDIA	Alexandra Carvalho
1.0	31/03/2017	Société du Canal de Provence	Isabelle Le Goff
1.0	31/03/2017	NTGS	Julio Estrella Molina
1.0	31/03/2017	CIRCE	Fernando Cirez
1.0	31/03/2017	SST Consult	David Toft
1.0	31/03/2017	Galileo Geosystems	Juan Barba Polo; Erik de Badts; Javier Sanchis
1.0	31/03/2017	SGI Studio Galli Ingegneria	Marino Balzarini
1.1	18/10/2017	LNEC	Anabela Oliveira, Elsa Alves
1.1	31/10/2017	YOURIS	Elena Gaboardi, Alessandra Barbieri, Veronica Meneghello

Deliverable 10.2: Data Management Plan



**Disclaimer**

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 689239.

The sole responsibility for the content of this report lies with the authors. It does not necessarily reflect the opinion of the European Union. The European Commission is not responsible for any use that may be made of the information contained therein.

## Executive Summary

This document presents the Initial Data Management Plan (DMP) for the WADI project. WADI has chosen to participate in the extended Open Research Data pilot. Following the recommendations provided by the European Commission (EU, 2016), in the scope of making data FAIR, this DMP provides an initial approach to the following topics:

- the handling of research data during and after the end of the project
- what data will be collected, processed and/or generated
- which methodology and standards will be applied
- whether data will be shared/made open access
- how data will be curated and preserved (including after the end of the project)

The WADI DMP was prepared using the Digital Curation Centre (DCC) DMP online tool (<https://dmponline.dcc.ac.uk/>), which provides a DMP template that match the demands and suggestions of the Guidelines on FAIR Data Management in Horizon 2020 (EU, 2016).

## Table of Content

1	Introduction.....	7
2	Data summary .....	8
3	FAIR data.....	17
3.1	Making data findable, including provisions for metadata .....	17
3.2	Making data openly accessible .....	18
3.3	Making data interoperable .....	19
3.4	Increase data re-use .....	19
4	Allocation of resources .....	20
5	Data security.....	21
6	Ethical aspects.....	22
7	References.....	23

## 1 Introduction

This document is developed as part of the WADI (Water-tightness Airborne Detection Implementation) project, which has received funding from the European Union's Horizon 2020 Research and Innovation programme, under the Grant Agreement number 689239. The Data Management Plan (DMP), integrated in Task 10.1 (Communication, Dissemination and Data Management Plan), represents the Deliverable 10.2 of Work Package10 (WP10) – Communication and dissemination. WADI has chosen to participate in the extended Open Research Data pilot.

The project WADI includes eleven work packages from which seven of them (WP2 to WP7 and WP9) will result in relevant information to be shared with the research and end-user communities. The purpose of the DMP is to support the data management life cycle for all data that will be collected, processed or generated by the WADI project. Following the recommendations provided by the European Commission (EU, 2016), in the scope of making data FAIR (findable, accessible, interoperable and reusable), this DMP provides an initial approach to the following topics:

- the handling of research data during and after the end of the project
- what data will be collected, processed and/or generated
- which methodology and standards will be applied
- whether data will be shared/made open access
- how data will be curated and preserved (including after the end of the project)

More developed versions of the Plan will be released at later stages.

The WADI DMP was prepared using the Digital Curation Centre (DCC) DMP online tool (<https://dmponline.dcc.ac.uk/>), which provides a DMP template that match the demands and suggestions of the Guidelines on FAIR Data Management in Horizon 2020 (EU, 2016).

## 2 Data summary

The main objective of WADI is to develop an innovative airborne water leak detection surveillance service to provide water utilities with adequate information on leaks in water infrastructures outside urban areas (rural areas) and optimise their performance in this field. Throughout the project, data will be produced and reused to support the development of research and innovation activities focusing on: defining end-user requirements for water leak surveillance services (WP2); defining system requirements including optical device coupling; optimisation and their application on aerial platforms (WP3); developing cost-effective data processing techniques, reliable data processing and a new interface for end-users (WP4); evaluating WADI's service feasibility using a performance matrix and its environmental and economic impacts (WP7); validating and demonstrating on both French and Portuguese sites (WP5, WP6), conducting a comprehensive legal and regulatory analysis (WP8); as well as developing pertinent market studies, marketing strategies, and business plans (WP9). Table 1.1 summarises the types of data produced / made available in different work packages of the WADI project. Table 1.2 presents a more detailed description of data and its use (or re-use), formats, standards and metadata, data availability (open, confidential) and the expected resources to store, curate and preserve (size, backup frequency, resources to maintain it after project ends, repository choices for open and private data).

	TYPE OF DATA	WP
WADI surveillance system development	Hyper-spectral/IR image database	WP3
	Optimized detection wavelengths	WP3
	Preliminary data from the platform for data processing development	WP3, WP4
	Imagery data for tests and validation	WP3, WP5, WP6
	Flight data of WADI process tests	WP5
	Flight data of WADI operational and surveillance tests	WP6
	Multi-spectral/IR image database	WP7
End-users demonstration sites information	Flights results, frequency, instrumentation used, fuel consumed, renting, staff costs	WP7
	Network maps	WP3, WP4, WP5, WP6, WP7, WP9
	Infrastructure data base	WP5, WP6, WP7
	Water company operational info	WP2, WP7
	Flow data	WP5, WP6, WP7
	Water losses data, leaks localisation	WP2, WP5, WP7
General information	Ground leak detection investigation	WP5, WP6, WP7
	Environmental data from satellites and other open sources	WP7
	Socioeconomic public reports analysis	WP7

Table 1. 1 Summary of types of data to be used / created in different work packages (WP)

Deliverable 10.2: Data Management Plan

Partner	Work Package	Type of data/ data description	Methodology of data production	Data documentation and how that will be made available	Data standards	Which data sets will be classified as Open Access	Which data sets will be classified as confidential	In case of datasets that are not shared, reasons for that	To whom data classified as Open Access could be useful	How will open data be shared / by whom and where	How will confidential data be archived / by whom and where	How is backup and versioning realised?	When data will be produced	When data will be placed in Open Access	How will the data be disseminated	How will data be available after the end of the project
OFFICE NATIONAL D'ETUDES ET DE RECHERCHES AEROSPATIALES (ONERA)	WP3 (task3.1)	Hyper-spectral/IR image database	From Airborne measurement	Not available	Standards are not available for this type of data. Data formats are owner format of the suppliers : NEO for hyperspectral camera and FLIR for IR camera	Not applicable	All	Data too complex for others to use it, data sets have very large size, and they cannot be exploitable without specific software owned by ONERA	Not applicable	Not applicable	By ONERA in its own servers	Process done by ONERA using their own resources and procedures	Data will be produced during WP 3-1	Not applicable	Not applicable	Data archive on ONERA server and repositories
	WP3 (task3.1)	Optimized detection wavelengths	Exploitation of database image	Available data will be reported in WADI report	Not applicable, wavelength standards do not exist	All data is open access	None	Not applicable	Researchers, end-users, service providers	Can be shared to all at the end of the WP3-1, through WADI repository at ZENODO	Not applicable	Process done using ZENODO's own procedures	Data will be produced during WP 3-1	At the end of the WP3	In conferences presentations	Through publication and available through ZENODO during the period that this repository offers for H2020 projects
	WP7 (task 3.1)	Multi-spectral/IR image database	Through Airborne measurements achieved on WP5 and WP6	WADI report	Standards will be defined later in WP4 with NTGS and LNEC partner	Images database	Not applicable	Not applicable	Researchers, end-users, service providers	Can be shared to all at the end of the WP7, using WADI repository at ZENODO	Not applicable	Process done using ZENODO's own procedures	Database will be produced during WP7	At the end of WP7	During WP10 process	Through publication and available through ZENODO during the period that this repository offers for H2020 projects

Table 1. 2Description of data to be used / created in different work packages, including standards and metadata, availability, storing resources for open and confidential data

Deliverable 10.2: Data Management Plan

Partner	Work Package	Type of data/ data description	Methodology of data production	Data documentation and how that will be made available	Data standards	Which data sets will be classified as Open Access	Which data sets will be classified as confidential	In case of datasets that are not shared, reasons for that	To whom data classified as Open Access could be useful	How will open data be shared / by whom and where	How will confidential data be archived / by whom and where	How is backup and versioning realised?	When data will be produced	When data will be placed in Open Access	How will the data be disseminated	How will data be available after the end of the project
AIR MARINE SARL (AIR MARINE)	WP5 (task 5.2)	Flight data of WADI process tests (validation of Water leak Airborne Detection on SCP infrastructure)	GIS output: data of the mission Onboard mission recording: raw data sensor output and navigation sensor output (synchronised metadata)	Internal technical document, proprietary repository	SHAPE format + Images (geotiff)	None	All: WADI consortium restricted	Flight data is protected and will be provided to the consortium partners on demand.	End-users, researchers, data providers,...	Not applicable	Confidential data is managed using Air Marine resources	Backup and versioning data is managed using Air Marine resources	Data produced according to the project schedule (demonstration flights in 2018)	Not applicable	Not applicable	Data archive on Air Marine server and repositories
	WP6 (task 6.2)	Flight data of WADI operational and surveillance tests (validation of Water leak Airborne Detection on EDIA infrastructure)	GIS output: data of the mission Onboard mission recording: raw data sensor output and navigation sensor output (synchronised metadata)	Internal technical document, proprietary repository	SHAPE format + Images (geotiff)	None	All: WADI consortium restricted	Flight data is protected and will be provided to the consortium partners on demand.	End-users, researchers, data providers,...	Not applicable	Confidential data is managed using Air Marine resources	Backup and versioning data is managed using Air Marine resources	Data produced according to the project schedule (demonstration flights in 2018)	Not applicable	Not applicable	Data archive on Air Marine server and repositories
LABORATORIO NACIONAL DE ENGENHARIA CIVIL (LNEC)	WP7 (task 7.3)	Infrastructure data for reliability analysis (water levels, discharge flows,...)	Field data acquisition	Internal end-users reports and databases. Will be provided by EDIA and SCP on request.	Not applicable	None	All: Infrastructure data for reliability analysis (water levels, discharge flows,...)	According to the Consortium agreement, infrastructure data is protected and will be provided to the partners on demand.	Not applicable	Not applicable	Infrastructure data is already managed by the end-users using their own resources	Not applicable	Data already exists. It will be used in WADI as part of the reliability data processing.	Not applicable	Not applicable	Data archive on LNEC server and repositories
	WP7 (task 7.3)	Environmental data from satellites and other open sources	Remote sensing and field data acquisition	In open repositories on the internet (Copernicus) and provided by other end-users (e.g. Water authority in Portugal)	Satellite data: standards used by Copernicus; Water authorities: distinct data formats and standards	All	Not applicable	Not applicable	End-users, researchers, general public,...	Available in open repositories	Not applicable	Procedures done by the data owners with their own resources	Data already exists. It will be used in WADI as part of the reliability data processing.	Already available	Through websites	Depending on data owner policies

Table 1. 2 Description of data to be used / created in different work packages, including standards and metadata, availability, storing resources for open and confidential data(continuation)

Deliverable 10.2: Data Management Plan

Partner	Work Package	Type of data/ data description	Methodology of data production	Data documentation and how that will be made available	Data standards	Which data sets will be classified as Open Access	Which data sets will be classified as confidential	In case of datasets that are not shared, reasons for that	To whom data classified as Open Access could be useful	How will open data be shared / by whom and where	How will confidential data be archived / by whom and where	How is backup and versioning realised?	When data will be produced	When data will be placed in Open Access	How will the data be disseminated	How will data be available after the end of the project
EDIA-EMPRESA DE DESENVOLVIMIENTO E INTRA-ESTRUTURAS DO ALQUEVA	WP3 and WP6	Digital data about EDIA network localisation	Working drawings	Internal documentation , available through request to EDIA project team, internet EDIA website	PDF maps, SHP files or other GIS format	All not included as restricted in the consortium agreement	Those included as restricted in the consortium agreement	Operational data from EDIA	Partners, public	By EDIA through its website	By EDIA in EDIA server	Using EDIA own resources	In website already available	After the analysis of WP6 flights (2019)	EDIA website	Depending on EDIA owner policies
	WP6	Flow data at specific locations	Flow meters	Internal documentation , available through request to EDIA project team	Distinct data formats and standards	Restricted in the consortium agreement	Those included as restricted in the consortium agreement	Operational data from EDIA	Partners, public	Not applicable	By EDIA in EDIA server	Using EDIA own resources	Data already exists	Not applicable	Not applicable	Depending on EDIA owner policies
SOCIETE DU CANAL DE PROVENCE ET D'AMENAGEMENT DE LA REGION PROVENCALE SA (SCP)	WP3 and WP5	Digital data about SCP network localisation	Working drawings	Internal documentation , available through request to SCP project team, or available at internet SCP website	PDF maps, SHP files or other GIS format	Data sets on localisation on 1/25000 scale without associated data (size, material, age...)	All others	Operational data from SCP	Partners, public	SCP internet website	SCP	Using SCP own resources	Data already exists	Data already exists	SCP internet website	SCP internet website
	WP5	Leaks localisation	Analysis of remote sensing acquisition	cf. partners (SGI, GG) who will produce this kind of data	cf. partners (SGI, GG) who will produce this kind of data	Data sets on localisation on 1/25000 scale in general map (without precise localisation)	Data sets with more precise presentation	Communication management with local residents	Partners, public	Not applicable	SCP, on our GIS and WADI partner who will produce the data	SCP, on our GIS and by WADI partners who will produce the data	During WP5	After the analysis of WP5 flights (2018)	SCP internet website	SCP internet website

Table 1. 2 Description of data to be used / created in different work packages, including standards and metadata, availability, storing resources for open and confidential data(continuation)

Deliverable 10.2: Data Management Plan

Partner	Work Package	Type of data/ data description	Methodology of data production	Data documentation and how that will be made available	Data standards	Which data sets will be classified as Open Access	Which data sets will be classified as confidential	In case of datasets that are not shared, reasons for that	To whom data classified as Open Access could be useful	How will open data be shared / by whom and where	How will confidential data be archived / by whom and where	How is backup and versioning realised?	When data will be produced	When data will be placed in Open Access	How will the data be disseminated	How will data be available after the end of the project
NEW TECHNOLOGIES GLOBAL SYSTEMS SL (NTGS)	WP3 (task 3.2)	Preliminary data from the platform for data processing development	Engineering	Internal reports. Data will be provided by NTGS on request	PDF maps	none	All	According to the Consortium agreement, infrastructure data (imagery) is protected and will be provided to the partners on demand.	Not applicable	Not applicable	Data owners in their own repositories	NTGS own backup	M18	Not applicable	Not applicable	Not applicable
	WP4 (task 4.3)	Preliminary data from the platform for data processing development	Design	Internal reports. Data will be provided by NTGS on request	PDF maps	none	All	According to the Consortium agreement, infrastructure data (imagery) is protected and will be provided to the partners on demand.	Not applicable	Not applicable	Data owners in their own repositories	NTGS own backup	M19	Not applicable	Not applicable	Not applicable
FUNDACION CIRCE CENTRO DE INVESTIGACION DE RECURSOS Y CONSUMOS ENERGETICOS (CIRCE)	WP7 (task 7.1)	Infrastructure information for building inventories for both demos: Pipes info, pumping consumption, chemical consumption, Flights results, Frequency, instrumentation used, fuel consumed, renting, staff costs	Information will be obtained from WP 2	Data bases will be internally build and stored according to WP 2 information.	As in WP2	None, unless end-users (EDIA and SCP) allow sharing the info	All data will be considered as confidential unless end-users (EDIA and SCP) allow sharing the info	According to the Consortium agreement, infrastructure data is protected	Not applicable	Not applicable	Internally in databases (included in SimaPro software)	Backup will be done weekly according to CIRCE procedures	Databases will be developed during the WP7 framework, once WP2 tasks will start	Not applicable	Not applicable	Not applicable
	WP7 (task 7.2)	Environmental data from satellites and other open sources and socioeconomic public reports analysis	Type of ecosystem will be defined according to MAES methodology internally, according to the information obtained from demo sites definition	In open repositories on the internet (Copernicus) and provided by other end-users (e.g. Water authority in Portugal and France)	Water authorities: distinct data formats and standards	All	None	Not applicable	End-users, researchers, local inhabitants	By means of open repositories	Not applicable	Backup will be done weekly according to CIRCE procedures	Databases will be developed during the WP7 framework, once WP2 tasks will start	At the end of the project	By means of project website	By means of project website

Table 1. 2 Description of data to be used / created in different work packages, including standards and metadata, availability, storing resources for open and confidential data(continuation)

Deliverable 10.2: Data Management Plan

Partner	Work Package	Type of data/ data description	Methodology of data production	Data documentation and how that will be made available	Data standards	Which data sets will be classified as Open Access	Which data sets will be classified as confidential	In case of datasets that are not shared, reasons for that	To whom data classified as Open Access could be useful	How will open data be shared / by whom and where	How will confidential data be archived / by whom and where	How is backup and versioning realised?	When data will be produced	When data will be placed in Open Access	How will the data be disseminated	How will data be available after the end of the project
SST-CONSULT ADAM STACHEL, RAFAL STANEK, DAVID ANDREW TOFT (SST - Consult)	WP9 (task 9.1)	Data on WADI results obtained from validation (WP5&6), results (WP 7) and legal analysis (WP 8). Desk study for market conditions.	Data on WADI results obtained from validation (WP5&6), results (WP 7) and legal analysis (WP 8)	Internal project documentation and for communication with funders. Not publicly available directly from project.	Distinct data formats and standards	Restricted in the consortium agreement	Those included as restricted in the consortium agreement	Operational data from end-users	Partners, public	Not applicable	In end-users servers	Using end-users own resources	During WP9	Not applicable	Not applicable	Data archive on end-users server and repositories
	WP9 (task 9.2)	Data on WADI results obtained from validation (WP5&6), results (WP 7) and legal analysis (WP 8). Desk study for market conditions.	Data on WADI results obtained from validation (WP5&6), results (WP 7) and legal analysis (WP 8)	Internal project documentation and for communication with funders. Not publicly available directly from project.	Distinct data formats and standards	Restricted in the consortium agreement	Those included as restricted in the consortium agreement	Operational data from end-users	Partners, public	Not applicable	In end-users servers	Using end-users own resources	During WP9	Not applicable	Not applicable	Data archive on end-users server and repositories
	WP9 (task 9.3)	Based on market strategy	Based on market strategy	Based on market strategy	Distinct data formats and standards	Restricted in the consortium agreement	Those included as restricted in the consortium agreement	Operational data from end-users	Partners, public	Not applicable	In end-users servers	Using end-users own resources	During WP9	Not applicable	Not applicable	Data archive on end-users server and repositories

Table 1. 2 Description of data to be used / created in different work packages, including standards and metadata, availability, storing resources for open and confidential data(continuation)

Deliverable 10.2: Data Management Plan

Partner	Work Package	Type of data/ data description	Methodology of data production	Data documentation and how that will be made available	Data standards	Which data sets will be classified as Open Access	Which data sets will be classified as confidential	In case of datasets that are not shared, reasons for that	To whom data classified as Open Access could be useful	How will open data be shared / by whom and where	How will confidential data be archived / by whom and where	How is backup and versioning realised?	When data will be produced	When data will be placed in Open Access	How will the data be disseminated	How will data be available after the end of the project
Galileo Geosystems S.L. (GG)	WP3 (task 3.3)	Imagery data for tests and validation	RPAS data acquisition	Will be provided by GG on request	End-users standards for geographic information	None	All: high resolution imagery from infrastructures and preliminary data report	According to the Consortium agreement, infrastructure data (imagery) is protected and will be provided to the partners on demand.	End-users, researchers	Available in restricted repositories	High resolution imagery archived by end-users	Procedures done by end-users with their own resources	M18	Not applicable	Not applicable	Not applicable
	WP5 (task 5.2)	Imagery data for tests and validation	RPAS data acquisition	Will be provided by GG on request	End-users standards for geographic information	None	All: high resolution imagery from infrastructures and preliminary data report	According to the Consortium agreement, infrastructure data (imagery) is protected and will be provided to the partners on demand.	End-users, researchers	Available in restricted repositories	High resolution imagery archived by end-users	Procedures done by end-users with their own resources	M23	Not applicable	Not applicable	Not applicable
	WP6 (task 6.2)	Imagery data for tests and validation	RPAS data acquisition	Will be provided by GG on request	End-users standards for geographic information	None	All: high resolution imagery from infrastructures and preliminary data report	According to the Consortium agreement, infrastructure data (imagery) is protected and will be provided to the partners on demand.	End-users, researchers	Available in restricted repositories	High resolution imagery archived by end-users	Procedures done by end-users with their own resources	M31	Not applicable	Not applicable	Not applicable
SGI STUDIO GALLI INGEGNERIA SRL (SGI)	WP1	Commercial (expectations from WADI)	Questionnaires	Aggregated analysis, rough data; internal repository	None	None	All	Either part of Partners' (EDIA and SCP) commercial strategy and competitiveness or protected according to the Consortium Agreement	Not applicable	Not applicable	Questionnaires archived by end-users and SGI	Not applicable	Within M2	Not applicable	Not applicable	Not applicable
	WP2	Water losses data, water company operational info	Questionnaires	Aggregated analysis; internal repository	None	None	All	Either part of Partners' (EDIA and SCP) commercial strategy and competitiveness or protected according to the Consortium Agreement	Not applicable	Not applicable	Questionnaires archived by end-users and SGI	Not applicable	Within M4	Not applicable	Not applicable	Not applicable

Table 1. 2 Description of data to be used / created in different work packages, including standards and metadata, availability, storing resources for open and confidential data(continuation)

Deliverable 10.2: Data Management Plan

Partner	Work Package	Type of data/ data description	Methodology of data production	Data documentation and how that will be made available	Data standards	Which data sets will be classified as Open Access	Which data sets will be classified as confidential	In case of datasets that are not shared, reasons for that	To whom data classified as Open Access could be useful	How will open data be shared / by whom and where	How will confidential data be archived / by whom and where	How is backup and versioning realised?	When data will be produced	When data will be placed in Open Access	How will the data be disseminated	How will data be available after the end of the project
SGI STUDIO GALLI INGEGNERIA SRL (SGI)	WP5	Network maps, infrastructure dB, flow measurements, GLD investigation	The pilot area/s will be identified with SCP. SCP will provide network maps, infrastructure characteristics and flow data through GIS, dB, xls or other media according to availability. Investigation will be planned by SGI together with SCP, who will provide gauged data accordingly. SGI will process investigation data and submit relevant report.	Documented through maps, spreadsheets, dB	To be defined	To be defined	To be defined	Operational data from Water Companies	Anyone interested in direct check of WADI performance	WADI repository at ZENODO	Internal WADI project repository	To be defined	M20-M23	M23	Training, public reports, conferences	Reports, conference / training papers, dissemination material
	WP6	Network maps, infrastructure dB, flow measurements, GLD investigation	The pilot area/s will be identified with EDIA. EDIA will provide network maps, infrastructure characteristics and flow data through GIS, dB, xls or other media according to availability. Investigation will be planned by SGI together with EDIA, who will provide gauged data accordingly. SGI will process investigation data and submit relevant report.	Documented through maps, spreadsheets, dB	To be defined	To be defined	To be defined	Operational data from Water Companies	Who is interested in direct check of WADI performance	WADI repository at ZENODO	Internal WADI project repository	To be defined	M24-M30	M23	Training, public reports, conferences	Reports, conference / training papers, dissemination material

Table 1. 2 Description of data to be used / created in different work packages, including standards and metadata, availability, storing resources for open and confidential data(continuation)



## 3 FAIR data

### 3.1 Making data findable, including provisions for metadata

WADI will produce and reuse a variety of data types, from images to time series and georeferenced information in GIS format, covering a broad range of areas (remote sensing, hydrology, hydraulics,..). Metadata will be produced for all data, using standards when available. Standards such as ISO19115 (<http://rd-alliance.github.io/metadata-directory/standards/iso-19115.html>) or the OGC Sensor Observation Service (SOS) Interface Standard (<http://rd-alliance.github.io/metadata-directory/standards/observations-and-measurements.html>) are expected to be adopted. Consistency between metadata for similar data sets will be sought when standards are not available. Elements to be included in the metadata include a clear description of the data, the institution and person of contact responsible for the data creation, its format, creation date and possible modifications, data units and georeferencing (when applicable) and a number of keywords (metatags). The choice of adequate keywords will be included to promote and ease the discoverability of data. These keywords will include a number of fixed, common keywords in WADI's scientific area and several new, free keywords that can help attract researchers from other areas to use and adapt WADI's results to their scientific fields.

For all open data in the project, ZENODO will be used as the project's open data repository. ZENODO provides Digital Object Identifiers for all data sets, thus guaranteeing that all open data in WADI will have persistent and unique identifiers.

For consistency and promotion of data discovery, consistent naming conventions will also be used and agreed among the partners (to be defined later).

Open access publication will also be sought, with direct links to the underlying data sets deposited in ZENODO.

LNEC will be responsible for uploading data and other items in ZENODO, through the project designated data manager. Each partner will provide the datasets and publications to be integrated in the ZENODO repository to the data manager, dully informed on its access policy. The datasets and the expected time of availability and access policies are described in Table 1.2. For publications subject to embargo periods (due to the publishers' policies), the data manager will upload them in ZENODO as soon as the embargos are finished. The coordinator will inform LNEC as soon as the datasets are delivered to youris.com, for data management monitoring purposes.

In the scope of the surveillance system developed in WADI, a succession of data sets will be produced, creating several databases of images at different stages of development and processing, from the raw data from cameras to the processed and quality-certified

images included in the end-user application. This sequence can be labelled as several versions of a single dataset or it can be identified and managed as different datasets. Regardless of the approach chosen by the partners for this data, a clear versioning policy will be adopted and linked with detailed metadata and supporting documentation.

### 3.2 Making data openly accessible

WADI will create or reuse a variety of data sets, which have different natures and correspondingly distinct access privileges. Part of these privileges was set up in the project's Consortium Agreement. These different access privileges are described in detail in Table 1.2 and are reviewed here in a concise manner. Table 1.2 also provides a detailed description of all aspects related to dataset management. A short overview on data access policies and availability is presented here.

Regarding end-users' infrastructure data (such as digital information regarding the networks' location in very fine detail, the customer's data and the location of leaks), these are classified as confidential in the Consortium agreement to protect personal information (as described in the D11.1 POPD - Requirement No. 3) as well as the security of the companies' assets. Infrastructure data at a granular scale is already openly available at the companies' website, thus fulfilling its usefulness for scientific or other public purposes. Regarding classified data, it is kept at the companies' own repositories, fulfilling their own policies on data backup and preservation, and will be maintained by these entities after the end of the project.

Regarding the data from the surveillance system developed at WADI (Tables 1.1 and 1.2), different access policies are defined according to the Consortium agreement. These data sets include mostly several databases of images at different stages of development and processing, ranging from the raw data derived from the hyperspectral, multi-spectral and I/R cameras to the processed and quality-certified images. The raw data from the several types of cameras have an immense size and require specialized, proprietary software to be accessed. Therefore, it is not made available, although it will be fully documented in reports and scientific publications. In contrast, the optimised detection wavelengths (developed in WP3) and the multi-spectral/IR image database to be developed during the demonstrations in Portugal and France will be openly available through ZENODO.

In-situ ground data is also part of the WADI service, but given the information's sensitivity and the restrictions posed by the infrastructure's competitive operation, it will not be open. Like all confidential data in WADI, its preservation and maintenance during and after the project will be handled by the data owners and/or the end-users.

Deposits in ZENODO will include the data, their metadata and their documentation. For most data sets, access is granted through generalised use software such as ArcGIS or similar.

### 3.3 Making data interoperable

All data developed in WADI will be fully documented and accompanied with detailed metadata supported by a set of select keywords, to facilitate automatic discovery and integration of WADI data for other purposes. Besides usual metadata fields, technical aspects such as units (complying with SI standards) and spatial and temporal references will be supplied. All data will be provided in generally used extensions, adopting well established formats (csv, shapefiles, image formats,...) whenever possible which will also facilitate its use by other parties. The exception will be the raw data from the optical devices which can only be interpreted using ONERA's own software.

### 3.4 Increase data re-use

Open data availability will occur as soon as possible in WADI while respecting the team's publication targets. Typically, open data will be available for publication in ZENODO at the end of respective WP (Table 1.2), and its publication will occur within one month. The team expects that this fast publication of data created in WADI will promote its reuse by other researchers and end-users, thereby contributing to the dissemination of WADI methodology and tools.

The usefulness of the data for third parties is closely linked to the perception of quality and robustness of the available data. Therefore, a dedicated task was defined in WP4 to validate data and standards compliance, setting the stage for the data reliability analysis performed in WP7. In this last analysis, other sources of data will be combined with images to provide a quality index for the leak detection data generated in WADI. All these methods will be expected to contribute to a long term usefulness of this data.

## 4 Allocation of resources

In WADI there is a considerable amount of confidential data (see the aforementioned reasons). This data will be managed by the partners responsible for its creation and/or by the end-users (Table 1.2). Therefore its maintenance, backup and versioning and long term preservation will be guaranteed by their own resources and at their own expense.

A repository in ZENODO was created (<https://zenodo.org/communities/h2020-wadiproject/>) for the projects' open data, therefore ensuring data availability, backup and versioning. Long term preservation will be guaranteed for the lifetime of the ZENODO repository (<https://zenodo.org/policies>). This is currently the lifetime of the ZENODO's host laboratory CERN, which currently has an experimental programme defined at least for the next 20 years. After the end of that period, the data will only be kept at the data owners' servers and repositories.

Publications featuring the data will be produced in the project (specifically by the research partners) and will be made available through open access (using open access journals or journals selected for a short embargo period). This channel will provide a long term availability of data and data analysis.

The partners expect the data to have immense value in short to medium term to support the optimised management of water mains and to reduce water losses. The methodology and data products developed in the project are expected to have a large impact in this field, through generalised application at European and other international sites. Moreover, we expect that the scientific and technological developments may serve as a basis in other fields of application, with the data demonstrating and promoting WADI's exceptional quality of service. The long term usefulness of the data collected during WADI will depend on the technological advances in this area. The resolution achieved with today's technology may prove lacking in the future, given the advances in camera quality and the supporting data's infrastructure.

## 5 Data security

Open data security will be addressed in WADI taking advantage of ZENODO's services of secure storage, backup and preservation and protected transfer mechanisms. Regarding the confidential data, different approaches will be used by each data-owner institution, but common rules apply. Data will be housed on servers under direct management of the institution's personnel to be installed in already provisioned data centres. These data centres are expected to be equipped with various features ranging from secure physical access, air conditioning, generators and fire extinguishing measures. Typically, hardware / electricity failure are addressed with redundant hardware and generators.

Access to data under different permission conditions (read-only, read-write, etc.) are granted to users and authorized computers by project managers or to whomever this task is delegated, according to a well-defined protocol. Confidentiality is assured by additional methods, encryption and anonymization to name a few, depending on their nature and final applications.

Taking in account the size of the data at stake that requires regular backup (be it either for security versus a hardware failure or for archival purposes), a sequence of regular full backups, differential backups and incremental backups on an increasingly frequent basis are envisaged, and following already installed procedures. The physical media used to store the data will be maintained in secure locations. Access to these backups is limited to the personnel authorised to use the backup system, and as a general rule, not authorised for external sources.

All data transfers should be encrypted to render all stolen/lost data useless. Encryption methods are to be specified at a later date.

## 6 Ethical aspects

The WADI partners are to comply with the ethical principles as set out in Article 34 of the Grant Agreement, which states that all activities must be carried out in compliance with:

- a) Ethical principles (including the highest standards of research integrity – as set out, for instance in the European Code of Conduct for Research Integrity – and including, in particular, avoiding fabrication, falsification, plagiarism or other research misconduct)
- b) Applicable international, EU and national law.

Activities raising ethical issues must comply with the “ethics requirements” set out in Annex1 of the Grant Agreement. The WP11 “Ethics requirements” aims to follow-up the ethical issues applicable to the WADI project implementation. It includes: informed consent procedures that will be implemented by each partner that foreseen activities that require such procedures (Deliverable 11.1); administrative clearance procedures and approvals for flights to be carried out in WP3 in France and Portugal (Deliverable 11.2); copies of confirmation/notification by the National data Protection Authority (Deliverable 11.3); and information on the procedures implemented for data collection, storage, protection, retention, destruction and confirmation and compliance with national and EU legislation (Deliverable 11.4).

Legal and regulatory issues also have a dedicated WP (WP8 “Legal and regulatory aspects analysis, including IPR protection”). This WP will provide legal guidance to the project partners, in particular with regard to the authorisations for executing the flight tests in real environment (WP5 and WP6 pilots), the compliance with relevant rules and procedures such as the applicable rules for privacy protection and processing of personal data in relation to the unmanned airborne vehicles use and for all issues with regard to the intellectual property rights of the project results. It will also develop the legal and regulatory framework for the commercialisation of the project results. The WP8 includes: Project legal framework (Deliverable 8.1); WADI service provision legal framework (Deliverable 8.2); WADI service provision terms and conditions (Deliverable 8.3).

## 7 References

European Commission (2016) - H2020 Programme. Guidelines on FAIR Data Management in Horizon 2020. Version 3.0, July 2016.

([http://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/hi/oa\\_pilot/h2020-hi-oa-data-mgt\\_en.pdf](http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf))