Project Acronym:	Hydroptics
Grant Agreement number:	871529 (H2020-ICT-2019-2)
Project Full Title:	Photonics sensing platform for process optimisation in the oil industry

DELIVERABLE

D2.5 – Report on GDPR and legal aspects: Final

Dissemination level	PU – Public
Type of Document	Report
Contractual date of delivery	30/11/2023
Deliverable Leader	George Athanasiou, Sotirios Michagiannis, Vasiliki Andriopoulou (DBC)
Status & version	Final
WP / Task responsible	DBC
Keywords:	GDPR compliance, legal issues

Deliverable Leader:	DBC
Contributors:	George Athanasiou, Sotirios Michagiannis, Vasiliki Andriopoulou (DBC)
Reviewers:	David Gachet (ALPES)
Approved by:	David Gachet (ALPES)

Document History			
Version	**Date**	**Contributor(s)**	**Description**
V1.0	18/12/2023	DBC	Complete version
V1.1	22/12/2023	David Gachet (ALPES)	Minor corrections
V2.0	22/12/2023	David Gachet (ALPES)	Final version, approved for Submission 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.

1. Executive Summary

The document is the final version of the common legal framework for HYDROPTICS's consortium to take into consideration during research and realization phases of the project. Following the framework that was set with the first version of this deliverable on month 24, it provides additional guidance on compliance with the privacy and data protection legislation, especially in terms of data deletion and the way the data subjects could successfully exersice their privacy related rights. The deliverable also acts as a conclusion, presenting the measures that had been implemented throughout project's lifce cycle and ensuring data privacy and integrity.

2. Table of Contents

3. Data Summary

This section outlines the data/information handled by partners during the project (resulted from data mapping), describing the data that were collected, generated, re-used and/or processed in research activities. As indicated above, this is a first, high-level overview of such research data, as it is based on early-stage information collected from partners.

The main objective of HYDROPTICS project is to develop a set of integrated sensors, making use of advanced photonics subsystems, aimed at optimizing the processes of the oil industry. The device will be validated in real industrial settings, for oil extraction and oil refining processes. The HYDROPTICS platform will perform: 1) oil in water measurements, 2) corrosion inhibitor concentration measurements, 3) oil droplets and suspended solids in water measurements, 4) industrial process optimization, based on simulation of processes through digital twins, as well as data assimilation from the readings coming from the sensors.

Moreover, a key vision of HYDROPTICS is to elaborate how data provided by these advanced photonic sensors can be combined with readily available process data, and a digital twin of the process apparatus to gain in-depth process understanding. Digitalization of process data, data fusion, machine learning and artificial intelligence shall enable a new level of process optimization yielding high and constant product quality despite fluctuating process conditions.

Throughout the life cycle of the project, the processing of personal data was limited to specific categories of personal data and certain data subjects, nevertheless DBC, being in charge of the monitoring of the processing operations, proceeded to all the necessary actions in order to ensure data privacy and integrity in compliance with the provisions of the General Data Protection Regulation (GDPR) and the applicable national, data protection related, legislations.

The right to personal data protection can be affected by collecting and in general processing of personal data in HYDROPTICS project, actions which fall under the scope of GDPR. As a result, the HYDROPTICS's partners carefully and anticipatory defined the types of data, means and purposes of processing, the roles of every partner applied in the project., informed the data subjects, whenever applicable, regarding the processing operations, their purpose, the storage period of the personal data

After the assessment done in the second period of the project, the project collected, generated and/or used the broad categories of research data/information shown in Table 1 below:

Table 1: Categories of research data handled.

Data category	Data category content/examples	Expected data use
Contact, administrative, and other details of and data/information about partners, as well as relevant documents.	Data may include, among others, names; titles; organizations; email addresses; telephone numbers; membership, participation and contacts in committees, bodies and/or associations; affiliations; payment details. The data was/will be collected from partners during the project, as appropriate, and will encompass data of the partner itself and/or its representatives.	Data is collected and processed to allow communication and liaisons between partners in the course of the project and enable the organization and execution of project-related activities (e.g., but not limited to, organization of meetings, payments, other research activities, etc.).
Contact details of and data/information about external stakeholders.	This category may include, among others and as required in each case, names; email addresses; organizations and role in organizations; informed consent and opt-out forms; checkboxes about privacy and terms of service;	This data will be used to plan and carry out communication and dissemination activities (e.g., stakeholders' engagement, newsletter, registration in events), workshops, events and other project/research activities.

Data category	Data category content/examples	Expected data use
	signatures; data about participants in workshops. This category also encompasses data gathered from the website (e.g., related to usage statistics, registered users, etc.). The data will be collected, for example, by partners organizing and engaging in activities involving external stakeholders during the project.	
Agendas, presentations, minutes, notes, signature lists, recordings and other documents and data from meetings, workshops, and events.	This category includes data generated as part of research activities of the project and events organized in the context of the project.	Such data will be used to, among others, enable and/or document research activities of the project.
Information from interviews, surveys, questionnaires, structured feedback, records, recommendations from workshops or other activities, and other information collected or generated from partners and external stakeholders through contacts, as well as (internal) documents and reports.	Research data in this category comprise a variety of information and materials (to be) collected or generated during research activities. Information is/will be collected from (representatives of) partners and/or external stakeholders through, *inter alia*, questionnaires, interviews, case studies and other activities during the project, and will be generated by partners as part of their research activities.	Such data will be used to carry out tasks under the project and draft relevant deliverables.
Materials and data(sets) used/generated in the context of the design and/or development of technologies, components and services, the drawing of inferences, and/or the testing, evaluation and validation of models, techniques, technologies, components, solutions and services and the overall framework.	This category may include, for example, information and data from/related to measurements undertaken during project activities, as well as other metrics (such as data related to operational parameters of a mobile device, e.g., performance metrics); user/customer/citizen data/information; data about preferences, transactions, interactions and usage; survey data; design data; processed quality information and analysis results; other (electronic) documents (e.g., certifying and containing data or representing policies). This category encompasses both existing data (e.g., data from surveys held in the past, other historical data, public research data), as well as data that will be	Such data will be used, among others, for carrying out technical research activities of the project, and for testing, evaluating and validating results (e.g., through the pilot demonstrations to be executed).

Data category	Data category content/examples	Expected data use
	collected/generated during the project. Such new data collection/generation may take place, for example, and depending on the specific case, synthetically; through sensors deployed in the context of pilot demonstrations; internal communication buses; smartphones or computers via applications/software development kits ("SDKs") integrated with applications; surveys; forms filled by users; dedicated platforms (e.g., used for manually entering data or for uploading documents); application programming interfaces ("APIs"); tools/instruments/software used for measurements; experimentation while designing and prototyping. This category may also encompass technical documents, guidelines, plans and reports generated as part of such processes, as well as other documentation.	
Concepts, designs, and software.	This category encompasses, for example, concepts on software architecture and design (e.g., functions, components, data flow, etc.), software artefacts, source code (e.g., in textual and/or binary format).	Such data will be used for carrying out technical research activities of the project, including as part of the pilot demonstrations.
Bibliographical materials, literature, and similar publications.	Such materials may comprise, among others, reports, journal articles, books, websites, legislation, case law, existing technical data, etc. Publicly available materials will be collected as necessary and appropriate during the project through the respective source, e.g., (online) libraries, journals, websites, etc.	Such data will be obtained and used as background of and to support research activities undertaken as part of the project.

Following the presentation, in Table 1 above, of the broad categories of research data handled in HYDROPTICS, some further information is provided below regarding this data. As is the case with the identification of research data, the information provided below is preliminary and subject to change and/or further specification as the project progresses.

Types, formats, and size of data

The above categories of research data are expected to be in a variety of types, including (as a preliminary, non-exhaustive estimate, and subject to clarification, addition and change, as the project progresses) textual, numeric (e.g., int, float), alphanumeric, binary, boolean, image, audio/video, geolocation, datetime, and a variety of formats, such as docx, csv, json, pdf, xls, npy, png, xml. Given that the project is still at an early stage, it is not possible to estimate with certainty the total size of the data that will be generated. With regards to the expected size of datasets of individual partners, the preliminary estimates provided by partners vary, depending on the partner's role and activities in the project, while some partners have also indicated that at this stage the expected size of their datasets is unknown.

Data origin and potential modifications

The project will generally both use newly collected/generated data and will re-use existing data. The decision as to the data to be used will be taken by partners bearing in mind their tasks and activities under the project. Relevant considerations include, for instance, whether a particular project activity necessitates the collection of new data or can rely on existing data, as well as whether datasets meeting the requirements partners seek for the performance of their research activities are available. Some partners have indicated that they will not re-use existing data or that they do not at this stage expect that they will re-use existing data, while others have indicated that they are planning on doing so. Some partners are considering, but have not yet determined, whether they will use existing data. Collection of new data throughout the project will be done as necessary for each specific task. Existing data used by partners may comprise their own historical data or may be third-party data (e.g., reports, technical data, etc.), although, as of this stage of the project, partners have indicated that they are still considering the data they will be using. Existing data will potentially be re-used for research activities, data model and algorithmic development and training, validation of the performance of HYDROPTICS technologies, and as background for project work. Existing technology offerings are also expected to be used. The period covered by each dataset varies – with some expected to cover (part of) the duration of the project, while others extending in the past, either for a specified time period (e.g., data collected in MM.YY) or for an undefined period (e.g., when it comes to journal articles, websites, etc.).

Manipulations to some data(sets) are expected to take place, depending on the activities to be carried out under each task, as well as the evolution of project activities and technical solutions developed. For instance, newly collected data may be added to a dataset, obsolete data may be removed, or data may be updated. Existing data may be modified by partners before being (re-)used in the project, if required for the purposes of carrying out research activities. Datasets may be anonymized or pseudonymized, e.g., to be made available to other partners for the execution of their tasks under the project. Manipulations may also be executed to constitute data to be inserted into a given system processable in terms of quality and quantity.

Information regarding personal data

While some partners will not process personal data (or it is not yet clear whether the datasets they will use will include personal data),[1] it is envisaged that certain partners will collect and process personal data during project activities. Personal data that will likely be processed during the project may concern, for example, the following categories of stakeholders (the broader categories and the personal data processed to be confirmed at a later stage):

[1] Article 4(1) of the GDPR defines personal data as "any information relating to an identified or identifiable natural person ('data subject')". It also states that "an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person". Thus, only anonymized data (after the point of anonymization) or information that does not refer to natural persons falls outside the scope of the GDPR. Whether a person is identifiable can only be assessed on a case-by-case basis. When data is pseudonymized, meaning that personally identifiable information, e.g., the individual's name, is substituted with a unique identifier not connected to their real-world identity, the GDPR still applies.

- Partners (including names, organizations, email addresses, signatures, etc.);
- External stakeholders participating in workshops, interviews or other events and activities held by partners in the course of the project, or being informed about project activities and results (e.g., names, organizations, email addresses, signatures, etc.); and
- Customers/users/citizens involved in pilots (for example, but not limited to, name, gender, date of birth/age, nationality, preferences, biometric data, behavioural data such as data about movement, device usage/information and transactions).

While the datasets to be used by each partner for project activities are still to be defined, it is expected that some personal data will be pseudonymized or anonymized, while other will not be (e.g., because it would not be possible to carry out the activity in question with anonymized or pseudonymized data). Further information will become available as the project progresses and is expected to be addressed in relevant tasks/deliverables.

Certain partners are expected to only engage in "limited" personal data processing (e.g., contact details of other partners), while others may process personal data to a larger extent (e.g., as part of the piloting activities of the project). Some partners anticipate processing special categories of and other sensitive personal data (e.g., facial biometric data for identity verification, data required for immigration processes). Appropriate consideration and due respect shall be given to the applicable legal framework and the legal and ethical guidance provided by DBC in the research ethics and compliance protocol of the project. Given that guidance provided in the Protocol is expected to influence partners' data processing decisions and activities (e.g., but not limited to, with regards to the processing of certain personal data, such as personal data of vulnerable people or children), it is imperative that discussions take place among the respective partners before processing of personal data commences (e.g., as part of the pilot demonstrations). Based on the information currently available, personal data to be included in datasets handled by the partners will generally not be transferred to countries outside the EU/EEA,[2] although a partner has indicated that this may happen according to their organization's privacy policy and another partner has indicated that it is not yet known whether this will happen.

To the extent the processing of personal data is involved, it shall be ensured that it will take place in accordance with applicable EU, international and national law on data protection, notably the GDPR. [5][1] In particular, the data controller – that is, the natural or legal person who, "alone or jointly with others, determines the purposes and means of the processing of personal data" [5] – shall ensure respect for the seven fundamental principles relating to the processing of personal data set out in Article 5 of the GDPR, namely the principles of:

- Lawfulness, fairness and transparency;
- Purpose limitation;
- Data minimization;
- Accuracy;
- Storage limitation;
- Integrity and confidentiality; and
- Accountability.

In other words, it shall be ensured that personal data is:

- Processed lawfully, fairly and in a transparent manner;
- Collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes;
- Adequate, relevant and limited to what is necessary in relation to the purposes for which it is processed;

[2] As has been flagged to partners, a data transfer does not entail actually "sending" the data to a non-EU/EEA country. If a partner or service provider is located outside the EU/EEA and is able to access the personal data collected, this also constitutes a "data transfer".

- Accurate and, where necessary, kept up to date, with every reasonable step being taken to ensure that personal data that is inaccurate, having regard to the purposes for which it is processed, is erased or rectified without delay;
- Kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the data is processed; and
- Processed in a manner that ensures appropriate security of the data, including protection against unauthorized or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organizational measures. [5]

Finally, the data controller shall be responsible for and be able to demonstrate compliance with all the above.

Data utility

Certain research data collected, generated and/or processed during the project activities will not have, as such, a utility outside the project, being relevant and useful only for the execution of research activities during HYDROPTICS and for the preparation of the relevant deliverables. Other research data may have a utility outside the project, e.g., for researchers and academic community, for other EU-funded projects, for members of a particular industry/community, for third parties (private and public) pursuing solutions for similar projects as those tackled by the project's use cases, or to act as examples of the HYDROPTICS technology and document how it works.

4. Data storage

Following the data summary, this section focuses on data storage, access, and security.

While handling data, including storing and sharing data, the consortium considered and complied with requirements, obligations and standards set out in applicable legislation and guidelines, including – but not limited to – the GDPR. To the extent that the storage, sharing or other processing of personal data is involved, the fundamental data protection principles briefly outlined in the previous section should be respected. Further details about the legal and ethical standards and principles the partners complied with are provided in the following section of this document and will be further developed during the project.

4.1 Data storage, quality, and security

Data are stored by the partner(s) owning/providing each dataset. However, data storage was also provided by the technical partner hosting the production environment used for pilots (although they will not provide datasets for the project). Certain data and information were also stored on the project's common repository provided by the project coordinator. Generally (though non-exclusively), it is anticipated that the partner(s) owning the dataset will control access to this dataset, and will be in charge of collecting, storing and deleting the data. However, it may be the case, for example, that a partner that is not responsible for collecting the data is responsible for deleting it (e.g., in the context of pilot demonstrations, where entities participating in pilots may be in charge of collecting data, with a partner being responsible for deleting it). Furthermore, data subjects have control over their data, in line with applicable legal rules.

Various types of storage were used during the project, including local/on premise/on device storage, (proprietary) distributed storage (using cloud storage as a data backend), and cloud storage (which may either be arranged and used individually by a partner – e.g., Azure, AWS, Hetzner, OneDrive/Office 365, etc. – and/or be the project's common repository provided by the project coordinator). A combination of storage solutions was also used depending on the specific research activities undertaken and/or a partner's choice, a decision which may also be influenced, among others, by the sensitivity of the data in question. It is possible, for example, that while the full dataset is stored locally, some limited data may be stored on the cloud, or that certain data/documents are stored locally, while examples and demonstrators are uploaded on the cloud, or that non-sensitive project data is stored on the cloud for collaborative work, while other data is stored locally. Source code was stored on GitHub. Further decisions with regards to the storage are expected to be taken during the project. Partners also will keep back-ups

of research data, locally and/or on the cloud. Such back-ups were planned at appropriate intervals depending on the partner (ranging, e.g., from hourly to monthly).

Partners handling data adopted appropriate measures to ensure data integrity, quality and confidentiality. In addition, data security is imperative, and partners protected data and information they hold by adopting necessary security measures and mitigating any risks.[3] Overall, ensuring data confidentiality (i.e., that data is only available to those authorized and is protected against unwanted exposure and tampering), integrity (i.e., that data is protected and not altered to ensure that it is reliable, accurate and complete) and availability (that is, ensuring that authorized users can access the data in a timely and uninterrupted manner whenever necessary for carrying out activities under the project), in a balanced manner and in line with carrying out project activities is important. [6][7] Measures were, therefore, be put in place, as necessary and appropriate, by partners to make sure that data access is restricted only to the intended audience (e.g., by adopting procedures for identification, authentication and authorization), as well as to prevent intentional or accidental destruction or modification of data (e.g., by maintaining back-ups and carrying out system audits).

Data integrity, quality, confidentiality, and security measures that have been identified by the partners and which may be adopted, as appropriate in each case, include:[4]

- Encryption at rest and encryption in transit methods/protocols;
- Integrity file system checks;
- Access controls (for example, implementation of appropriate restrictions to data access; role-based access control; appropriate measures for authentication and authorization, e.g., implementation of single sign-on ("SSO") with multi-factor authentication ("MFA"), use of SSH keys for authentication, etc.);
- Data only being accessible from the organization's cooperate network through the organization's laptops;
- Use of virtual private network ("VPN") to access data located on the organization's servers;
- Access to personal computers with password;
- Definition of conditions and policies under which data, research infrastructure and related tools and applications can be used, instructing users to follow set guidelines on how to use the tools and services that handle data, and enforcement of security, trust management and acceptable use policies;
- Use of cryptographic means for data generation (e.g., p-ABC signatures);
- Storage and/or sharing of documents that do not contain personal or confidential data;
- Local storage in hard drives with no internet access;
- Use of encrypted computer disks;
- Storage of data on the partner's infrastructure only with appropriate access control and restrictions;
- Firewalls to ensure network security;
- Use of tools with appropriate security measures;
- Protection of internal services communication by virtual private cloud ("VPC");
- Validating input data to ensure accuracy and veracity of information regarding recorded values;
- Replicating databases in production;
- Regular backups (which will also ensure data recovery);
- Periodic recovery tests to ensure recoverability;
- Access to back-ups only by information technology ("IT") services;
- Implementation of pre-processing techniques during data collection and at period sequence;

[3] Security is a fundamental principle with regards to personal data, and Article 5(1)(f) of the GDPR requires that appropriate technical or organizational measures shall be taken to ensure that personal data are protected "against unauthorized or unlawful processing and against accidental loss, destruction or damage". Thus, when processing personal data, partners shall have appropriate security measures and controls in place.

[4] These are an overview of measures identified by individual partners in their individual responses, and each partner envisages to take some of these measures according to their specific circumstances and needs.

- Appropriate measures for physical security (e.g., physical access to premises for authorized persons only and through access control, guards in facilities);
- Regular security checks and audit controls;
- Opting for servers and services located in the EU to make sure that data handling is GDPR-compliant.

Furthermore, to ensure data integrity and quality, partners responsible for gathering information from other partners (e.g., use case providers) were in frequent communication. In addition, confidentiality rules binding partners as per the Grant Agreement ("GA") and the Consortium Agreement ("CA") are relevant when it comes to data confidentiality. Finally, it should be noted that internal security policies and procedures were put in place by certain partners also address such aspects, hence being relevant in the context of HYDROPTICS, and so are the security measures of the project's common repository provided by the project coordinator.

4.2 Data availability and sharing between partners

Certain datasets held by partners were not made available to other partners. Such datasets may be, e.g., those containing sensitive personal data or those that would not be useful to other partners for their activities under the project. In fact, it is also possible that certain datasets may only be available to part of the partner's team that will manage the data (which will be ensured through strong access control), and logging mechanisms will provide appropriate reporting as to who is accessing the data or any abnormal behavior. Although certain datasets may be inaccessible by partners other than their owner, results derived from their use may be made available to all partners, through the project's repository and/or through publications.

Nevertheless, data availability and sharing of some form between partners took place, to carry out research activities under the project. Two broad "types" of data sharing took place:

- Data sharing between use cases partners and (certain) technical partners; and
- Data sharing between partners carrying out similar or complementary activities under the project and/or being involved in the same tasks.

In addition, certain data/information were accessible by the consortium (e.g., when all partners shall contribute data/information) through the project's repository.

In the first case, use cases partners made (a curated version of) their proprietary datasets available to (certain) technical partners for the purpose of carrying out project activities. Technical partners assisted use cases partners, e.g., with the organization and execution of the pilots, as well as with dataset management, including, but not limited to, preparing specific datasets for use in the project, adding persistent identifiers, implementing standard metadata access protocols and, if necessary, transforming data for compliance with the FAIR Principles. Technical partners utilized use cases providers' data to, e.g., carry out analyses, generate source code and models, develop procedures, test techniques, etc., which may then be used in the context of the project and/or support project activities by other partners (e.g., pilot demonstrations). For datasets made available by use cases partners to technical partners, it is generally anticipated that the use case partner owning the dataset-controlled access to the dataset, including access mechanisms and processes implemented. Use cases partners controlled other aspects in relation to their dataset, though this is to be determined on a *per pilot* basis as the project progresses (e.g. while some partners may provide anonymized data to the project, others may require technical partners' involvement in this). Matters relating to access to/use of use cases partners' datasets by technical partners also depend on the technology used in a specific case and the specific activities concerned. It has also been suggested that the potential use of data spaces (e.g., data spaces based on International Data Space ("IDS") standards) in the project could also address some matters regarding access to/usage of datasets during the project, as data providers can have control over what they share and with whom.

With regards to data sharing, to the extent necessary and appropriate, between partners having similar or complementary tasks under the project, this took place to allow benefiting from synergies, to carry out technical or other work under the project and/or to utilize project resources more effectively (e.g., to avoid duplicating efforts of the consortium).

Broadly speaking, partners requiring access to another partner's data to carry out defined activities under the project were given access to that data. However, certain non-sensitive data may be available to all partners, e.g., through the project's common repository provided by the project coordinator, accessible by partners. Partners shall ensure that data sharing took place in accordance with legal requirements. Consequently, any necessary conditions for or restrictions to access to or use of the data (e.g., authorization) were set out, as appropriate, which is generally anticipated to be done by the partner owning the dataset. Information about access to data was also be provided to ensure that partners can undertake their research activities under the project. Furthermore, partners collaborated, as necessary and appropriate, to arrange aspects surrounding data access and sharing between them for the needs of the project (e.g., to come to an agreement with regards to data specifications, such as formats, restrictions, etc., to be used). In addition, partners adopted measures to ensure that data to be used by other partners during the project will be available to the latter when needed. Anonymization or pseudonymization of datasets performed, though, as of this stage in the project.

4.3 Archiving, preservation, and deletion of data

Applicable organizational, legal, and regulatory requirements, rules and obligations were taken into account in determining the approach to such matters.

Data was stored until it is clear that they will not be analysed again for project activities and/or until the project ends and the final review has been undertaken. Data was subsequently deleted and/or discarded from the storage that has been used during the project. Certain data and research results may, however, be kept (for a certain time) and/or archived after the project, while taking necessary and appropriate measures (e.g., blurring of personal data before archiving). Datasets published/shared with third parties (e.g., if open access is provided) and published results were kept after the project, and source code that has been generated during the project by certain partners may remain open-source on GitHub, while respecting the IPR strategy of the project. In addition, examples used to demonstrate technology, e.g., in the project's website or project deliverables, are expected to remain public.

5. Data principles in HYDROPTICS

The FAIR Guiding Principles are high-level principles developed by a range of stakeholders from academia, industry, funding agencies and scholarly publishers, with the purpose of creating guidance for researchers wishing to increase the findability and, ultimately, re-usability of their data (importantly, not only by individuals, but also by machines). Therefore, the FAIR Principles do not themselves constitute standards or specifications – instead, they are a guide to the FAIRness of data, helping researchers evaluate whether their choices are making their digital research assets FAIR. [8] The FAIR Guiding Principles comprise four elements, which are related, but independent and separable: findability, accessibility, interoperability, and re-usability. These principles are meant to be followed "in any combination and incrementally", considering the context and special circumstances of each case. These principles can be applied not only to data, but also to non-data assets. [8]

Responsible research data management in line with the FAIR Principles is a mandatory open science practice for H2020-funded projects, notably through the use of plans and open access to research data under the principle "as open as possible, as closed as necessary". [1] An important observation should be made at this stage: FAIR data does not equal open data (that is, data that is publicly available for everyone to access and re-use). Data can – and shall be – FAIR even if access is restricted. [1] Such restrictions may be due to necessary and legitimate reasons (e.g., protection of personal data, protection of IPRs, trade secrets, etc.).

The following sub-sections provide some information about individual partners' considerations with regards to each of the four elements of the FAIR principles.

5.1 Making data findable, including provisions for metadata

The first element of the FAIR Guiding Principles is "findability". The following steps lead to data being findable: (1) "(meta)data are assigned a globally unique and persistent identifier"; (2) "data are described with rich metadata"; (3) "metadata clearly and explicitly include the identifier of the data it describes"; and (4) "(meta)data are registered

or indexed in a searchable resource". [8] Assigning a globally unique and persistent identifier – that is, an identifier that cannot be re-used/assigned without referring to the specific data for which it was initially assigned, and that is not rendered invalid over time – is of utmost importance in achieving other aspects of FAIRness. [9] Having rich metadata, including descriptive information as to the context, quality, condition and/or characteristics of the data, allows researchers and, importantly, computers to find data based on the information provided by its metadata, thus facilitating re-use. [10] Finally, registering or indexing the (meta)data – e.g., in repositories or specialized engines – ensures that it can be discovered on the internet and, consequently, that it can be re-used, as others can be made aware of the data's existence. [11]

The partners considered measures related to the findability of their data, as appropriate and considering the specific circumstances of each case, although at this stage of the project the concrete measures to be taken towards this goal cannot be identified with certainty. Measures considered, and potentially taken, as appropriate, include assigning a persistent identifier to data and/or metadata, providing rich metadata (e.g., in terms of type of data, timestamps, etc.) and search keywords in the metadata to optimize the possibility for discovery and potential re-use. In addition, specified naming conventions were followed, version numbers provided, and metadata will be offered in a way that can be harvested and indexed. The use of trusted open access data repositories (e.g., Zenodo) were considered.

5.2 Making data accessible

Once data has been found, the next step towards potential data re-use is to know how such data can be accessed. The FAIR Guiding Principles indicate that data will be "accessible" when, (1) "(meta)data are retrievable by their identifier using a standardized communications protocol"; which (1.1) "is open, free, and universally implementable", and (1.2) "allows for an authentication and authorization procedure, where necessary"; and (2) "metadata are accessible, even when the data are no longer available". [8] In other words, FAIR data access entails that access is mediated without specialized or proprietary tools or communication methods (e.g., protocols that have limited implementations or poor documentation) – notwithstanding the possibility that the access protocol is not fully mechanized, in case of, e.g., highly sensitive data. [12] Instead, the protocol used should be free and allow anyone with a computer and an internet connection to potentially access at least the metadata. [13] As explained above, however, FAIR data are not open data: hence, what is envisaged under the "accessibility" principle is that the exact conditions under which the data is accessible should be provided – in other words, ideally, it should be possible for a machine to automatically understand the requirements for access and then either automatically execute the requirements or alert the user about the requirements (e.g., the need to create a user account to authenticate). [14]

Taking the above considerations into account, the partners made an assessment, as to the data to be made accessible to third parties, including the data to be made openly available. Based on that assessment (and in line with any applicable rules), partners also considered the conditions of access to such data and how to appropriately specify such conditions.

Certain data included as examples, e.g., in technical documentations of solutions developed through HYDROPTICS. It is anticipated that this data will be made accessible accompanying such documentation (e.g., in a code repository, the project's deliverables, etc.). It is also anticipated that some research data will not be accessible to third parties as such. However, processed information, resulting activities, analyses, models, or other results are expected to be made available in, e.g., deliverables disseminated to the public, reports, journal articles, conference proceedings and other technical documentation of the project solutions. Certain results may also be made available through open pre-print repositories (such as, hal and eprint.iacr). Code may be hosted on GitHub in a repository that supports read access by any user, certain technologies developed are expected to be available as open source on GitHub and certain data may be made accessible through an Open API exposure. Research data underpinning scientific publications may be made available to third parties – under the premise "as open as possible, as closed as necessary" – at the time of the publication, in which case it is expected that it will be deposited in public repositories (e.g., gitlab or GitHub) or in the repository associated with the publication. To the extent shared, such data will generally be in a widely used format. In some cases, certain technologies, methods, or software tools may

be needed to access the data. If available, it is considered that standardized protocols will be followed. Finally, metadata may be described in relevant publications that present project results and, to the extent that the data is published alongside the publication, metadata may also contain information to enable the user to access the data. Metadata reported in publications is expected to remain accessible even if published research data is no longer available.

5.3 Making data interoperable

For data to be re-used, it usually needs to be integrated with other data, and to interoperate with applications or workflows for analysis storage and processing. [15] The "interoperability" principle entails that: [8] (1) "(meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation" – which would allow humans to exchange and interpret each other's data, and machines to read the data without the need for specialized or ad hoc algorithms, translators or mappings (which means that they should know the other system's data exchange formats); [16] (2) "(meta)data use vocabularies that follow FAIR principles" – meaning that the vocabulary used in describing the dataset shall be documented and resolvable using unique and persistent identifiers, with the information being easy to find and accessible by those who use the dataset; [17] and (3) "(meta)data include qualified references (i.e., cross-references explaining their intent) to other (meta)data" – in other words, it makes and sufficiently describes meaningful scientific links between (meta)data resources, enriching the contextual knowledge about the data (e.g., specifying that one dataset builds on another, stating that complementary information is found in another dataset). [18] Thus, this principle refers to technical interoperability.

Broadly speaking, partners made data with a utility outside the project interoperable. Data held by the partners is expected to generally use widely known and community-endorsed vocabularies, standards, formats or methodologies (e.g., json, xml, ngis-ld,), and necessary information to process the data and understand what it is by associating a publication with its metadata may be provided. Certain partners store data in an open data format, which can be opened with any programming language, and while making available the format specifications. In addition, specific ontology mappings are expected to be provided, as necessary. To the extent relevant, qualified references to other data may be included.

5.4 Increasing data re-use

Increasing and optimizing data re-use is the ultimate goal of the FAIR principles. [15] To ensure that data is re-usable, "(meta)data [shall be] richly described with a plurality of accurate and relevant attributes. [8] This means that adequate and appropriate labels attached to the data should allow the (re)-user (human or machine) to decide if the data will indeed be useful for them in the particular context (by considering, for example, the context under which the data was originally generated). For this to happen, aspects such as the purpose for which the data was generated/collected, the date of generation/collection, who prepared the data, the software used, etc. should be described. [19] Moreover, (meta)data need to be "released with a clear and accessible data usage license", indicating clearly the conditions under which the data can be re-used; and it shall be "associated with detailed provenance" – i.e., it should be clear where the data comes from, who generated or collected it, how it has been processed, whether it has been published before and whether it contains data from other sources. Finally, it shall "meet domain-relevant community standards", as having data that is of the same type, that is organized in a standardized manner, that uses well-established and sustainable file formats, that is accompanied by documentation that follows a common template and that uses a common vocabulary makes it easier to re-use datasets. [8]

The partners need to take a number of steps to increase the re-usability of their research data that may have a utility outside the project:

- Thoroughly documenting the provenance of data using appropriate standards;
- Documenting software and providing Git instructions;
- Including descriptions of the data in shared documents and describing formats, designs, and applications in technical documentations (e.g., methodology, fields used, purpose, etc.);

- Making available readme-files alongside the data, as well as notebooks to illustrate the use of data, instructions and guidelines;
- Making available through publications results of analysis carried out on the data, which will allow to validate further data analysis by comparing outcomes;
- Publishing examples alongside technical descriptions of the solution developed through public deliverables, repositories and/or research publications.

In addition, the use of appropriate common types of data sharing/re-use licenses were considered (e.g., Creative Commons). Quality assurance procedures were defined during the project, and, among others, may include, e.g., downloading published data to check that it corresponds to data analyzed to generate the results (verification) or performing risk assessment for tools developed following applicable frameworks.

6. HYDROPTICS compliance

In this section, the main concepts, and the most important Articles of the GDPR are thoroughly analyzed, while the basic principles that rule the processing of personal data as well as the main rights that the GDPR provides to the data subjects are also being mentioned. Furthermore, in this section the legal bases under which data is processed, the privacy policy which may be implemented by data controllers. Special reference has been made to the manner in which the security of processing is ensured in the project, but also to the manner in which we deal with a personal data breach situation. Moreover, the definition of the DPIA has been analysed and the cases in which it might be performed in the project have been mentioned.

6.1 The purpose of the GDPR and its core concepts

The purpose of the GDPR is to protect fundamental rights and freedoms of natural persons and in particular their right to the protection of personal data, while also ensuring the “free movement of personal data”. The GDPR sets out the EU regulatory framework for the processing of personal data. According to Article 3, the Regulation applies to the processing of personal data in the context of activities of an establishment of a controller or processor in the EU; to the processing of personal data of data subjects who are in the EU by a controller or processor not established in the EU, where the processing activities are related to offering goods or services to such data subjects in the EU or monitoring their behaviour as far as it takes place within the EU; as well as to personal data processing by a controller not established in the EU but in a place where EU Member State law applies on the basis of public international law.

In accordance with Article 1 of the GDPR, the Regulation lays down rules relating to the protection of natural persons with regard to the processing of personal data and rules relating to the free movement of personal data. The Regulation protects fundamental rights and freedoms of natural persons and in particular their right to the protection of personal data. The free movement of personal data within the Union shall be neither restricted nor prohibited for reasons connected with the protection of natural persons with regard to the processing of personal data.

For a more effective understanding of the Regulation, the following definitions are considered indispensable, provided in Article 4 of the GDPR:

“Personal data” means any information relating to an identified or identifiable natural person (“data subject”). An identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural, or social identity of that natural person.

“Processing” means any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure, or destruction.

The processing may involve special categories of personal data. Processing of personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation shall be prohibited. All this data is sensitive and belongs to special category of personal data. One of the cases when processing of sensitive data is allowed is when the data subject has given explicit consent to the processing of this specific personal data for one or more specified purposes.

"Data processor" is defined in the GDPR as the natural or legal person, public authority, agency, or other body which processes personal data on behalf of the controller.

"Data controller" is defined in the GDPR [5] as the natural or legal person that, alone or jointly with others, determines the purposes and means of the processing of personal data. It is possible that a controller may involve another actor, known as a data processor, to process personal data on behalf of the controller and in accordance with the ways and purposes specified by the latter. This difference can be difficult to translate into the complexity of modern relationships but what is decisive is the scope of decision-taking power – i.e., who decides the "why" and "how" personal data shall be processed. Data controllers have direct obligations by handling data of data subjects, but data processors will be under direct obligations as well, such as the obligation to maintain a written record of processing activities carried out or a duty to notify the controller on becoming aware of personal data breach without undue delay. Data controllers must provide transparent information to data subjects at the time personal data is obtained, in a clear, comprehensive, and easily accessible way. The data subject must be informed about his/her rights, the way data is going to be processed, for which reason, but also, for example, about the period for which data is going to be stored.

Although the status of each partner (data controller or data processor) is not clear at the moment since we are still in the beginning of the project and many partners still do not have a clear knowledge regarding their duties, it is expected, during the progress of the project, that they will process personal data according to GDPR, such as ID and contact information, information regarding their work or marital status or any other data that could help identify a person, directly or indirectly. Moreover, some partners are also expected to process special categories of personal data (according to Article 9 of GDPR), such as biometric data for example.

As a result, the partners of the project should be in a position to recognize whether they are expected to process personal data and thus, need to respond properly to all the preliminary actions being taken to assure data safety by the authorized partners. Also, those who are due to engage in data processing activities, should respect the GDPR provisions, having into consideration the basic principles of the Regulation and the rights and freedoms of the data subjects.

6.2 General principles of data protection and rights of the data subjects under the GDPR

The basic principles that are required to be observed when processing personal data by data controllers and processors are set out below. The legal bases for the processing of personal data and the main rights of data subjects when processing their data are also extensively described. Those are considered to be the cornerstone of data safety and the partners of the project are obligated to respect them throughout their data processing activities.

The basic principles that are required to be observed during the processing of personal data from data controllers and processors are, according to Article 5 of the Regulation, the following:

- **The principle of lawfulness, transparency, and fairness of the processing of personal data.** Personal data must be processed in a lawful, transparent, and fair manner. The lawfulness of processing is ensured, in accordance with Article 6 of the Regulation, in cases where the prior consent of the data subject to the processing of his/her data for one or more specified purposes has been obtained, the processing is necessary for the performance of a contract or for compliance with a legal obligation of the controller arising from another rule of law, the

processing is necessary to safeguard a vital interest, and finally processing is necessary for the purposes of the legitimate interests pursued by the data controller, including but not limited.
Transparency is ensured by providing to the data subject all information on the processing in a concise, transparent, and comprehensible manner. Fairness means that processing must be done in ways that data subjects would reasonably expect and not in ways that have unjustified adverse effects on them.

- **The principle of purpose limitation.** Personal data shall be collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes.
- **The principle of data minimization.** Personal data shall be adequate, relevant, and limited to what is necessary in relation to the purposes for which it is processed.
- **The principle of accuracy of personal data.** Personal data shall be accurate and, where necessary, kept up to date. Every reasonable step must be taken to ensure that personal data that is inaccurate, having regard to the purposes for which it is processed, is erased, or rectified without delay.
- **The principle of storage limitation.** Personal data shall be kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data is processed.
- **The principle of integrity and confidentiality.** Personal data must be processed in a way that guarantees appropriate security of personal data, including its protection from non-authorized or unlawful processing and accidental loss, destruction, or damage by using appropriate technical or organizational measures.
- **The principle of accountability,** under which the controller and processors are responsible for, and must be able to demonstrate compliance with, the principles relating to the processing of personal data.
- **The principle of proportionality**, which requires that there must be a connection between data kept and the purpose for which it is collected.

According to Article 6 of the Regulation, in order for the processing of personal data to be lawful, it must be based on a legal basis. The GDPR requires any organization processing personal data to have a valid legal basis for that processing activity. The GDPR provides six legal bases for processing personal data: **consent of data subject**, **performance of a contract** (the data processing should be lawful where it is necessary in the context of a contract or the intention to award a contract), **a legitimate interest pursued by the data controller or by a third party** (the legitimate interest assessment is carried out in three stages. First, the actual legitimate interest of the controller is assessed. Secondly, it is assessed whether the envisaged processing is strictly necessary for its fulfilment and thirdly, whether the processing overrides the interest of the data subject not to have his/her data processed for reasons of privacy), **a vital interest of the data subject or of another natural person** (processing is necessary in order to protect the vital interests of the data subject or of another natural person), **a legal requirement** (the processing is necessary for compliance with a legal obligation of the data controller), and **a public interest** (processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested on the data controller).

Article 7 of the Regulation prescribes conditions for demonstration of consent. “Consent” of the data subject means any freely given, specific, informed, and unambiguous indication of the data subject’s wishes by which he/she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her. When consent is given in a written form, it should be clearly distinguishable from the rest of the information provided to the data subject. The informed consent shall leave no possible space for ambiguity or confusion for the data subject, and it should make clear that the aim is to collect and process their personal data. The consent should use clear and plain language, which means that its content shall not be hidden behind complex legal formulas. It is necessary to inform the data subject of the possibility to withdraw their consent at all times and without need for any uneasy procedures. Consent to process sensitive data will have to be explicitly given by a data subject and they must be given the capability to withdraw consent effortlessly. This is confirmed by Article 7(3) of the Regulation, which states that the data subject shall have the right to withdraw his or her consent at any time and it shall be as easy to withdraw as to give consent. When the data subject has given explicit consent to the processing of sensitive personal data for one or more specific purposes, then the processing of such personal data is permitted. Existing consent given under previous rules may still be valid, but only given that they meet the new stricter requirements.

The withdrawal of consent shall however not affect the lawfulness of processing of data based on consent given before its withdrawal and the data subject shall be informed thereof.

Before personal data is collected from the data subject, the data controller shall provide him/her with a variety of information prescribed in Articles 13 and 14 of the GDPR, such as the identity of the controller, contact details, the purpose of the processing of data (lawful processing), the period for which data is going to be stored, the existence of the right to request from the controller access to and rectification or erasure of personal data, the right to lodge a complaint with a supervisory authority, etc.

The GDPR further acknowledges a range of rights for data subjects whilst the existence of these rights should be brought to the attention of the data subject in the explicit and clear manner in the informed consent. Our partners should make sure that the data subjects will be able to exercise their rights properly and respond to any relevant requests of the data subjects. These rights include, among others, the following (Articles 15-22 GDPR):

- **Right to information (requirement of transparency).** In accordance with Articles 12, 13 and 14 of the Regulation, the data subject has the right to know the full identity and contact details of those who collect data, either directly from him or indirectly (e.g., through cookies of websites visited). The data subject must also know exactly what data is collected by the data controller, for what purpose, for how long it is kept and to which recipients it is transmitted, if any. The information to the user must be provided in plain language, without any ambiguities, terms, and conditions and without any unclear legal or technical terms.
- **Right to access, erasure, rectify or restrict data.** The data subject has the right (Articles 12 and 15 of the Regulation) to receive full and thorough information as well as oral or written confirmation of the processing, description of the purposes, categories of data and recipients. In addition, the possibility and procedure for submitting a request for rectification, erasure, restriction of processing and for lodging a complaint with the competent authority should be made known to the data subject. The data subject should also be provided with a copy of his/her data. Moreover, the data subject has the right (Articles 5(1)(d), 16 and 17 of the Regulation) to request, within a reasonable period of time, the correction of inaccurate personal data and the completion of incomplete data. The data subject shall bear the burden of proving his or her true identity. The data subject has the right (Articles 5(1)(d), 17 and 19 of the Regulation) to request the erasure of his/her personal data without having to invoke any prejudice to him/her. This is a relative and not an absolute right which might be infringed when there is an overriding legal ground for the retention of personal data. Ultimately, the data subject has the right (Articles 4(3), 12, 18 and 19 of the Regulation) to request the restriction of the processing of his/her data when he/she contests its accuracy, when the processing is unlawful or when the data is no longer needed by the controller.
- **Right to data portability to other data controller** – this is a right to receive the personal data concerning the data subject, which he/she has provided to a controller, in a structured, commonly used, and machine-readable format and have the right to transmit this data to another controller.
- **Right to object against further processing.** In principle, Article 21 GDPR prescribes that the data subject has the right to object the further processing of their data at any time, unless the controller has a justified aim to continue the processing, which overrides the rights and interests of the data subject. It should be up to the data controller to prove that its compelling legitimate interests possibly override the interests or fundamental rights and freedoms of the data subject. However, it is very difficult to prove that processing of data overrides the rights and interests of the data subject. The situations in which such overriding can occur is when there is an epidemic, threat to public health, threat to national security, etc.

4.3 Data protection policy

Under Article 24 of the GDPR, the data controller shall implement data protection policies as part of appropriate technical and organizational measures to demonstrate compliance where proportionate in relation to processing activities. In this sub-section, some proposed measures will be presented, in order for the compliance of the project with the data protection related provisions to be assured.

6.4 Data mapping

Data mapping is crucial to the success of many data processes. In the scope of data protection regulations, the purpose of data mapping is firstly to consider whether personal data, as defined in the GDPR, is being processed and if so, to further examine other factors related to data protection principles such as for example the lawfulness of the processing, the respect of the subjects' rights and possible threats to the subject's rights and freedoms related to any act of processing.

In order to achieve this, the a questionnaire was sent to the partners, asking them to provide information about the data expected to be handled during the project, how such data would be handled, and how the FAIR principles would be taken into consideration. In the questionnaire there are also questions related to the handling of research outputs other than data, as well as allocation of responses and ethics matters.

After the evaluation of the answers received, we were able to witness whether any personal data is being/will be processed and if this is the case, whether further measures should be taken regarding this. The answers are also an indicator of the knowledge that partners have related to data protection principles, the level of their compliance in relation to their project activities and the measures they have implemented (or plan to implement) in order to achieve data security and integrity. Thus, it is important that the partners answer such questions while having the necessary knowledge regarding basic GDPR provisions, as provided in this section, in order to ensure that their answers are informed.

6.5 Data protection policy

As any type of technical and organizational measures to demonstrate compliance, the data protection policy was implemented considering the nature, scope, context, and purposes of processing, as well as the risks of varying likelihood and severity for the rights and freedoms of natural persons. In other words, the data protection policy is a set of principles, rules, and guidelines that informs how the organization will ensure ongoing compliance with data protection laws. The policies should recognize the data protection principles and the rights of individuals set out by the GDPR and explain how they are put into practice in relation to the processing carried out by the organization.

The template provided in **Annex 1** is to be implemented and detailed by project partners processing personal data with regards to their activities in HYDROPTICS. The partners should make sure that they implement those example provisions and guidelines throughout any act of processing of personal data during the project. They should be properly informed regarding basic GDPR rules and principles, have the necessary knowledge to deal with any data protection related issues they may face and most of all, be willing to seek for guidance or help in any data protection related case.

6.6 Data protection officers

Articles 37-39 of the GDPR refer to the appointment of a data protection officer and describe in which circumstances the appointment of such an officer is recommended and necessary. Appointment of a DPO is necessary if data is processed by a public authority, or if the core activities consist of processing operations which require regular and systematic monitoring of data subjects on a large scale, or if the core activities consist of processing on a large-scale special category of data.

The partners of the HYDROPTICS project are obligated to record an internal analysis to determine whether a DPO should be appointed, so that they are able to prove that consideration was given to the nature, scope, context, and purposes of the processing, as well as the risks of different probability of occurrence and severity to the rights and the freedoms of natural persons. This analysis is part of the documentation required under the principle of accountability, could be requested by the supervisory authority, and should be updated when deemed necessary.

Moreover, the term of "regular and systematic monitoring of data subjects" provided in the DPO provisions may also be a factor on this project. According to the Article 29 Working Party ("WP29") "Guidelines on Data Protection Officers ('DPOs')", regular monitoring is the process which is ongoing or occurring at particular intervals for a

particular period, recurring or repeated at fixed times or constantly or periodically taking place. Systematic monitoring of data subjects could be the process occurring according to a system, pre-arranged, organised, or methodical, taking place as part of a general plan for data collection or carried out as part of a strategy. [22]

To conclude, we suggest the partners to carefully examine whether the appointment of a DPO is required to assure that the requirements of Articles 37-39 of the GDPR are being met.

6.7 Data management and measures

With the current sub-section, we will refer to the way personal data could be collected during the project, how this will be managed, described, analyzed, and stored and what mechanisms will be used to share and preserve data.

Data processing principles

Data collection, management and in general processing throughout the duration of the project, were guided through the following principles/measures:

- Data could be collected/processed mostly in an **anonymized form**. In this case, the questionnaires, interview guidelines and other used instruments, where possible, must not contain questions, whose answers could lead to the participant's identity – alone or in combination with other answers.
- For the purposes of individual tasks, data could be pseudonymized instead. In this case, the partners should justify their action and data subjects should be informed and provide the partners with the necessary consent, according to Article 6 GDPR, to the extent that this is the legal basis for processing. Data should be anonymized by the time data processing is finished.
- The legal basis of processing of project partners' personal data (names, communication information, etc.) for the purposes of the project (for example to communicate and cooperate with other partners of the project) could be performance of the contract, or another legal basis as appropriate.
- The legal basis of processing of personal data of any of the participants to the project (stakeholders, citizens participating to pilot activities, etc.) could be consent according to Article 6 GDPR.
- Partners are strongly advised to refrain from data processing of special categories of personal data of Article 9 or 10 GDPR. In case the processing of such personal data is necessary for the purposes of the project, the partners should be provided with the necessary consent, in which the data subjects should be also informed regarding the necessity of such acts of processing.
- The **anonymity**, where possible, and the **privacy of participants** (stakeholders, citizens participating to pilot activities, etc.) must be respected. Personal information must be kept confidential and transferred to other project partners only if this necessary for a specific task related to the project. In this case, data subjects should be informed regarding the possibility of such acts of processing while providing the partners with their consent. Guarantees of confidentiality and anonymity given to the participants must be honoured, unless there are clear and overriding reasons to do otherwise.
- In case that the participants must be registered at the HYDROPTICS platform, they must not be registered with their name, if possible. E.g. an ID-code could be applicable instead of it. That **takes into account privacy considerations** and further, the ID-code helps to match answers of questionnaires and the data collected at the platform by the user.
- The data subjects themselves have their **data sovereignty**. Although the data has been provided by them and is being processed for the purposes of the project, the data subjects retain control over their data. So, for example, in case the data subjects request the deletion of their data, this has to be done without any undue delay.
- The participant is allowed to change/limit the access authorization of their data collected at the HYDROPTICS platform.
- All researchers have the duty of maintaining confidentiality of the collected data.
- The integrity of stored, processed and published data must be ensured by the researchers and the project consortium.
- Data that is collected by the participant at the HYDROPTICS platform must be treated with care:
  - Participants must be informed that the data could be used for the project.

- Participants must be informed in which way the data could be used.
- The participants must be informed when the collected data will be deleted.
- Appropriate measures, namely cryptography and physical security measures, must be taken by the partners to process data in secure manner.
- In case the participant withdraws from the pilot, the collected data at the platform must be deleted or the access to them must be impossible for others, without any undue delay.

Security of processing

The GDPR does not stipulate exact ways of data security, rather it gives minimum recommendations and urges the data controllers to make decisions on which technical and organizational measures to take depending on the state of art, the cost to implement, the varying likelihood of risk, but also the fundamental rights and freedoms at stake. Technical and organizational measures are for instance (Article 32 GDPR):

- The pseudonymization and encryption of personal data. Pseudonymization of data means replacing any information which could be used to identify an individual with a pseudonym, or, in other words, a value which does not allow the individual to be directly identified. It is a technical and organizational measure to ensure non-attribution to an identified or identifiable person and, according to the GDPR, **pseudonymized data are "personal data" since they can indirectly help identify the data subject with the use of additional information.** On the contrary, anonymized data (anonymization is the process of removing personal identifiers, both direct and indirect, that may lead to an individual being identified) cannot help identifying the data subject since all the personal identifiers have been removed and as such, is not considered to be "personal data" according to the Regulation.
- The ability to restore the availability and access to personal data in a timely manner in the event of a physical or technical incident.
- A process for regularly testing, assessing, and evaluating the effectiveness of technical and organizational measures for ensuring the security of the processing.

The partners, to be able to demonstrate compliance with this Regulation, should establish internal policies and implement measures, which specifically respond to the principles of data protection by design and by default. Data protection by design means that implementation of appropriate technical and organizational measures, such as pseudonymization, which are designed to implement data protection principles, such as data minimization, shall be done in an effective manner. The necessary safeguards into the processing shall be integrated from the very outset to meet the requirements and protect the rights of data subjects (Article 25 GDPR). Data privacy by default means that only personal data which are necessary for each specific purpose of the processing are processed. That obligation applies to the amount of personal data collected, the extent of their processing, the period of its storage and its accessibility.

HYDROPTICS considered the implementation of the following techniques [23]:

- **Directory replacement**: A directory replacement method involves modifying the name of individuals integrated within the data, while maintaining consistency between values, such as "postcode + city".
- **Scrambling**: Scrambling techniques involve a mixing or obfuscation of letters. The process can sometimes be reversible. For example: "Annecy" could become Yneanc.
- **Masking**: A masking technique allows a part of the data to be hidden with random characters or other data. For example, pseudonymization with masking of identities or important identifiers. The advantage of masking is the ability to identify data without manipulating actual identities.
- **Personalized anonymization**: This method allows the user to utilize their own anonymization technique. Custom anonymization can be carried out using scripts or an application.
- **Blurring**: Data blurring uses an approximation of data values to render their meaning obsolete and/or render the identification of individuals impossible.

While the project evolves, the consortium will compare the proposed techniques and decide to implement one or more techniques according to the needs of the project.

Data minimization

The data minimization principle comprises that data has to be adequate, relevant and limited to what is necessary for the purposes for which it is processed. This implies that:

- Data collected, processed, analyzed, and archived should not be held or further processed, unless this is essential for reasons that were stated in advance.
- Data collection and processing should only include as much data as is required to successfully answer the research question(s).
- Data collected for one purpose can be repurposed only under strict restrictions. The processing of personal data for purposes other than those for which it was initially collected should be allowed only where the processing is compatible with the purposes for which the personal data was initially collected. In such a case, no legal basis separates from that which allowed the collection of the personal data is required. Moreover, further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes should be considered to be compatible lawful processing operations in such cases. In order to ascertain whether a purpose of further processing is compatible with the purpose for which the personal data was initially collected, the controller, after having met all the requirements for the lawfulness of the original processing, should take into account, inter alia: any link between those purposes and the purposes of the intended further processing; the context in which the personal data has been collected, in particular the reasonable expectations of data subjects based on their relationship with the controller as to their further use; the nature of the personal data; the consequences of the intended further processing for data subjects; and the existence of appropriate safeguards in both the original and intended further processing operations (Recital 50 GDPR).

In the HYDROPTICS project, the following guidelines are provided to the partners for them to conclude whether the data minimization principle is respected:

- When collecting personal data, ask yourself for which purpose you collect the data, how you are planning to use the data, and whether there is a way of achieving this purpose without having to collect the personal data. Document the choices you make in this process.
- Only collect the personal data that is strictly necessary to achieve the purpose, i.e., answering the research question(s). The partners should not collect personal data that is not compatible to the purposes of the processing.
- It is advisable not to keep the personal data stored longer than necessary to achieve the purpose, i.e., answering the research question(s), being able to prove validity of research outcomes, complying with legal obligations, etc.
- De-identification of personal data reduces the chance of identification. Anonymization (see previous sub-section) is the process in which you delete all information that may lead to identification of an individual. Consider indirect indicators and combinations of indicators as well, as these may lead to identification as well. Once personal data is properly anonymized, the data does not fall within the scope of the GDPR anymore.
- Another form of de-identification is pseudonymization (see previous sub-section), which offers a (temporary) solution when personal data is necessary to keep (for instance for longitudinal research or accounting for scientific integrity), but the personal data itself is redundant in the daily routine of processing and analyzing data. Pseudonymization refers to the process of replacing personal identifiers with codes that are stored in a different file on a different location. Keep in mind that pseudonymized data still remains personal data and therefore the GDPR still applies to this data.
- Repurposing personal data becomes an issue in the case the purpose is formulated in a (too) restrictive way in the informed consent procedure (e.g., “for this research project” or “by the involved researchers”). If you reasonably expect repurposing of personal data in the near future, we advise you to make sure your consent is formulated wide enough (e.g., “for research purposes” or “by researchers employed by a scientific organization”). However, data may only be processed for specified and explicit purposes (Article 5 GDPR).

Data breaches notification obligation

According to Article 33 GDPR, in the case of a personal data breach, the controller shall without undue delay and, where feasible, not later than 72 hours after having become aware of it, notify the personal data breach to the supervisory authority competent in accordance with Article 55, unless the personal data breach is unlikely to result in a risk to the rights and freedoms of natural persons. Where the notification to the supervisory authority is not made within 72 hours, it shall be accompanied by reasons for the delay. The processor shall notify the controller without undue delay after becoming aware of a personal data breach. The notification referred above shall at least:

- Describe the nature of the personal data breach including where possible, the categories and approximate number of data subjects concerned, and the categories and approximate number of personal data records concerned.
- Communicate the name and contact details of the DPO or other contact point where more information can be obtained.
- Describe the likely consequences of the personal data breach.
- Describe the measures taken or proposed to be taken by the controller to address the personal data breach, including, where appropriate, measures to mitigate its possible adverse effects.

6.8 Data protection impact assessment

One of main elements of the GDPR, introduced in Article 35, is the need to perform a DPIA in specific situations. Although not specifically described in the GDPR, a DPIA is considered as a process designed to describe the data processing and assess its necessity and proportionality. DPIAs are designed to help manage the risks to the rights and freedoms of natural persons resulting from the processing of personal data by assessing them and determining the measures to address them. DPIAs are important tools for accountability, as they help controllers not only to comply with requirements of the GDPR, but also to demonstrate that appropriate measures have been taken to ensure compliance with the Regulation.5 In other words, a DPIA is a process for building and demonstrating compliance and to avoid possible consequences of non-compliance which can cause a fine up to 4% of worldwide turnover for a company.

The GDPR places obligations on both the data controller, which “alone, or jointly with others, determines the purposes and means of the processing of personal data”, and the data processor, who processes personal data on behalf of the controller. However, the subject responsible for carrying out a DPIA is data controller. [24] If the processing is wholly or partly performed by a data processor, the processor should assist the controller in carrying out the DPIA and provide any necessary information.

When a DPIA is required

A DPIA [25] is required under the GDPR any time a new project is beginning that is likely to involve “a high risk” to other people’s personal information. The variety of discussion running nowadays related to impact assessments and the related GDPR direction around DPIAs demonstrate the difficulties to define borders which separate the need to do mandatory assessment (DPIA) from the simple adherence to GDPR principles. In general, the suggestion made by authorities is to run the assessment if not sure to be directly or indirectly involved in one of the cases which require that.

Furthermore, it is important to maintain compliance during processes, so an assessment will be a periodic action to be performed. The aim of those periodic assessment documents is to provide evidence of methodology, procedures and related outcomes that will be part of the compliance process activated by partners to provide their specific assessments related to Legal, Social, Ethics and Liability issues.

[5] See also Recital 84 GDPR: “The outcome of the assessment should be taken into account when determining the appropriate measures to be taken in order to demonstrate that the processing of personal data complies with this Regulation”.

In line with the risk-based approach embodied by the GDPR, carrying out a DPIA is not mandatory for every processing operation. [24] The GDPR demands that a DPIA be carried out "where a type of processing in particular using new technologies, and taking into account the nature, scope, context and purposes of the processing, is likely to result in a high risk to the rights and freedoms of natural persons."

However, there is no "silver bullet" method for carrying out impact assessments [26]: "What matters is the choice of an appropriate assessment method allowing for the best understanding and treatment of possible consequences of the envisaged initiative. These methods can range from qualitative or quantitative risk management to scenario planning, to scientific foresight, supported by a compliance check with relevant legal and otherwise regulatory requirements (e.g., technical standards)." [26] The GDPR sets out the minimum features of a DPIA (Articles 35(7) and Recitals 84 and 90):

- A description of the envisaged processing operations and the purposes of the processing;
- An assessment of the necessity and proportionality of the processing;
- An assessment of the risks to the rights and freedoms of data subjects;
- The measures envisaged to "address the risks" and "demonstrate compliance with this Regulation". [24]

A DPIA may concern a single data processing operation or could be used to assess multiple processing operations that are similar in terms of nature, scope, context, purpose, and risks. [24]

As indicated above, according to the Regulation, "where a type of processing in particular using new technologies, and taking into account the nature, scope, context and purposes of the processing, is likely to result in a high risk to the rights and freedoms of natural persons, the controller shall, prior to the processing, carry out an assessment of the impact of the envisaged processing operations on the protection of personal data". While this passage makes it clear that a DPIA is required by law under certain conditions, it is unhelpfully light on specifics. To help clarify the situation, here are some concrete examples of the types of conditions that would require a DPIA: [25]

- If you are using new technologies;
- If you are tracking people's location or behaviour;
- If you are systematically monitoring a publicly accessible place on a large scale;
- If you are processing personal data related to "racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation";
- If your data processing is used to make automated decisions about people that could have legal (or similarly significant) effects;
- If you are processing children's data;
- If the data you are processing could result in physical harm to the data subjects if it is leaked.

Additionally, the WP29 "Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679" provide the following criteria that shall be considered to define the need for DPIA:

- Evaluation or scoring;
- Automated decision-making with legal or similar significant effect;
- Systematic monitoring;
- Data processed on a large scale;
- Sensitive data or data of a highly personal nature;
- Matching or combining datasets;
- Data concerning vulnerable data subjects;
- Innovative use or applying new technological or organizational solutions;
- When the processing in itself "prevents data subjects from exercising a right or using a service or a contract".

In other cases, where the high-risk standard is not met, it may still be prudent to conduct a DPIA to minimize your liability and ensure best practices for data security and privacy are being followed during the progress of the project.

Although the project is not running "systematic" actions as described in GDPR regulation, and no sensitive personal data is expected to be processed in any way, being a research project in which some pilots will run for a limited period of time in a contained space, some actions carried out (e.g piloting activities, dissemination, etc.) could partially go under some of criteria specified bellow. In view of these, we suggested all partners involved as data controller to complete a dynamic DPIA. The GDPR makes it clear (Article 35 and recitals 89[6] and 91[7]) that the use of a new technology, defined in "accordance with the achieved state of technological knowledge" (recital 91), can trigger the need to carry out a DPIA. This is because the use of such technology can involve novel forms of data collection and usage, possibly with a high risk to individuals' rights and freedoms. Indeed, the personal and social consequences of the deployment of a new technology may be unknown. A DPIA will help the data controller to understand and to treat such risks.

For the purposes of compliance of the project activities with the GDPR provisions, we proceeded to an assessment in order to indentify risks that we may face as a result of the project related processing activities. The table below presents the identified data protection risks and measures to mitigate them.

[6]https://gdpr-info.eu/recitals/no-89/
[7]https://gdpr-info.eu/recitals/no-91/

RISKS RELATED TO THE PROTECTION OF PERSONAL DATA						
Risk ID	**Name**	**Description**	**Probability of occurrence**	**Impact**	**Risk response plan**	**Responsible partner**
LAWFULNESS, FAIRNESS AND TRANSPARENCY						
1.	Consent lacks informativeness	HYDROPTICS involves a wide range of technologies developed and operated by different partners. The technologies are connected to each other and operated both separately and commonly. Additionally, the project involves different data subjects and different pilots. The variety of all these elements as well as the complexity of technologies might create difficulties for a data subject to understand the flows of their personal data and subjects involved in the processing. This affects both lawfulness and transparency of data processing.	Possible	Severe	For different groups of subjects and for different pilots, the processing activities and the roles of the processing entities were defined. The consent forms varied depending on the data subjects and their role in the project (pilot, activities). Informational sheets were provided in addition to consent forms. At the stage of signing the informed consent, the data subject were asked if they fully understand its content. In case of non - understanding, the missing information was provided and any issues clarified. For example, the project's technology developers explained to the data subject how the technology in question works and how it processes their data. All this was monitored during the whole period of processing of data at every stage of their participation in the project. Moreover, material on the website and media about the project were served as additional source of information/clarification for data subjects.	DBC + all partners
PURPOSE LIMITATION						
2.	Purpose of data processing is not clearly defined	The purpose of personal data processing is conducting the research activities in the project. However, due to the complexity of the project, the mentioned purpose is deemed to be too wide and might lack sufficient pecification.	Possible	Severe	The general purpose was layered to sub - purposes and accompanied with clear description of the project and its goals (in informational sheets, on the website). This ensured that the purpose is detailed enough to determine what kind of processing is and is not included within the specified purpose, and to allow that compliance with the law can be assessed and data protection safeguards applied.	DBC
3.	Processing of personal data outside the scope of the	The project's pilots will engage end users working at end user HYDROPTICS partners. In this case, some of their personal data is already being processed	Probable	Severe	While engaging end users in pilots or other project activities, the conditions of their data processing (including purpose, legal basis, processing activities) was defined separately from the existing processing	DBC + all partners

	purpose it was collected for	by the respective partners. Depending on the description of initial purposes of data processing, it might be incompatible with processing activities in the project.			activities in their organisation. This enabled compatibility with the purpose.	
DATA MINIMIZATION						
4.		Hydroptics will collect data via different means and different technologies, which will be processed by different partners. It might happen that data collected by one partner for its purposes is provided to another partner but this data is not needed to achieve the goals of those partners	Possible	Severe	For every processing activity the scope of the data necessary to achieve the purpose of processing was defined. Additionally, the list of partners involved in that processing activity as well as their respective roles was specified. The process of filtering and cleaning of data was applied. Moreover, the data that is processed for aggregation purposes was pseudonymized.	DBC + all partners
INTEGRITY AND CONFIDENTIALITY						
5.	Insufficient security of data processing, transfer and storage	Hydroptics's technical architecture is complex and will include different layers and several means of and processing data (several types of devices, hardware, middleware). This all might create the security risks such as risks of data loss, breach of confidentiality.	Possible	Severe	The measures to ensure security of processing include data pseudonymisation and anonymization, secure middleware, data filtering and cleaning, encryption.	All partners
STORAGE LIMITATION						
6.	Different periods of data storage	Hydroptics includes different partners processing different types of data and with regards to different processing activities. Partners might store the data for different periods of time.	Probable	Severe	The Hydroptics partners shall agree on the minimum and maximum periods for storing personal data (might vary for different processing purposes)	DBC + all partners
ACCOUNTABILITY						

7.	The roles of partners are not clearly defined	Involvement of almost all partners in processing of data with respect to different purposes and activities creates the risk of lack of accountability ('everyone is responsible for everything'='no one is responsible')	Probable	Severe	All partners shall define their role (controller/processor of personal data), the partners they cooperate with and how. They will specify the purposes of data processing, types of data and relevant activities.	DBC + all partners
8.	Access to data by unauthorized subjects	HYDROPTICS includes different companies, organisations and universities. While some representatives are continuously involved in the project activities and are informed on the necessary procedures, other employees might get access to the data not being aware of the rules of its protection.	Probable	Severe	The HYDROPTICS partners provided the information on the person responsible for data protection in their organization (name and contact details). In compliance with art. 30 of the GDPR, HYDROPTICS partners shall keep the record of processing activities describing the type of data processed, by whom (including the person within organization) and for which purpose. The scope and amount of people having access to the personal data shall be limited.	All partners
RESPECT OF DATA SUBJECTS' RIGHTS						
9.	Limited right to erasure of personal data	HYDROPTICS will apply technologies similar to blockchain that make the erasure of data of a specific data subject technically challenging	Probable	Severe	To solve this issue, HYDROPTICS used the type of technology that allows to delete the pieces of the information from the chain without deleting the whole chain (e.g., IOTA). It proposes the use of structure digital identities to store all information gather for a data subject in a private tangle of sorts. What this means is that when TensorFlow preprocessors identify private information related to any known data subject, the information is placed on a special private tangle that belongs to the data subject. This private tangle is analogous to the IOTA SSI structure in that it stores private information about the data subject. The data subject to which the data belongs has then the ability to keep or delete the whole private tangle, without affecting the integrity of the main tangle, thus allowing it to forget the private data, in an efficient manner. This enabled compliance with the requirement of data erasure.	All partners
10.	Limited data portability	It is not defined if the data processed within HYDROPTICS might be technically	Remote	Severe	The control of data portability shall be carried out at the development and validation stages of HYDROPTICS's technical architecture.	All partners

		transferred to another data controller under the request of data subject				

During the life cycle of the project, we were able to ensure that such issues or any other relevant ones did not arise, thus ensuring the GDPR compliance of the project. In terms of the consent forms and their context, the instructions presented in both the privacy related deliverables (D2.2 and the current one) were implemented and the consent forms had the necessary context, ensuring that the data subjects had provided their consent freely and by having all the necessary information regarding the project's purposes and the processing of their personal data available. The project partners proceeded to the processing of the personal data both of the rest of the partners and of the relevant stakeholders according to the project's purposes and limited the processing operations to be absolutely important ones and according to their allocated tasks, while simultaneously remaining in contact with the data subjects regarding any additional information were to be provided to them. Finally, a personal data deletion plan has been conducted in order for the personal data of the data subjects to be successfully deleted following the end of the project, ensuring that no unauthorized and unlawful processing activities could take place in the future.

6.9 Ethical issues and societal concerns in HYDROPTICS

For all activities funded by the EU, ethics is an integral part of research from beginning to end, and ethical compliance is seen as pivotal to achieve real research excellence. There is clear need to make a thorough ethical evaluation from the conceptual stage of the proposal not only to respect the legal framework but also to enhance the quality of the research. Ethical research conduct implies the application of fundamental ethical principles and legislation to scientific research in all possible domains of research. The process to assess and address the ethical dimension of activities is called the Ethics Appraisal Procedure, in order to ensure that the provisions on ethics are respected.

General provisions

The partners must carry out the action in compliance with the ethical principles (including the highest standards of research integrity) and the applicable international, EU and national law. They must respect the fundamental principle of research integrity, as set out in the European Code of Conduct for Research Integrity. While carrying out with the project, the partners should:

- Respect human dignity and integrity;
- Ensure honesty and transparency towards research subjects and notably get free and informed consent (as well as assent whenever relevant);
- Protect vulnerable persons;
- Ensure privacy and confidentiality;
- Promote justice and inclusiveness;
- Minimise harm and maximising benefit;
- Share the benefits with disadvantaged populations, especially if the research is being carried out in developing countries;
- Maximise animal welfare, in particular by ensuring replacement, reduction and refinement in animal research;
- Respect and protect the environment and future generations.

According to Article 19(1) of the Regulation (EU) 2021/695 of the European Parliament and of the Council of 28 April 2021, "particular attention shall be paid to the principle of proportionality, to the right to privacy, the right to the protection of personal data, the right to the physical and mental integrity of a person, the right to non-discrimination and to the need to ensure protection of the environment and high levels of human health protection".

The partners should be capable to predict the cases in which ethical issues may be raised. Before the beginning of an activity as such, each partner must have obtained any ethics committee opinion required under national law and any notification or authorisation for activities raising ethical issues required under national and/or European law.

Research integrity

In order to ensure the necessary level of research integrity, the partners must follow principles listed below and ensure that the people carrying out research tasks comply with the European Code of Conduct for Research Integrity. The fundamental research integrity principles are:

- Reliability in ensuring the quality of research reflected in the design, the methodology, the analysis and the use of resources;
- Honesty in developing, undertaking, reviewing, reporting and communicating research in a transparent, fair and unbiased way;
- Respect for colleagues, research participants, society, ecosystems, cultural heritage and the environment;
- Accountability for the research from the idea to publication, for its management and organization, for training, supervision, and mentoring, and for its wider impacts.

Ethics and data protection

In this sub-section, we will mostly refer to ethical issues that may be raised and are related to data protection. Other research sectors in which ethical issues may arise but we do not consider them related to the current project are research on human embryos and foetuses, human cells, or tissues, human science research and research related to the environment, the health and safety of people or animals.

Data protection is both a central issue for research ethics in Europe and a fundamental human right. The right to data protection is enshrined in the EU Charter of Fundamental Rights and the Treaty on the Functioning of the European Union, which give effect to individuals' right to privacy by providing them with control over the way information about them is collected and used.

In research settings, data protection imposes obligations on researchers to provide research-data subjects with detailed information about what will happen to the personal data that they collect. It also requires the organizations processing the data to ensure the data is properly protected, minimized, and destroyed when no longer needed. Depending on the setting or information in question, the failure to protect personal data against loss or misuse can have devastating consequences for the data subjects. It may also have serious legal, reputational, and financial consequences for the data controller and/or processor. Many recent examples of unethical research practices have involved the unauthorized collection and/or (mis)use of personal data, resulting in enforcement action by regulators.

It should be highlighted that the fact that research is legally permissible according to data protection legislation does not necessarily mean that it will be deemed ethical. This especially applies in cases in which the processing of personal data raises higher ethics risks, when it involves: processing of "special categories" of personal data (formerly known as "sensitive data"), processing of personal data concerning children, vulnerable people or people who have not given their consent to participate in the research; complex processing operations and/or the processing of personal data on a large scale and/or systematic monitoring of a publicly accessible area on a large scale; data processing techniques that are invasive and deemed to pose a risk to the rights and freedoms of research participants, or techniques that are vulnerable to misuse; and collecting data outside the EU or transferring personal data collected in the EU to entities in non-EU countries.

In case of higher-risk data processing, a detailed analysis of the ethics issues raised by the project methodology is needed, which should comprise an overview of all planned data collection and processing operations, identification, and analysis of the ethics issues that these raise and an explanation of how you will mitigate these issues in practice.

Processing of "special categories" of personal data, according to Article 9 GDPR, should be treated with caution since, apart from the strict legislative provisions of GDPR, ethical issues may arise too. Project partners who are due to process special categories of data should first of all be able to justify and determine the necessity of their action according to the principles of accountability and purpose limitation, taking into consideration also the nature of the project and the individual tasks.

The process of securing the explicit and informed consent of data subjects is of utmost importance. The partners of the project should explain to research participants-data subjects what the research is about, what their participation in the project will entail and any risks that may be involved. Only after they have conveyed this information to the participants – and they have fully understood it – we can seek and obtain their express permission to include them in the project.

The consent should be provided according to the GDPR. This requires consent to be given by a clear affirmative act establishing a freely given, specific, informed, and unambiguous indication of the subject's agreement to the processing of their personal data. This may take the form of a written statement, which may be collected by electronic means, or an oral statement. Data subjects should be properly informed with the use of clear and plain language regarding their rights, all the data processing activities, the principles of data processing and the way they can withdraw their provided consent at any time. As a minimum, this should include the identity of the data controller and, where applicable, the contact details of the DPO, the specific purpose(s) of the processing for which the personal data will be used, the subject's rights as guaranteed by the GDPR and the EU Charter of Fundamental Rights, in particular the right to withdraw consent or access their data, the procedures to follow should they wish to do so, and the right to lodge a complaint with a supervisory authority, information as to whether data will be shared with or transferred to third parties and for what purposes and how long the data will be retained before it is destroyed.

Moreover, whenever the legal basis of the processing is the consent of the data subject, our partners should limit data processing to the purposes under which the consent was provided by the data subject. Repurposing or re-use of previously obtained personal data is for starters prohibited since it is a clear violation of Article 5(1)(a) GDPR, although, as mentioned above, the processing of personal data for purposes other than those for which it was initially collected should be allowed only where the processing is compatible with the purposes for which the personal data was initially collected. In such a case, no legal basis separate from that which allowed the collection of the personal data is required.

If our partners are going to use data that is publicly available, they must provide details of the source(s) and confirm that the data is openly and publicly accessible and may be used for research purposes. In the case data from social media networks will be used, our partners must assess whether those people actually intended to make their information public. Overall, in such cases, it should be examined whether the data subject had any reasonable expectation of privacy while making their data publicly available, alongside other factors, such as the terms of use and the privacy policy of the data controller of such platform.

Finally, ensuring data safety and integrity is a mandatory in terms of ethics. We have already highlighted to our partners the need of implementing the necessary technical and organizational measures in order to ensure data security and integrity. The use of anonymization and pseudonymization techniques during the collection and in general processing of personal data is of utmost importance for starters, while at the same time the necessary measures should be taken also to ensure that data are securely stored. Considering the "risk-based approach" of the GDPR, the necessary security measures should be implemented depending on the degree of the risk to the rights of the data subjects, and thus our partners should be accordingly informed and able to encounter any possible threats that may arise.

To conclude, our project partners should proceed with caution regarding any data processing actions that may raise ethical issues, such as the processing of special categories of data or the repurposing of data. Other potential cases in which ethical issues may arise could be the processing of data of children or vulnerable people but, taking into consideration the nature of each task of the project, our partners should refrain from any act of processing of such personal data.

7. Relevant Regulatory Frameworks in Turkey & Switzerland

The present chapter provides the overview of national legislations in non-EU Member States of HYDROPTICS's partners.

Turkish Data Protection Law (DPL)

Turkish Data Protection Law (DPL) was enacted in 2016. Turkey's supervisory authority, The Personal Data Protection Board (DPB), is still publishing assorted regulations and communiqués relating to it, as well as draft versions of secondary legislation.

Although it stems from EU Directive 95/46/EC, DPL features several additions and revisions. It does, however, contain almost all of the same fair information practice principles, except that it does not allow for a "*compatible purpose"* interpretation and any further processing is prohibited. Where the subject gives consent that data may be compiled for a specific purpose, the controller can then use it for another purpose as long as further consent is obtained, or if further processing is needed for legitimate interests.

The grounds for processing under DPL are similar to GDPR – saving that explicit consent is needed when processing sensitive and non-sensitive personal data. Inevitably, this is much more time-consuming. Such a burdensome obligation would initially make it seem that DPL provides a higher level of data protection compared to GDPR, but DPL's definition of explicit consent also has to be compared to GDPR's regular consent. '*Freely given, specific and informed consent* 'is common to both, while GDPR further requires '*unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her'*.

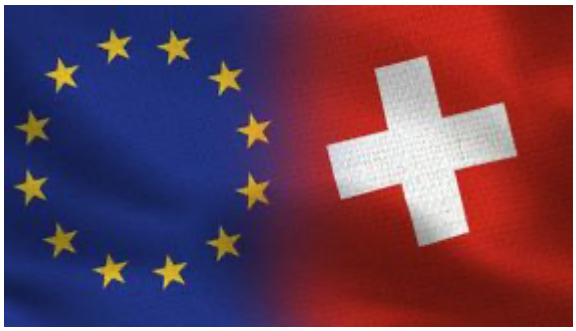
While DPL consent might appear to be less onerous than GDPR, no DPB enforcement action has yet occurred: interpretation of explicit consent therefore remains uncertain. Under DPL, the processing grounds for sensitive personal data are notably more limited than under GDPR – with the exception of explicit consent, the majority of sensitive personal data can be processed, but only if it is currently permitted under Turkish law. The sole exception is data relating to public health matters.

Equally burdensome under DPL is the cross-border transfer of personal data to a third country. As determined by the DPB, the country of destination must have sufficient protection – either that, or parties must commit to provide it. DPL also states that: "*In cases where interests of Turkey or the data subject will be seriously harmed, personal data shall only be transferred abroad upon the approval of the Board by obtaining the opinion of relevant public institutions and organizations"*. Under this provision, data controllers must decide whether a transfer could cause serious harm, and if it does, they need to obtain DPL approval. However, it is unclear how these interests might be determined.

Controllers have to maintain internal records under GDPR, whereas DPL does not make any general requirement to register with the data protection authorities. Instead, it has one notable point, DPL and GDPR are in harmony:

just as not complying with GDPR requirements carries substantial penalties, so does any breach of Turkish provisions. [8]

Switzerland, the Federal Act on Data Protection (FADP)



In Switzerland, the Federal Act on Data Protection (FADP) protects the privacy and the fundamental rights of natural and legal persons when their data is processed. It sets out the requirements for permissible data processing in accordance with the rule of law and therefore protects against possible abuses.[9]

Although Switzerland is not a member state of the European Union (EU), the country has been traditionally connected to the European legislation, in order to ensure a free flow of capital between the two regions. Following the provisions of the General Data Protection Regulation (GDPR), applicable at the level of the EU starting with 25th of May 2018, Switzerland had to modernize its national legislation concerning the protection of data.

The FDAP first passed in 1992 and is currently undergoing review to bring it closer to the standards set out in the GDPR. While the FDAP and the GDPR share many similarities, there are some important differences.

Perhaps most notably, while the GDPR only recognizes natural persons to be "data subjects" the FDAP recognizes both natural and legal persons.

The EU's proposed ePrivacy Regulation, would extend privacy rights to legal persons in much the same way as the FDAP.

Some of the other key differences between the GDPR and the FDAP are set out below:[10]

[8] Turkish Data Protection vs. GDPR: Spot the Difference (finance-monthly.com)

[9] GDPR and Switzerland - TermsFeed

[10] DPR and Switzerland - TermsFeed

	GDPR	**FDAP**
Scope	All private persons, businesses, charities, and local, national and EU-level public organizations processing personal data in the EU.	All private persons, businesses, charities, and federal government organizations **based in Switzerland**. Public bodies at the cantonal (regional) level are subject to local data protection laws.
Data subject	Natural persons only.	Natural persons ***and*** **legal persons** (e.g. corporations).
Standard of consent	Must be freely given, specific, informed, unambiguous, and given via a clear, affirmative action.	Must be "**given voluntarily on the provision of adequate information**." Must be given expressly when the processing concerns "sensitive personal data or personality profiles."
Maximum fine	4% of annual global turnover or €20 million.	250,000 Swiss Francs (CHF) (approximately €235,000).
Data subject rights	Right to be informed Right of access Right to rectification Right to erasure Right to restrict processing Right to data portability Right to object Rights related to automated decision-making	Only "**the right to information**" is explicitly set out in the FDAP. This is similar in scope to the "right of access" under the GDPR. Data subjects can also **request the rectification and erasure** of their personal data through Swiss civil law.
Data breach notification requirements	For a serious data breach likely to risk the "rights and freedoms" of individuals: inform the Data Protection Authority within 72 hours at the latest. For a very serious data breach that is likely to cause "high risk to the rights and freedoms" of individuals: inform the affected individuals without undue delay.	No formal data breach notification requirements.

8. Data Management Measures in HYDROPTICS

Data collection and management in HYDROPTICS were guided through the following principles/measures:

- The **privacy of participants** was respected. Personal information was kept confidential and guarantees of confidentiality were given to the participants.
- In the case of the participants registered at the HYDROPTICS platform, they did so with an ID-code instead of their personal information, which **guaranteed the confidentiality and the implementation of the data minimization principle**.
- The participants themselves had their **data sovereignty**. In case the participant wants the deletion of their data, the procedure described on the relevant sub section above was followed.
- The participant was able to change/ limit the access authorization of their data collected at the HYDROPTICS platform.
- Only information pertinent to piloting activities were collected.
- All researchers had the duty of confidence regarding collected data.
- The integrity of stored, processed, and published data was ensured by the researchers and the project consortium.
- Data that were collected by the participants at the HYDROPTICS platform were treated with care:
  - Participants were informed that the data could be used for the project.
  - Participants were informed in which way the data could be used.
  - The participants were informed who has the data sovereignty.
  - The participants were informed when the collected data will be deleted.
  - Appropriate measures, namely cryptography and physical security measures, were implemented by the researcher to protect the collected personal data.
  - Appropriate measures, namely cryptography and physical security measures, were implemented by the researcher to store and process data in secure manner.
  - In case the participant withdraws from the pilot, the collected data at the platform must be deleted or the access to them must be impossible for others, without any undue delay.

Pseudonymization is a technique that was implemented in order to reduce the chance that personal data records and identifiers lead to the identification of the natural person (data subject) whom they belong too. Identifiers make identification of a data subject possible. Pseudonymisation enhances privacy by replacing most identifying fields within a data record by one or more artificial identifiers, or pseudonyms. There can be a single pseudonym for a collection of replaced fields or a pseudonym per replaced field.

9. Consent within HYDROPTICS

9.1 Data collection activities

The data collection activities performed within HYDROPTICS events, strictly adhere to EC regulation as well as the legislation of individual Member States and Associated Countries. The following specific cases for data collection are immediately identified:

- **The collection of personal, non-sensitive data within the HYDROPTICS public events (workshops, pitstops etc.).** According to WP1, the organization of the HYDROPTICS advisory board workshops, demonstrations and public dissemination events, as a means to receive valuable stakeholder and end-user feedback within the

project's lifecycle. The collection of data via questionnaires or surveys, required within the HYDROPTICS workshops, **only entailed the collection of personal, non-sensitive data, as it is described in detail above in the legal basis subsection.** In such cases, the participants were presented **with a consent form, available both in English and the local languages** of the locations where each HYDROPTICS event were held. The personal data was pseudonymized whenever this was possible, taking into consideration its potential use for the purposes of the project.

- **Written and Audio/Visual documentation of the HYDROPTICS demos/pilots & dissemination events.** The HYDROPTICS demos were extensively documented by means of collecting written and photographic evidence, as well as audio/video capture. As previously stated, the participants of the project's pilots were debriefed and fully notified of all the pilot-related activities, including the documentation activities. Consent forms were made available to the participants available both in English and the local language in the event location. Volunteers were able to withdraw from these activities at any given time.

HYDROPTICS did not foresee the need to collect any kind of special categories of personal data of article 9 GDPR or personal data related to criminal offenses or convinctions of article 10 of GDPR and eventually no relevant need arised.

9.2 Consent requirements for the HYDROPTICS pilots

Informed Consent requires three elements: (i) **voluntary participation**, (ii) **competence** and (iii) **comprehension**. In order to conform to the requirements, set in place by the Nuremberg Code, the Declaration of Helsinki, the APA Ethics Code and relevant EU legislation, the Informed Consent forms included, at minimum, the following information:

- A statement that HYDROPTICS involves research subjects and an explanation of the main purpose.
- The expected duration of the subject's participation in the pilot activity.
- A description of the procedures to be followed with focus on the experimental procedures.
- A statement that participation is voluntary.
- Information about who is organising and funding the research.
- A description of any reasonably foreseeable risk, discomfort, or disadvantages. (First Aid and medical care will be available during the demonstration.)
- A description of any benefits to the subject or to others, which may reasonably be expected from the research, thus avoiding inappropriate expectations.
- A statement describing the procedures adopted for ensuring data protection/confidentiality/privacy including duration of storage of personal data and curation procedures.
- A description of handling of incidental findings.
- A reference to whom to contact for answers to pertinent questions about the research and research subjects' rights, and whom to contact in the event of a research-related injury to the subject.
- A statement offering the subject the opportunity to ask questions and to withdraw at any time from the research without consequences.
- An explanation of what will happen with the data or samples at the end of the research period and if the data/ samples are retained or sent/sold to a third party for further research.
- Information about what will happen to the results of the research.

Finally, the participants had to date, sign and initial the form, declaring that:

- They understand the purpose of the pilot,

- They have been given all the information that they have asked for,
- They agree to participate to the pilot,
- They understand that they reserve the right to ask for clarifications during the pilot and that they can withdraw at any given time.

Prior to the pilots, participants were guided through the consent form and all pilot activities by qualified research staff. The data subjects received any additional information they requested related both to their participation in the project and the processing of their personal data, thus ensuring that they had available all the information needed in order to participate freely and unambigiously.

9.3 Data lifecycle

Furthermore, participants were fully informed about information handling during all the stages of the data lifecycle, including:

- Where and how this information was stored. The partner responsible for the data collection and processing was responsible to ensure the security of the facilities and the confidentiality of the data, with support from the Project Coordinator and the Technical Coordinator.
- Who had access rights to it. Qualified research personnel had access to the data gathered from the participants after they have been pseudonymised. Consent forms were only accessed by the Coordinator and/or the Data Protection Officer of the organisation collecting and processing data.
- How long the (personal) data will be stored. Only during the project lifecycle. The related partner and the Coordinator is responsible to delete and destroy data sets after the project's conclusion, following the provisions of the personal data deletion plan mentioned above.
- How it will be pseudonymized and processed. The Data Protection Officer of the partner involved in these activities, will be responsible to pseudonymise any collected data sets.

The signed consent statements are being held on archive by the Project Coordinator and the involved partner and are relayed to the Project Officer (all relevant documents are part of the annual project management reports). Furthermore, **participants were properly informed regarding their ability to withdraw their consent on both participating in the project and having their personal data processed, and they way they could exercise this right of theirs**. Any partner involved in data collection and data processing was required to **provide assurance that the data will not be mishandled or utilized outside the expected scope of the project, along with the verification of their respective national Data Protection Authorities, when deemed necessary.**

10. Data Mapping

For the purposes of monitoring project activities to ensure GDPR compliance, DBC provided the partners and project participants with a questionnaire (see Annex III) and conducted interviews with all the project partners. As already indicated in the Grant Agreement, HYDROPTICS adopts a specific template for the definition of the regulatory compliance specifications for each component. The template related to any components that store or process personal data (internal or external to HYDROPTICS) and add security metadata.

The questions included in the questionnaire aimed to help document processing activities which relate to personal data processing. The partners were asked to fill in the questionnaire and provide the following information:

- Work packages and tasks, for the implementation of which, they need to process personal data.
- Type of personal data processed (e.g. name, surname, data of birth, address, health information, IP address, race, occupation, etc.).

- Type of processing (e.g. collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure, or destruction etc.).
- Automatic or manual way of processing.
- Role in the processing, either as data controllers or data processors.
- Identification of possible data processors.
- Source of personal data (directly from the individual or from another source).
- Categories of data subjects.
- Purpose of processing.
- Legal basis of processing (e.g. consent or other).
- Means to gain consent by a data subject.
- Place where personal data are stored and retained.
- Third parties which may offer hosting services for personal data.
- Retention time.
- Technical and organizational measures applied by the partners.

The partners completed the questionnaire by considering their activities. As a result, their answers indicated that, during the life cycle of the project, personal data processing for working on the project deliverables did not appear so often. Most of the partners did not process personal data, rather anonymized data. It shall be mentioned that besides processing of personal data of third parties, within the project processing of the personal data of the partners and the data subjects working for them took place for the purposes mentioned above on section 2.

From careful and thorough review of the answers provided by the partners, we can infer that processing of personal data took place in the cases mentioned below. It is noteworthy that the list is not restrictive, rather subject to changes and updates as the project proceeds.

- Project management: Personal data are processed for the purposes of the project management. Personal data refer mainly to data of the participants and data subject which work for the partners and who are part of the consortium. The legal basis for the processing of such data is mainly the performance of a contract, that would be the performance of the obligations agreed upon in the Grant Agreement.
- Trials: For the purposes of working on some deliverables (e.g. validation of results, testing etc.) the partners may engage real users and perform use case trials. In such a case, it is imperative that the data subjects shall give their informed consent as indicated by the current legislation.
- During the HYDROPTICS public events, workshops, public dissemination events: During such events, the organizing committee had to collect and process personal data of the participants. The data included information about the name, profession, contact details etc. of individuals and photographs of the events, which may include data subjects. The processing of personal data during such events was subject to a strong data protection policy and individuals were informed beforehand about the way their data are processed and provided their consent to participating in written form. It shall be mentioned that the processing of data during such events followed the guidelines set in the Grant Agreement – Ethics and Societal Impact Section (pg. 238-243).

11. Conclusion

This deliverable is the final report of our acitivities in GDPR compliance and legal issues management in Hydroptics project and provides a clear view regarding the processing operations thourghout the project's life cycle. The defined methodology (compliance framework of Hydroptics) on D2.2, published during month M24 of the project, was followed in detail during both pilot tests and in general the project related activities. As presented above, all the necessary procedures have been followed and all the necessary technical and organizational measures have been implemented in order to ensure GDPR compliance and data privacy and integrity.

Additionally, having into consideration the fact that the project is closing to the final stages of its life cycle, we have conducted and present with the current deliverable a plan of how to handle requests related to the rights of the data subjects in general and a plan of action regarding any personal data erasure requests that it is predicted to be filed as a result of the end of the project. We consider that the partners have received all the information needed for the forementioned procedures and during the review process, any additional information can be porovided to them, in order to ensure that they meet their GDPR related requirements and ensure compliance.

12. References

[1] European Commission (2022), *Horizon Europe (HORIZON) Programme Guide (Version 2.0)*, https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/programme-guide_horizon_en.pdf, retrieved on 03-01-2023.

[2] European Commission (2021), *Horizon Europe Data Management Plan Template (Version 1.0)*, https://enspire.science/wp-content/uploads/2021/09/Horizon-Europe-Data-Management-Plan-Template.pdf, retrieved on 03-01-2023.

[3] Science Europe (2021), *Practical Guide to the International Alignment of Research Data Management*, DOI: 10.5281/ZENODO.4915861.

[4] Consortium of European Social Sciences Data Archives (CESSDA) (2019), *Adapt your Data Management Plan: A list of Data Management Questions based on the Expert Tour Guide on Data Management*, https://static-archive.cessda.eu/content/download/4302/48656/file/TTT_DO_DMPExpertGuide_v1.2.pdf, retrieved on 03-01-2023.

[5] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), *OJ L 119*, 4 May 2016, pages 1-88.

[6] Jones, T. (2022), *What Is CIA Triad? – The Backbone Of Information Security*, https://medium.com/nerd-for-tech/what-is-cia-triad-the-backbone-of-information-security-712659c7206e, retrieved on 08-03-2023.

[7] GitHub, *Confidentiality, Integrity, Availability (CIA)*, https://github.com/0xsanny/guides/blob/master/src/pages/security/confidentiality-integrity-availability/index.md, retrieved on 08-03-2023.

[8] Wilkinson, M. D. et al. (2016), *Comment: The FAIR Guiding Principles for scientific data management and stewardship*, *Scientific Data, 3,* 160018.

[9] GO FAIR, *F1: (Meta) data are assigned globally unique and persistent identifiers*, F1: (Meta) data are assigned globally unique and persistent identifiers, https://www.go-fair.org/fair-principles/f1-meta-data-assigned-globally-unique-persistent-identifiers/, retrieved on 03-01-2023.

[10] GO FAIR, *F2: Data are described with rich metadata*, https://www.go-fair.org/fair-principles/f2-data-described-rich-metadata/, retrieved on 03-01-2023.

[11] GO FAIR, *F4: (Meta)data are registered or indexed in a searchable resource*, https://www.go-fair.org/fair-principles/f4-metadata-registered-indexed-searchable-resource/, retrieved on 03-01-2023.

[12] GO FAIR, *A1: (Meta)data are retrievable by their identifier using a standardised communication protocol*, https://www.go-fair.org/fair-principles/metadata-retrievable-identifier-standardised-communication-protocol/, retrieved on 03-01-2023.

[13] GO FAIR, *A1.1: The protocol is open, free and universally implementable*, https://www.go-fair.org/fair-principles/a1-1-protocol-open-free-universally-implementable/, retrieved on 03-01-2023.

[14] GO FAIR, *A1.2: The protocol allows for an authentication and authorisation procedure where necessary*, https://www.go-fair.org/fair-principles/a1-2-protocol-allows-authentication-authorisation-required/, retrieved on 03-01-2023.

[15] GO FAIR, *FAIR Principles*, https://www.go-fair.org/fair-principles/, retrieved on 03-01-2023.

[16] GO FAIR, *I1: (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation*, https://www.go-fair.org/fair-principles/i1-metadata-use-formal-accessible-shared-broadly-applicable-language-knowledge-representation/, retrieved on 03-01-2023.

[17] GO FAIR, *I2: (Meta)data use vocabularies that follow the FAIR principles*, https://www.go-fair.org/fair-principles/i2-metadata-use-vocabularies-follow-fair-principles/, retrieved on 03-01-2023.

[18] GO FAIR, *I3: (Meta)data include qualified references to other (meta)data*, https://www.go-fair.org/fair-principles/i3-metadata-include-qualified-references-metadata/, retrieved on 03-01-2023.

[19] GO FAIR, *R1: (Meta)data are richly described with a plurality of accurate and relevant attributes*, https://www.go-fair.org/fair-principles/r1-metadata-richly-described-plurality-accurate-relevant-attributes/, retrieved on 03-01-2023.

[20] *D1.1 Project Management HandBook*,

[21] Regulation (EU) 2021/695 of the European Parliament and of the Council of 28 April 2021 establishing Horizon Europe – the Framework Programme for Research and Innovation, laying down its rules for participation and dissemination, and repealing Regulations (EU) No 1290/2013 and (EU) No 1291/2013, *OJ L 170,* 12.5.2021, pages 1–68.

[22] Article 29 Data Protection Working Party (2017), *Guidelines on Data Protection Officers ('DPOs'), Adopted on 13 December 2016, As last Revised and Adopted on 5 April 2017*, 16/EN WP 243 rev.01.

[23] European Union Agency for Cybersecurity, *ENISA proposes Best Practices and Techniques for Pseudonymisation*, https://www.enisa.europa.eu/news/enisa-news/enisa-proposes-best-practices-and-techniques-for-pseudonymisation, retrieved on 03-03-2023.

[24] Article 29 Data Protection Working Party (2017), *Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679*, Adopted on 4 April 2017, as last Revised and Adopted on 4 October 2017, 17/EN WP 248 rev.01.

[25] Wolford, B., *Data Protection Impact Assessment (DPIA)*, https://gdpr.eu/data-protection-impact-assessment-template/, retrieved on 03-03-2023.

[26] Kloza, D. et al. (2017), *Data protection impact assessments in the European Union: complementing the new legal framework towards a more robust protection of individuals*, *d.pia.lab Policy Brief 2017/1*.

[27] Intersoft consulting, *Recital 89 Elimination of the General Reporting Requirement*, https://gdpr-info.eu/recitals/no-89/, retrieved on 03-03-2023.

[28] Intersoft consulting, *Recital 91 Necessity of a Data Protection Impact Assessment*, https://gdpr-info.eu/recitals/no-91/, retrieved on 03-03-2023.

[29] Council Decision (EU) 2021/764 of 10 May 2021 establishing the Specific Programme implementing Horizon Europe – the Framework Programme for Research and Innovation, and repealing Decision 2013/743/EU, *OJ L 167I* , 12.5.2021, pages 1–80.

[30] European Commission, Directorate-General for Research and Innovation (2021), *Horizon Europe Strategic Plan 2021-2024*, Publications Office, https://data.europa.eu/doi/10.2777/083753, retrieved on 03-01-2023.

[31] European Commission, Shaping Europe's digital future: *A European Strategy for data*, https://digital-strategy.ec.europa.eu/en/policies/strategy-data, retrieved on 03-01-2023.

[32] ALLEA – All European Academies (2017), *The European Code of Conduct for Research Integrity, Revised Edition*, I

Annex I: Data protection policy

1. Introduction

1.1. This document represents the policy of [partner's name] (hereinafter – **'we'/'us'/'Organization'**) with regards to the processing of personal data within the HYDROPTICS project.

1.2. HYDROPTICS project is a research project currently run under the Horizon Europe Framework Programme under the Grant Agreement no. 871529.

1.3. The EU-funded HYDROPTICS project through its technologies aims to design a trustworthy environment acting as a gatekeeper to information and data flows. Citizens and public/private organizations will be empowered to act and interact providing data both online and offline. HYDROPTICS will focus its activities on 3 main pillars: (i) the deployment of trustworthy, accountable and privacy-preserving data sharing technologies and platforms; (ii) the creation of data governance models and frameworks; (iii) the improvement of data availability, quality and interoperability – both in domain-specific settings and across sectors.

2. Scope of the policy

2.1. This policy only considers the processing of personal data by the Organization concerning its participation in the HYDROPTICS project. Other processing activities carried out by the Organization are outside the scope of this policy.

2.2. The participation of the Organization in the HYDROPTICS project will include [the type and description of activities, pilots where the Organization is involved, other relevant details].

2.3. The Organization's activities in HYDROPTICS will include the following processing of personal data: [types of personal data and data subjects, processing activities, and purpose(s) of processing].

2.4. Data will be processed according to "the principles relating to processing of personal data" of Article 5 GDPR, while the legal basis for processing personal data is defined by Articles 6 or 9 GDPR, depending on the personal data that will be processed and the circumstances under which the process is taking place.

3. Data protection principles

We support the principles set out by the GDPR by the following measures:

3.1. **lawfulness, fairness, and transparency of processing**: before and during the processing of personal data based on any legal basis stated in Article 6 GDPR, we provide the data subjects with information sheets describing the legal basis on which the processing is being done, the personal data that is going to be processed as relevant to the project and the processing activities. We inform them about their rights (mentioned in part 4 of this policy), providing them with all the necessary information regarding the way the subjects could proceed to any request associated with their rights and the way partners are obligated to respond.

3.2. **purpose limitation**: we only process personal data that is necessary to reach the goals of the project.

3.3. **data minimization**: we only collect and process personal data that is strictly necessary to conduct our activities in HYDROPTICS.

3.4. **accuracy of data**: we update/modify/erase the data upon request of the data subject or upon other discovery of its incorrectness.

3.5. **storage limitation**: we store the data during the term of the project and for [term] after its finishing; we irrevocably delete the data or anonymize it after the end of its processing.

3.6. **integrity and confidentiality**: we limit the scope of people having access to personal data to those who work in our organization and participate in the project; we store the personal data separately and use authentication procedures to control the access to it; [other applicable security measures to be added];

3.7. **accountability**: we use this policy to set and demonstrate compliance with the GDPR [other ways to be compliant and demonstrate it if necessary].

4. Rights of data subjects.

We respect the rights of data subjects specified in the GDPR, including:

4.1. the right to ask us what data is being collected about the data subject and how this data will be used in connection with the HYDROPTICS project ("right to access").

4.2. the right to lodge a complaint with a supervisory authority ("right to complain").

4.3. the right to request us to correct any of data subject's personal data that is inaccurate ("right to rectification").

4.4. the right to request us to erase, without undue delay, data subject's personal data ("right to erasure" also "right to be forgotten"), unless such a request would render impossible or seriously impair the achievement of the objective of that processing – including the impairment or invalidation of the research. According to Article 17(3) GDPR, such request can be denied in cases that processing is necessary: (a) for exercising the right of freedom of expression and information; (b) for compliance with a legal obligation which requires processing by Union or Member State law to which the controller is subject or for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller; (c) for reasons of public interest in the area of public health in accordance with points (h) and (i) of Article 9(2) as well as Article 9(3); (d) for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) in so far as the right referred to in paragraph 1 is likely to render impossible or seriously impair the achievement of the objectives of that processing; or (e) for the establishment, exercise or defence of legal claims.

4.5. the right to request us to restrict the processing of data subject's personal data ("right to restriction of processing").

4.6. the right to receive the personal data related to the data subject which he/she has provided to us and to transmit this data to another controller ("right to portability"); and

4.7. the right to object, at any time, to us regarding the processing of data subject's personal data ("right to object").

5. Other provisions

5.1. This policy is effective as of [date] till the end of the processing activities [the period after the end of the HYDROPTICS project to be defined].

5.2. This policy will be reviewed annually and updated if needed until the end of its effect.

5.3. We propose to the partners to strongly consider the appointment of a Data Protection Officer in cases they proceed to operations which require regular and systematic monitoring of data subjects on a large scale, especially if special categories of data are being processed.

5.4. All our employees having access to personal data specified herein, will be informed on this policy and other measures expected from them to be compliant with the GDPR.

5.5. The person controlling the implementation of this policy and other measures to comply with the GDPR from the side of the Organization is [name and contact details of the Organization's employee representing HYDROPTICS].

5.6. In case of questions regarding data protection rules and implementation of this policy the Organization will consult with [Name Surname (Entity Name), e-mail:]