DELIVERABLE

D1.3 – Data management plan v2

Dissemination level	PU – Public
Type of Document	ODPR: Open research Data Pilot
Contractual date of delivery	30/11/2023
Deliverable Leader	ALPES
Status & version	V2.0
WP / Task responsible	WP1 / ALPES
Keywords:	Data, Data management plan

Deliverable Leader:	ALPES
Contributors:	Argyros Alexandros (ALPES), David Gachet (ALPES)
Reviewers:	David Gachet (ALPES)
Approved by:	David Gachet (ALPES)

Document History			
Version	**Date**	**Contributor(s)**	**Description**
V1.0	26/10/2023	ALPES	First complete version
V1.1	19/12/2023	ALPES	First complete version to be reviewed
V1.2	21/12/2023	ALPES	Final complete version
V2.0	21/12/2023	ALPES	Final version, after review by the Coordinator

Executive Summary

This document explains how the data was generated, used, saved, and shared between the Partners or published as 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 were generated, what formats were used to share the data between the Consortium members, which data can and/or was shared outside of the Consortium, and which data could not 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 has been a living document that has evolved during the lifespan of the project.

This DMP is committed to the principles of FAIR (Findable, Accessible, Interoperable, and Reusable) data to maximize the value of our research data, promoting transparency, collaboration, and innovation in line with the objectives of the Horizon 2020 framework.

Table of Contents

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 the partners' need to preserve the intellectual property rights and related benefits derived from project results and activities.

This document addresses the following questions:

- What types of data does 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 does the shared data have?
- How is data exploited and/or shared/made accessible for verification and re-use?

The data that can be shared is 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 has evolved during the lifespan of the project. The following are basic aspects that have been dealt with for said data to be shared:

2. Types of data generated within the project

The HYDROPTICS project did not collect any direct personal data. This facilitated the process of data collection and management.

The type of data that were 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 varies 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, hyperspectral images, prototype design/drawings.
- **Field tests data**

- Prototype outcomes, operational results, future improvement directions

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

2.1. Administrative data

This type of data includes all types 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 documents.

This category of data were saved and stored on the dedicated HYDROPTICS data sharing platform, that all the partners have access to.

Said data can be found on the ownCloud server of the project in various text forms, namely **.rtf, .docx, .pdf** and **.odt**

2.2. Dissemination and communication data

These data are related to all the information related to dissemination material including logo, brochures, posters, banners, events, workshops, and all significant results, achievements, or news released in press. They also include all the scientific publications that were or will be published. An important point is that scientific publications are open access, accessible on the project webpage , since the published information has been already payed through an EU project. The data used in the scientific publications are also accessible through open access. Videos that the project has generated were saved on the project webpage. In this category, there has been also significant amount of data generated related to the business model development of the HYDROPTICS product, in particular, business canvas models and business plans.

This category of data was saved and stored on the dedicated server ownCloud of the HYDROPTICS project, accessible by all partners. The file formats include **.jpeg, .png, .svg, .ods, .odt, .docx, .pdf, .xlsx** depending on the file usage.

2.3. Simulations and Modelling data

These type of data were generated and filed throughout the entire lifetime of the project. In the first phase of the project, prior to the fabrication of the platform, a significant amount of simulations was performed for designing various sub-systems. In particular, a complete set of simulations was 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 of various systems in the oil refineries were performed to generate digital-twins of the real systems. This helped understand all the processes in the refineries and to design Hydroptics systems optimized for them.

Complete sets of simulation were performed for integrated beam-combiner on a silicon wafer, in order to understand and to get the best design.

This category of data was saved on the corresponding Partner servers. If the data were relevant to scientific publications, they were also stored on the shared platform.

Data file formats include **.cfd4a, .cfd, .txt, .xlsx, .ods**

2.4. Experimental data

During the first phase of the project different sub-systems of the final prototype were developed and tested, hence there was significant amount of data generated during the fabrication, characterization and prototype integration of sub-systems. In particular, a complete set of characterization was performed including LIV-curves, dispersion measurements, optical and RF spectra measurements, gain spectrum measurements and optical beam measurements. Dual-comb integrated driving electronics tests were performed towards developing a truly integrated electronics system. Characterization of various samples from the refinery fields was performed in order to identify various molecules and their corresponding absorption spectra. Dual-comb tests were performed. They included optical and RF spectra measurements, time domain optical power measurements. Spectroscopy measurements with beam combined dual-DFB system were performed with different samples. Tests on RF injection into QCL lasers were performed aiming to understand and model an optimized design of the lasers for efficient RF injection. Complete set of beam-combined tests was performed for integrated beam-combiner in a silicon chip with integrated laser sources. This included soldering tests of the lasers, alignment tests with the laser, mounting of lasers with pre-aligned wafers, beam combiner tests. UV-Vis-NIR Hyperspectral imaging experiments were performed in order to get a complete data set of the samples that were measured. To facilitate hard particle spectroscopy, ultrasound particle manipulation tests were perfomed using sound waves to concentrate the interested particles in a limited volume.

This category of data was saved on the corresponding Partner servers. If the data were relevant to scientific publications, they were also stored on the shared platform.

Data file formats include **.cfd4a, .cfd, .txt, .xlsx, .ods, json, .xml, .pptx .tiff, .gds**

2.5. Field tests data

During the final phase of the project, a significant amount of field tests were performed with the prototypes, and data generated from these field tests were properly filed and saved. The generated data related to detected oil-in-water and solid particles content in water.

This category of data was saved on the corresponding Partner servers. If the data were relevant to scientific publications, they were also stored on the shared platform.

Data file formats include **.txt, .xlsx, .ods, json, .xml, .pptx, .tiff, .gds**

3. Data Formats

It was decided that all the reports (official and unofficial/internal), documents, internal meeting notes and minutes, meeting agendas, dissemination documents, press releases, or dissemination material, should be saved under one of the following formats **.docx, .pdf**. All the graphical presentations of results, plots or figures were saved in one of the following formats **.jpg, .png, .pdf, .eps**. The video material that was created during the project had one of the following formats **.avi, .mp4, .wmv**.

Each partner generated and stored experimental, simulation and design model data under its preferred format. However, a sharable version of data should be generated and made available to the Consortium and/or the EC in case of request, under one of the following formats **.csv, .txt .xlsx, .docx, .xml, .json, .pptx**.

4. Data Management plan

All the data were saved on the hardware of Partners generating the data (unless otherwise agreed). Each Partner is held responsible for its own data, for securely saving them, insuring they are backed up, and for sharing them, together with relevant information, with other members of the Consortium, as well as with the EU, if needed. If data were to be shared between Partners, these data were uploaded on the dedicated OwnCloud data sharing

platform and removed once there was no more need for sharing. If a certain dataset was shared with the public, it was uploaded on the project webpage, and kept being accessible to everyone. More details on data sharing policy in the following Chapter (5).

The list of dara generated during the project was consolidated and saved in a dedicated document that can be found on the Hydroptics' ownCloud data sharing platform. This document describes data, the way they were generated, as well as their format, size, storage location, ownership, and private/public status. If needed, this document will be updated after the end of the project.

5. Data Sharing

5.1. Data sharing platform within the Consortium

OwnCloud was selected as a platform for data sharing between Consortium partners. It provides a secure data sharing platform, with servers being securely guarded. Moreover, this platform complies with the GDPR regulations. A dedicated server and the services of a service provider have been purchased in order to securely govern, and provide a secure data sharing platform. The capacity of this particular sharing platform is limited to 100 GB, hence this platform is not meant to save big data. This is mainly a data sharing platform to be used by the Consortium members.

All the documentation related to the project has been shared on this platform, as well as some data necessary for the creation of other files. In particular, all the deliverables, minutes of meetings, meeting agendas, internal reports, all the legal documents (contracts, CA, GA) are shared on the platform. Each Partner has access by a maximum of 4 users, to add and modify the shared platform.

5.1.1. OwnCloud

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

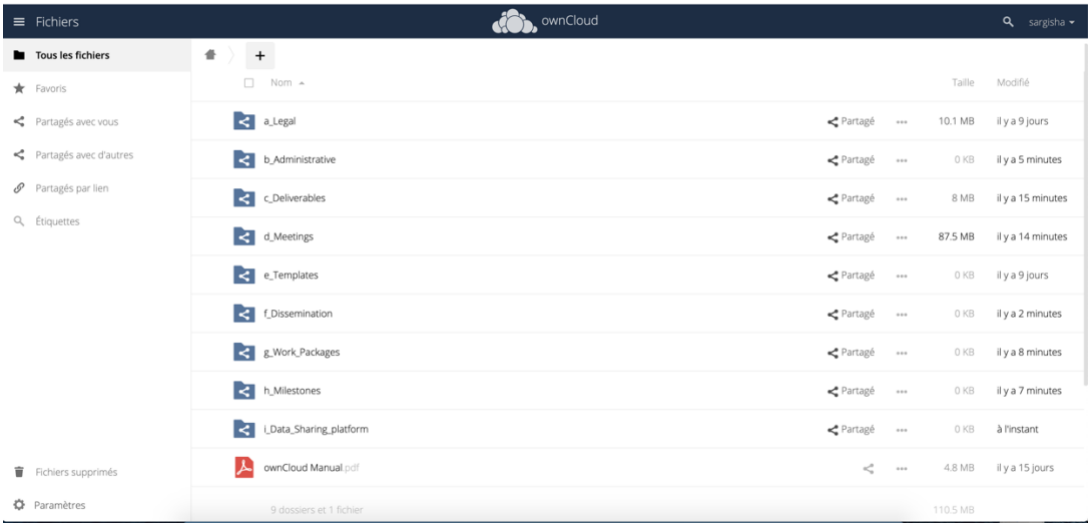

Fig. 1: Main menu of the OwnCloud data sharing platform.

Each Partner can connect up to four people to the platform, with maximum capacity of 5 Gb per Partner. Files 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 were published and hosted as per individual availability on the project's public website i.e., www.hydroptics.eu. Partners generating the data were also encouraged to publish the sharable data on other online repositories.

The Hydroptics website is user-friendly. It will be modified in due time to accommodate additional sections (pages) where the publishable data will be stored. The consortium made sure that available data can be easily recovered by any interested party.

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

The data were 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 made available to the public domain, enriched with the necessary metadata and open to web crawlers for search engine listing, so they are available to the public through standard web searches.

All available datasets will be downloaded in their entirety.

Early and open sharing of research: Early detection of opportunities was enabled through the continuous monitoring of dissemination activities automatically triggering openness assessment, through the philosophy of "As open as possible. As closed as necessary". The project served as an opportunity to gradually implement the practice of depositing pre-prints, posters, presentations in open repositories at their submission to journal/event committee when feasible. The practice of accompanying information and data in digital format resulted in speed sharing.

6. Ethical and Legal issues

6.1. Ethical Issues

The HYDROPTICS partners have complied with the ethical principles as set out in Article 34 of the Grant Agreement, which, among other, stated 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.

6.2. Confidentiality

All HYDROPTICS partners were required to keep any data, documents, or other material confidential during the implementation of the project and for additional period beyond the initial four years, in accordance with Article 36 of the Grant Agreement. Further detail on confidentiality can be found in Article 36 of the Grant Agreement. In addition, all partners complied with the General Data Protection Regulation 2016/679. Under the regulation, the data controllers and processors are fully accountable for the data processing operations. Any violation of the data subject rights may lead to sanctions as described in Chapter VIII, art.77-84.

7. Conclusions

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