

Project Acronym:
Grant Agreement number:
Project Full Title:

Hydroptics
871529 (H2020-ICT-2019-2)
Photonics sensing platform for process optimisation in the
oil industry

 Ref. Ares(2020)2820474 - 01/06/2020



DELIVERABLE

D1.2 – Data management plan v1

Dissemination level	PU – Public
Type of Document	ODPR: Open research Data Pilot
Contractual date of delivery	31/05/2020
Deliverable Leader	Sargis Hakobyan (ALPES)
Status & version	V2.0
WP / Task responsible	WP1 / ALPES
Keywords:	Data, Data management plan

Deliverable Leader:	Organization
Contributors:	Sargis Hakobyan (ALPES)
Reviewers:	Nikos Bakalos (NTUA), Tibor Bereczki (SAL)
Approved by:	Sargis Hakobyan (ALPES)

Document History			
Version	Date	Contributor(s)	Description
V1.0	07/05/2020	ALPES	First complete version
V2.0	26/05/2020	ALPES	Final version, after review by the Coordinator

This document is part of a project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 871529. It is the property of the HYDROPTICS consortium and shall not be distributed or reproduced without the formal approval of the HYDROPTICS Management Committee. The content of this report reflects only the authors' view. EC is not responsible for any use that may be made of the information it contains.

Executive Summary

This document explains how the data will be generated, utilized, saved, and shared between the Partners or published as an open access. Data management plan is an important aspect of the project, hence a careful analysis of the data content and flow has been done. In particular, it has been identified which type of data will be generated, what formats will be used to share the data between the Consortium members, which data can and/or will be shared outside of the Consortium, and which data cannot be shared.

Dedicated data sharing platform is described as well, and a detailed description of the structure and the operation of the last is described.

This is a living document and will evolve during the lifespan of the project.

Table of Contents

Executive Summary	2
Table of Contents	3
1. Introduction	4
2. Types of data generated within the project.....	4
2.1. Administrative data.....	5
2.2. Dissemination and communication data.....	5
2.3. Simulations and Modelling data.....	5
2.4. Experimental data	5
2.5. Field tests data.....	6
3. Data Formats.....	6
4. Data Management plan.....	6
5. Data Sharing	7
5.1. Data sharing platform within the Consortium.....	7
5.1.1. OwnCloud.....	7
5.2. Data sharing platform outside the Consortium.....	9
Conclusions.....	9

1. Introduction

This document deals with the research data produced, collected and preserved during the project. This data can either be made publicly available or not according to the Grant Agreement and to the need of the partners to preserve the intellectual property rights and related benefits derived from project results and activities.

This document will address the following questions:

- What types of data will the project generate/collect?
- What data is to be shared for the benefit of the scientific community?
- What data cannot be made available?
- What format will the shared data have?
- How will this data be exploited and/or shared/made accessible for verification and re-use?

The data that can be shared will be made available as Open access research data; this refers to the right to access and re-use digital research data under the terms and conditions set out in the Grant Agreement. Openly accessible research data can typically be accessed, mined, exploited, reproduced and disseminated free of charge for the user.

The Hydroptics project is pursuing the European Commission's vision that information already paid for by the public purse should not be paid for again each time it is accessed or used, and that it should benefit European companies and citizens to the full. This means making publicly-funded scientific information available online, at no extra cost, to European researchers, innovative industries and citizens, while ensuring long-term preservation.

The Data Management Plan (DMP) is not a fixed document, but evolves during the lifespan of the project. The following are basic issues that will be dealt with for the data that can be shared:

2. Types of data generated within the project

During the HYDROPTICS project it is not provisioned to collect any direct personal data. This facilitates the process of data collection and management.

The type of data that will be collected during the Hydroptics project have the following origins:

- **Administrative data**
 - Project management, project meeting notes, organizational documents, contracts, NDAs
- **Dissemination and communication data**
 - Brochures, workshops, posters, banners etc.
 - Reports on the business model and market study
 - Datasets, codes and other data related to a scientific publication
- **Simulations and modelling data**
 - Computational fluid dynamics of various refinery systems, simulations on the laser design, laser simulations, drawings of prototype designs, design of a beam-combiner, integrated laser and beam-combiner simulations, machine learning techniques implementation for data analysis, hyperspectral imaging data treatment.
- **Experimental data**
 - Data generated in the labs, experimental results, spectroscopic fingerprints, laser optical spectra, electrical RF spectra, laser optical beam data, prototype design/drawings.
- **Field tests data**

- Prototype outcomes, the operation results, future improvement directions

Therefore, there will be no personal data collected and/or utilized during the project. However, The Consortium has made clear that if a personal data were to be collected, a strict data protection policy should apply to these data. In particular, dedicated Consent Form has to be presented to the data owner and signed by the both parties.

2.1. Administrative data

This type of data includes all type of administrative documents, contracts and NDAs. In particular, all the project meeting notes, meeting agendas, decisions of the projects, mitigations (if any), risk identification and management, Consortium Agreement, NDAs with external bodies, any other type of legal docmunets.

This category of data will be saved and stored in the dedicated server for HYDROPTICS project, all the partners will have access to it.

2.2. Dissemination and communication data

This data is related to all the information related dissemination material including logo, brochures, posters, banners, events, workshops etc. All the significant results, achievements, news released in press. It also includes all the scientific publications that will be published. An important point is that the scientific publications will be free access in the webpage for the project since the published information has already been payed through an EU project. The data used in the scientific publications will also have free access. Videos of the project will be generated and saved in the project's webpage. In this category there will be also significant amount of data generated related to the buissness model development of the HYDROPTICS product. In particular, buissness canvas model, business plan.

This category of data will be saved and stored in the dedicated server for HYDROPTICS project, all the partners will have access to it.

2.3. Simulations and Modelling data

These type of data will be generated and filed throughout the entire lifetime of the project. In the first phase of the project, prior to the fabrication a significant amount of simulations are and will be performed for designing various sub-systems. In particular, a complete set of simulations will be made for the laser design to match the required wavelengths, as well as reducing the dispersion of the waveguides for optical frequency comb operation.

Computational Fluid Dynamics simulations will be performed of various systems of the oil refineries in order to generate a digital-twin of the real systems. This will help to understand all the processes in the refineries and try to design Hydroptics systems already optimized for the last.

Complete set of simulation will be performed for integrated beam-combiner in a silicon wafer, in order to understand and to achieve the best design for the last.

This category of data will be saved to the corresponding Partnes server. If the data are related to the scientific publication, it will be also stored in the shared platform.

2.4. Experimental data

During the first phase of the project different sub-systems of the final prototype will be developed and tested, hence there will be significant amount of data generated by fabricating sub-systems, characterizing, also when integrating the sub-systems towards having the prototype. In particular, complete set of characterization will be performed including LIV-curves, dispersion measurements, optical and RF spectra measurements, gain spectrum measurements and optical beam measurements. Dual comb integrated driving electronics tests will be performed towards developing a truly integrated electronics system. Characterization of various samples from the refinery fields will be performed in order to identify various molecules and their corresponding absorption spectra. Dual-

comb tests will be performed including optical and RF spectra measurements, time domain optical power measurements. Spectroscopy measurements with beam combined dual-DFB system will be performed with different samples. Tests on RF injection into QCL lasers will be performed aiming to understand and model an optimized design of the lasers for efficient RF injection. Complete set of beam-combined tests will be done for integrated beam-combiner in a silicon chip with integrated laser sources. This will include, soldering tests of the lasers, alignment tests with the last, mounting of lasers with pre-aligned wafers, beam combiner tests. UV-Vis-NIR Hyperspectral imaging experiments will be done in order to get a complete data set of the samples that are being measured, in order to facilitate the hard particle spectroscopy an ultrasound particle manipulations tests will be performed with using sound waves to concentrate the interested particles in a limited volume.

This category of data will be saved to the corresponding Partner's server. If the data are related to the scientific publication, it will be also stored in the shared platform.

2.5. Field tests data

During the final phase of the project there will be significant amount of field tests performed with the prototype, and related data generated from these field tests would be properly filed and saved. The data generated will be detected oil-in-water and various additives in water content, solid particles content in water. The format of the test field data will be finally designed and defined with the prototype, depending on exact targeted measurements that the prototype will execute.

This category of data will be saved to the corresponding Partner's server. If the data are related to the scientific publication, it will be also stored in the shared platform.

3. Data Formats

It is defined that all the reports (official and unofficial/internal), documents, internal meeting notes, minutes of meetings, meeting agendas, dissemination documents, press releases, dissemination materials, should have one of the following formats **.docx**, **.pdf**. All the graphical presentations of the experimental or other results, plots, figures should have one of the following formats **.jpg**, **.png**, **.pdf**, **.eps**. The video materials that will be created for the project will have one of the following formats **.avi**, **.mp4**, **.wmv**.

For the experimental data, simulation and design model data, each partner can generate and store the data in their preferred format, however, a version of the data for sharing will be generated and made available for the Consortium and/or EC in case of a request, with one of the following formats **.csv**, **.txt**, **.xlsx**, **.docx**, **.xml**, **.json**, **.pptx**.

4. Data Management plan

All the data will be saved in the hardware of the initial data generator Party (unless otherwise agreed), and it is the respective Partners responsibility to securely save the data, have a backup version, and be ready to provide it with the complete information to the other members of the Consortium or EU in case of a need. If the data has to be shared between all the Partners, that particular data will be uploaded to a dedicated OwnCloud data sharing platform and will be removed once there is no more need to have it shared. If a certain data has to be shared to the public, it will be uploaded to the project webpage, and will be accessible to everyone. More details on the data sharing can be found in the Chapter 5.

5. Data Sharing

5.1. Data sharing platform within the Consortium

OwnCloud is selected to be as a platform for data sharing between the Consortium. It provides a secure data sharing platform, with servers being securely guarded, moreover, the GDPR regulations are completely satisfied with this platform. A dedicated server and a service provider is purchased in order to securely govern, and provide a secure data sharing platform. The capacity of this particular sharing platform will be limited to 100 GB, hence this platform is not meant to save big data, this is more a platform for data sharing within the Consortium.

All the documentation related to the project should be shared in this platform, as well as some data if need be. In particular, all the deliverables, minutes of meetings, meeting agendas, internal reports, all the legal documents (contracts, CA, GA) should be shared in the platform. Each Partner will have maximum of 4 users to access, add and modify the shared platform. If need be, the capacity of the platform can be increased by purchasing more capacity.

5.1.1. OwnCloud

The sharing platform has the following address: <https://hydroptics.owncloud.de/>. The Main menu structure of the OwnCloud platform is generalized in the Fig. 1.

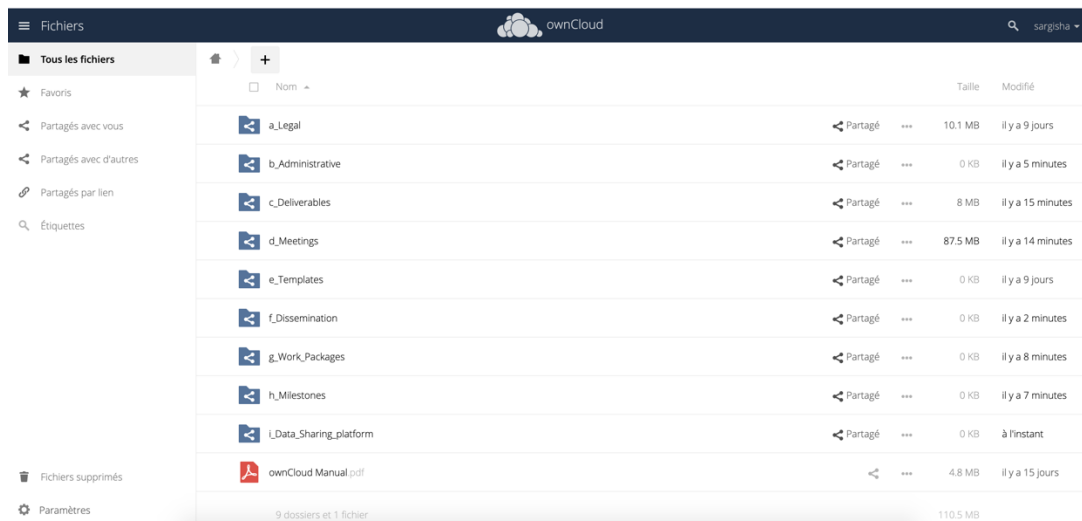


Fig. 1: Main menu of the OwnCloud data sharing platform.

Each Partner will have a possibility to connect 4 people to the platform, with maximum capacity of 5 Gb per Partner. Files will have different access/modify/delete/create policies for different Partners, depending on their relevance to the material. A detailed description of the main menu, and the policies can be found in Table 1.

Table 1: Detailed description of the file sharing platform, and the policies applied.

a_legal					
Contains all the legal files such as contracts, grant Agreement, consortium agreement, non-disclosure agreements, amendments, etc.					
Right to add	Coordinator				
Right to modify	Coordinator				
Right to delete	Coordinator				
b_Administrative					

Contains data related to general project management, timelines, list of executed tasks and tasks to be done.					
Right to add	Coordinator				
Right to modify	Coordinator				
Right to delete	Coordinator				
c_Deliverables					
Contains all the submitted deliverables with their respective files.					
Right to add	Coordinator	Respective Partners			
Right to modify	Coordinator	Respective Partners			
Right to delete	Coordinator				
d_Meetings					
Contains all the information for all the meetings of the project.					
Right to add	Coordinator	Respective Partners			
Right to modify	Coordinator				
Right to delete	Coordinator				
e_Templates					
Contains all the templates of the project, such as power point presentations, deliverables, meeting notes, minutes of meeting, meeting agenda etc.					
Right to add	Entire Consortium				
Right to modify	Coordinator				
Right to delete	Coordinator				
f_Dissemination					
Contains all the dissemination information. Brochures, Banners, Partners logo, webstie related data, scientific publications.					
Right to add	Entire Consortium				
Right to modify	Coordinator				
Right to delete	Coordinator				
g_Work_Packages					
Contains information on the work packages, status of the WP, timlines, achievements, executed tasks, upcoing tasks.					
Right to add	Coordinator	Respective Partners			
Right to modify	Coordinator	Respective Partners			
Right to delete	Coordinator	Respective Partners			
h_Milestones					
Contains information on the Milestones, status, timlines, achievements, executed tasks, upcoing tasks.					

Right to add	Coordinator	Respective Partners			
Right to modify	Coordinator	Respective Partners			
Right to delete	Coordinator	Respective Partners			
i_Data_Sharing_Platform					
This is a dedicated folder for sharing any information, data, files within the consortium members. Each Partners has a dedicated folder, where they will share information.					
Right to add	Coordinator	Respective Partners			
Right to modify	Coordinator	Respective Partners			
Right to delete	Coordinator	Respective Partners			

5.2. Data sharing platform outside the Consortium

All sharable data outside the Consortium will be published and hosted as per individual availability on the project’s public website i.e. www.hydroptics.eu. Partners generating the data are also encouraged to publish the sharable data on other online repositories.

The Hydroptics website has a friendly and easy to use navigation. It will be modified in due time to accommodate additional sections (pages) where the publishable data will be stored. The consortium will make sure that available data will be easily recoverable by any interested party.

The data will be made available on the website through adaptive webpages. The pages will cover the topics and project descriptive information to an appropriate level for each set of information or dataset.

The data will be formatted as per the description of each section, provided previously in this document, and will be presented for access along with the necessary links to download the appropriate software tools, if necessary.

The pages will be available to the public domain, enriched with the necessary metadata and will be open to web crawlers for search engine listing, so they will be available to the public through standard web searches.

All available datasets will be downloadable in their entirety.

Conclusions

The detailed data management plan has been presented with this document, specification of what data is expected to be collected within the projects lifetime, which data will be shared and which will stay confidential. A dedicated section is devoted on explaining how the data will be shared, a data sharing platform overview, also some details on the projects website sharing platform. It is important to mention that this is a living document and it might be modified throughout the project’s lifetime. All the Partners will be following the policies indicated in this document.