



A Cost-Effective Photonics-based Device for Early Prediction, Monitoring and Management of Diabetic Foot Ulcers

Deliverable	D2.4 Data Management Plan (DMP)
Work Package(s)	WP2 - Project Management
Task(s)	T2.4 - Data management plan
Dissemination Level	Public
Due Date	30-04-2020
Actual Submission Date	31-08-2020
WP Leader	METIS BALTIC
Task Leader	METIS BALTIC
Deliverable Leader	METIS BALTIC
Contact Person	Panagiotis Terzopoulos
Phone	
Email	panos@metisbaltic.com

Document History			
Revision	Date	Author(s)	Description
v0.1	30-04-2020	Yung Shin Van der Sype, Mahault Piéchaud Boura, Jos Dumortier (Timelex)	Document structure initialization, Editing
v0.2	12-05-2020	Siri Luthman, Murali Jayapala (imec)	Editing
v0.3	25-05-2020	Eftychios Protopapadakis (NTUA)	Editing
v0.4	26-05-2020	Giannis Lazarou (EXUS)	Editing
v1.0	31-08-2020	Jos Dumortier (Timelex) Panagiotis Terzopoulos (METIS)	Editing, Final Review



PHOTONICS PUBLIC PRIVATE PARTNERSHIP

This project is an initiative of the Photonics Public Private Partnership. It has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 871908

This document and the information contained within may not be copied, used or disclosed, entirely or partially, outside of the PHOOTONICS consortium without prior permission of the project partners in written form

Contents

Executive Summary	4
1 Introduction.....	5
1.1 Structure.....	6
1.2 Applicable Documents.....	6
1.3 Acronyms.....	6
2 Research data	7
2.1 Definition.....	7
2.2 Collection purposes.....	7
2.3 Data sets.....	8
2.4 Re-use of existing data and origin of the data.....	11
2.5 Expected size of the data.....	11
2.6 Data utility	11
3 Findable, Accessible, Interoperable and Re-usable data (FAIR data)	12
3.1 Making data findable, including provisions for metadata.....	12
3.2 Making data openly accessible	13
3.3 Making data interoperable	14
3.4 Increase data re-use (through clarifying license)	14
4 Allocation of resources	16
5 Data security	17
6 Ethical aspects.....	18
7 Conclusion	19
References.....	20

Tables

Table 1: Acronyms.....	6
------------------------	---

Executive Summary

The purpose of this Data Management Plan (DMP) is to set out the data management plan for the PHOOTONICS project. It provides information on which research data will be made openly accessible in the PHOOTONICS project, their findability, interoperability and re-usability.

This document is not a static document and will be updated throughout the project, if necessary.

It covers the following aspects:

- Description of the research data in the PHOOTONICS project,
- The methods which will be used in respect of the FAIR data principles, and
- The safeguards to ensure respect of the FAIR data principles in terms of allocations of resources, data security and research ethics.

1 Introduction

Data Management Plans (DMPs) are considered as a key element of good data management. This deliverable is a living documents and provides the first version of the PHOOTONICS project DMP. I will be periodically updated throughout the Action. It describes the data management life cycle for the research data to be collected, processed and/or generated by the PHOOTONICS project.

As part of making research data **Findable, Accessible, Interoperable and Re-usable** (FAIR)¹, this DMP includes information on:

- What categories of research data will be collected, processed and/or generated by the PHOOTONICS project ;
- Which research data will be made publicly accessible and by which open access model ;
- Which methods will be used in respect of the FAIR data principles for handling research data during and after the project ;
- Which additional safeguards will be implemented to ensure respect of the FAIR data principles in terms of allocations of resources, data security and research ethics.

Given that the PHOOTONICS project participates in the extended Open Research Data pilot, this DMP is a mandatory deliverable of the project. Article 29.3 of the PHOOTONICS project Grant Agreement requires open access to research data. More specifically, the consortium needs to ensure that **digital research data generated in the action** are deposited in a research data repository and the consortium needs to make it possible for third parties to access, mine, exploit, reproduce and disseminate (free of charge for any user) the data, including associated metadata, needed to validate the results presented in scientific publications, as soon as possible, and any other data, including associated metadata, as further specified in this DMP.

The first version of this DMP is delivered within the first six months of the project and is based on the DMP template provided by the European Commission and in line with its ‘Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020’.

This DMP is a living document in which information can be made available on a finer level of granularity through updates, as the implementation of the project progresses and when significant changes occur, such as (but not limited to) the use of new data, changes in consortium policies (e.g. new innovation potential, decision to file for a patent, etc.), or changes in consortium composition and external factors (e.g. new consortium members joining or old members leaving).

The DMP contains a brief description of the research data types which will be collected during the PHOOTONICS project, explaining the procedures used to collect or create them, it contains details on copyright and IP issues, and details related to data storage, access rights, retention periods, anonymisation and pseudonymisation, consent procedures, etc.

¹ European Commission, “[H2020 Programme, Guidelines on FAIR Data Management in Horizon 2020](#)”, v3.0, 26 July 2016.

1.1 Structure

This document is structured as follows:

Section 1: Introduction

Section 2: Research data

Section 3: Findable, Accessible, Interoperable and Re-Useable data (FAIR)

Section 4: Allocation of resources

Section 5: Data security

Section 6: Ethical aspects

Section 7: Conclusion

1.2 Applicable Documents

AD1 PHOOTONICS Grant Agreement

AD2 PHOOTONICS Ethics Summary Report

1.3 Acronyms

Table 1: Acronyms

Term	Definition
CIA	Confidentiality Integrity and Availability
DFU	Diabetes Foot Ulcers
DMP	Data Management Plan
FAIR	Findable, Accessible, Interoperable and Re-Useable data
IP	Intellectual Property
IR	Infrared
PU	public

2 Research data

In this section an overview of the PHOOTONICS research data is provided.

2.1 Definition

The notion ‘research data’ covers a broad range of types of information and can be defined as “*the information, in particular facts or numbers, collected to be examined and considered as a basis for reasoning, discussion, of calculation*”².

Given the context of the Horizon Open Research Data Pilot, this DMP focuses primarily on research data that are available in digital form.

Some examples of research data include statistics, results of experiments, measurements, survey results, interview recordings and images, and field-study observation results. Not considered as research data are trade secrets, commercially sensitive information, and confidential information.

The types of data covered by the Open Research Data Pilot include ‘underlying data’ which need to validate the results presented in scientific publications, including the associated metadata (i.e. metadata describing the research data deposited), as soon as possible, as well as any other data (for instance curated data not directly attributable to a publication, or raw data), including the associated metadata, as specified and within the deadlines laid down in the DMP – that is, according to the individual judgement by each project.³

Research data generated within the PHOOTONICS project regard mainly medical imaging data acquired during the pilot validation. It may also include the data from the development of the technological components (e.g. spectral measurements, electrical/electronic characteristics of devices, methods used, etc.).

2.2 Collection purposes

Research data will be collected and processed in the course of the project for the following purposes and in relation to the following project objectives:

Data collection will support the detection mechanisms development and the smooth integration of the individual detection components to the PHOOTONICS ecosystem. Collected data will be used for developing reliable detection mechanisms for DFU monitoring and reduce the equipment cost, by exploiting every bit of raw data information that is available. A more detailed list, demonstrating the data collection purposes, is presented below:

- a. Selection and parameterization of the pre-processing tools, which support the software platform by improving the discrimination capabilities over the raw data
- b. Training, validation, evaluation & fine-tuning for a variety of deep learning models, capable to handle multisensory data, i.e. RGB, thermal, Hyperspectral, etc.
- c. Development of decision support systems which generate meaningful additional information, e.g.

² European Commission, “Guidelines to the rules on open access to scientific publications and open access to research data in Horizon 2020”, 21 March 2017, 4.

³ https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/open-access_en.htm.

indicate regions of interest on the examined foot and detect pathological elements.

- d. Establish a safe and reliable communication interface with the rest of submodules. That way, data exchange will become more efficient, allowing for faster and uninterrupted results interpretation, through proper visualization tools.

Data collection will also support the validation and testing of the developed devices: The reliability and the performance of the developed device for DFU monitoring, management and early diagnosis will be verified, validated and assessed by the three university hospitals involved in the project. The goal of the clinical studies is to assess the discrimination capabilities of the new device and derive factors that can prove that such a medical tool can improve management and diagnosis of the disease.

Each sensor delivered as part of the WP3 will be supplied with sensor calibration and system files. During system integration (WP6), further calibration and characterization data of the sensor performance will be acquired. Data acquired will evaluate the accuracy of experimental measurements and aim to establish the detection limits of clinical parameters based on data acquired by the sensors when integrated into the Phootonics device. Digital data will be generated in WP7 as HSI data is acquired of DFU during the clinical studies.

2.3 Data sets

In the first months of the project, all project partners were asked to describe the specific data sets that will be processed in the course of the PHOOTONICS project.

At this early stage of the project, the following types of research data were identified, to include: Test results,

- Documents (text, Word), spreadsheets
- Laboratory notebooks, field notebooks, diaries
- Test responses
- Collection of digital objects acquired and generated during the process of research
- Database contents (text, images)
- Sensor calibration and data provided in xml and .raw format
- Processed and raw image data from feasibility studies and the clinical study performed with HSI sensor (.raw and .xml)
- Models, algorithms, scripts
- Contents of an application (input, output, logfiles for analysis software, simulation software, schemas)
- Methodologies and workflows
- Standard operating procedures and protocols

All partners commit to continuously keep track of the specific data sets processed under the tasks they are leading and to report them internally by completing the following table:

NTUA, WP 6, T6.1	
Data set	<i>Thermal images dataset</i>
Description	<i>Thermal images obtained using IR sensors</i>
File format	<i>.TIFF, .PNG or other type of image files format.</i>
Metadata	<i>Additional data may involve age, sex, disease information, regions of interest locations, or any other information that can be used.</i>
Data sharing	<i>Data derives from the employed thermal sensors, is shared with backend data based, and is used by deep learning models. Data is transferred after the removal of private meta-data that could lead to identification of the patient (anonymization of the data).</i>
Archiving and preservation	<i>Data are stored locally, during development phase, online or at backend during evaluation phase. Backups locally, NTUA's laboratory's PCs. Data which is received by EXUS is stored at two of EXUS's servers, one that runs the back-end and one for backup.</i>
Additional information	<p><i>All data will be image format files. There will be, also, txt files, preferably JSON formatted, containing additional information (metadata).</i></p> <p><i>Data quality and size varies according to the IR sensor used.</i></p> <p><i>Image data will be anonymized. Personal information will also be excluded from the metadata files since it is not required for the training of the algorithms.</i></p> <p><i>Are you generating the data or sourcing it from elsewhere? Are there certain terms and conditions applicable?</i></p> <p><i>Is the data digital or non-digital? Both?</i></p> <p><i>How will the data be created or collected? What instruments or tools will be used?</i></p> <p><i>What transformations will the data undergo?</i></p> <p><i>Will the data be updated or become redundant as you make revisions and produce subsequent versions?</i></p> <p><i>Are you processing information that falls outside of the scope of this DMP? E.g. confidential information.</i></p>

NTUA, WP 6, T6.1	
Data set	<i>Hyperspectral images dataset</i>
Description	<i>Hyperspectral images obtained using project specific sensors</i>

File format	<i>.TIFF or other type of image files format.</i>
Metadata	<i>Additional data may involve age, sex, disease information, regions of interest locations, or any other information that can be used.</i>
Data sharing	<i>Data derives from the employed hyperspectral sensors, is shared with backend data based, and is used by deep learning models . Data is transferred after the removal of private meta-data that could lead to identification of the patient (anonymization of the data).</i>
Archiving and preservation	<i>Data are stored locally, during development phase, online or at backend during evaluation phase. Backups locally, NTUA's laboratory's PCs. Data which is received by EXUS is stored at two of EXUS's servers, one that runs the back-end and one for backup.</i>
Additional information	<p><i>All data will be image format files. There will be, also, txt files, preferably JSON formatted, containing additional information (metadata).</i></p> <p><i>Data quality and size varies according to the hyper spectral sensor used.</i></p> <p><i>Image data will be anonymized. Personal information will also be excluded from the metadata files since it is not required for the training of the algorithms.</i></p>

NTUA, WP 6, T6.1	
Data set	<i>RGB images dataset</i>
Description	<i>RGB images obtained using project specific sensors</i>
File format	<i>.JPG, .PNG, or other type of image files format.</i>
Metadata	<i>Additional data may involve age, sex, disease information, regions of interest locations, or any other information that can be used.</i>
Data sharing	<i>Data derives from any RGB optical sensors, is shared with backend data based, and is used by deep learning models</i>
Archiving and preservation	<i>Data are stored locally, during development phase, online or at backend during evaluation phase. Backups locally, NTUA's laboratory's PCs. Data which is received by EXUS is stored at two of EXUS's servers, one that runs the back-end and one for backup.</i>
Additional information	<p><i>All data will be image format files. There will be, also, txt files, preferably JSON formatted, containing additional information (metadata).</i></p> <p><i>Data quality and size varies according to the hyper spectral sensor used.</i></p> <p><i>Image data will be anonymized. Personal information will also be excluded from the metadata files since it is not required for the training of the algorithms.</i></p> <p><i>Data will be obtained by the partners of PHOOTONICS, or found online, if</i></p>

<i>related data set archives become available within project's duration</i>

2.4 Re-use of existing data and origin of the data

Image data on diabetic feet will be used at any time, as they become available. The following usage scenarios are considered:

1. Training classification models: These models will be responsible for identifying if specific categories appear on a given image, e.g. existence of fissure, erosion or ulcer, ulcer severity, etc.
2. Training object detection models: Similar to classification models but also including localization. In this case, detections are marked using bounding boxes around them.

Semantic segmentation: This is a pixel level classification. Image masks are provided over the original image, depicting in detail the regions of interest. Only original data will be collected for the pilot validation. Such validation does not require any use of pre-existing or secondary data.

For the development of the interfaces and the DFU the partners will primarily use existing knowledge materials (tools, methods, etc.) that have been used and validated before and that are published and already available. These existing materials will be selected on the basis of their relevance according to the project objectives and will be digested and formatted based on the needs identified by the consortium. The proper acknowledgment of this material and of the work of others will be guaranteed by applying appropriate citation and quotation methods.

The research data of the PHOOTONICS project will be original unpublished data, except where clearly indicated otherwise. The original data will be collected by the project partners.

Existing knowledge materials will be collected by scientific collection methods, using databases and repositories available to the project partners.

2.5 Expected size of the data

Data size rough estimation can be provided at a later stage, when we have the first results from all the employed optical sensors.

The total size of the research data is difficult to estimate at this point in the project. The size of the research data will in any case be linked to the KPIs set out in the Description of Action, although the consortium aims to exceed such values for the purpose of maximizing the project impact.

For the clinical studies, the expected number of patients participating is 200 (50%-50% gender, 70%-30% type 1, type 2, 30%-70% younger and older than 35 years old).

Specifically for imec's HSI camera, although subject to further refinement of the system validation protocol, the acquisition of a raw and processed HSI data (.hdr, .raw and .xml files) with the imec camera from one patient is estimated to be 5GB; extending this to the full study yields approximately 1TB of HSI data acquired from DFU patients. Although imec's HSI camera may be used in initial feasibility studies, the size of the data cubes may, however, be adjusted according to the requirements of system integration of the HSI sensor.

2.6 Data utility

The PHOOTONICS project aims to gather, assess and deliver research data that can contribute to any research field that can use optical information including medical reasoning, context interpretation,

semantics, health care system, decision support systems, patient monitoring, etc.

All data contributing and /or validating to project research findings, will be made available via online repositories in accordance with the limitation set out in this DMP. PHOOTONICS site will provide links on these datasets.

3 Findable, Accessible, Interoperable and Re-usable data (FAIR data)

Under Horizon 2020's principle of open access to data, research data must be **findable, accessible, interoperable and re-usable** (FAIR). Compliance with these four foundational principles will contribute to the use of data in future research.⁴

3.1 Making data findable, including provisions for metadata

For project-internal data sharing, such as the sharing of unfinished reports and meeting minutes, the PHOOTONICS project uses a dedicated private cloud (<https://cloud.phootonics.eu/index.php/>) in order to efficiently manage the project information amongst the project partners and to enable the preservation of research data and appropriate versioning of the documents. Following the consortium agreement and guidelines for working on documents, each version of each document is referred to in a list at the beginning of the document.

Naming conventions are as follows:

The naming conventions used will be defined by the coordinator but will be based on the naming conventions for the internal project library, *i.e.* '[Name] [Year] [Title] [Conference is applicable] ([Type]).Extension'.

The use of the following format is proposed for each file/document: "Date (yyyymmdd)_project_company_filename_author_version". For example, the file containing the minutes from the kick-off meeting will be called: '2020XXYY_PHOOTONICS_Minutes of Kick Off Meeting_ZZZ_Final'

If multiple versions of a document exist, this will be indicated at the end of the name of the document by using 'v1', 'v2',... and 'final'.

Data acquired with the imec HSI cameras a given a user defined name followed by a suffix identifying the nature of the data (e.g. _dark, _white, _reflectance). Since the name of the data acquired with the HSI camera is user defined, it may straightforwardly be made to follow the naming convention defined by the project coordinator. For integration of the HSI sensor into the Phootonics device, meta and image data can be adjusted according to system needs and project requirements.

Metadata will be created and stored using .txt files, JSON formatted. The name of each field and the corresponding value range will be defined at a later stage, depending on the project's requirements.

All open research data, including metadata, will be integrated in EU open access repositories such as Zotero and OpenAIRE. The partners agree to provide adequate metadata within the data sets in order to ease the interpretation of the data and to increase the identification, discoverability, re-use and preservation

⁴ M. WILKINSON, M. DUMONTIER,... B. MONS, "The FAIR Guiding Principles for scientific data management and stewardship", *Scientific Data* 3, 2016, 1-9.

thereof. Metadata is structured information describing the characteristics of the sources. A distinction is made between:

- Descriptive metadata, such as title, abstract, author and keywords,
- Administrative metadata which is used to provide information to help manage a source, such as when and how it was created, file type and other technical information, and who can access it.

In order to increase the findability, the partners will also include keywords or key-phrases describing the subject or content of the data, including relevant terms of the field.

Specifically for imec's solution: Currently, raw image data acquired with the HSI camera is supplied with an HDR file in which the image acquisition parameters are specified. hdr files and image data are automatically given the same name in the imec proprietary data acquisition and analysis software; data is thus automatically associated with metadata. The user may specify the naming of the image files such that it matches any to be agreed naming convention. For integration of the HSI sensor into the Phootonics device, meta and image data can be adjusted according to system needs and project requirements.

Other information that the research data contain include the reference period, project funding information (e.g. logo and information about the Grant Agreement and the action/programme that funds the project, official project name and project ID), release policy including dissemination rules, information about the collection of the data such as the data source, geographical coverage of the data, language and file format.

3.2 Making data openly accessible

Materials generated under the PHOOTONICS project will be disseminated in accordance with the Grant Agreement. The project deliverables that are marked as 'PU' (public) will be made openly available on the project website and related platforms such as Zenodo, OpenAIRE, etc... In PU deliverables all data will be anonymized.

All other types of data intended to be open will also be made openly accessible by default, as far as not restricted by IP or other rights. Data sharing is UAB Metis's responsibility. A dedicated database will be developed for sharing the research data with the interested parties.

Certain data fall outside the scope of the open access strategy. These include different types of data that can either be used to identify individuals or that are of a commercially sensitive nature. As a consequence, personal data of research participants, project partners or other stakeholders, raw qualitative research data from draft reports, unfinished work, personal notes, plans for future research, preliminary analyses, peer reviews, and communication outside of a test setting, fall outside of the scope of the open access strategy.

In principle, the open research data will be made available with the lowest technical threshold possible. For practical purposes, technical limitations could follow from the use of the hosting site. While the repository is accessible through any standard browser and the program is free to download and free of charge, it requires users to create a profile in order to have access to the knowledge library.

In principle, access to the project data will not be restricted within the consortium. Nonetheless, in order to protect the identity of research participants and in order to encourage participants to contribute freely and truthfully, all reporting and communication relating to research participants will be in a pseudonymized manner based on the initials or case numbers of the participants. Identifiers will be stored in order to allow identification and traceability for research validations and follow-up, but such storage will be organized

separately from the research data and in adherence to state of the art confidentiality and security standards (including encryption, access logs and seals). If the document itself cannot be made secure with a password, they will be stored in an encrypted container with password protection (e.g. 7zip).

HSI data, from **imec's camera and sensors** are stored in ENVI standard: <https://www.harrisgeospatial.com/docs/ENVIHeaderFiles.html>. Standard and publicly available software may be used to read the data acquired with imec's camera.. Additionally, they are also readable with imec's proprietary software (HSI Suite) and Matlab SDK. Software manuals and commented code is available to use software to read and analyse acquired data.

3.3 Making data interoperable

The (meta)data that will be made open and re-useable will be in line with most widely used terminologies, standards and methodologies in the relevant fields in order to facilitate interoperability and interdisciplinary interoperability.

From a practical perspective, standard file formats will be used, considering the following guidelines:

- Non-proprietary and not tied to specific software,
- Open, documented standard,
- Common format used by the scientific community,
- Standard representation (Unicode, ASCII),
- Unencrypted,
- Uncompressed.

Specifically for imec's solution, HSI data is stored in .raw and HDR image format (.hdr) commonly used within the scientific community. Both file formats are readable with several publicly available software. Metadata relating to the acquisition and data analysis, as stored in the .hdr files, follows the naming conventions of the imaging community. Standard and publicly available tools can be used to open these data (e.g multibandread in matlab). If needed, dedicated software, in house developed by imec, will be shared with consortium partners.

3.4 Increase data re-use (through clarifying license)

Specific details on licensing will be provided in D9.6 'First Interim Exploitation Report', D9.7 'Second Interim Exploitation Report', D9.8 'Final Exploitation Report', D9.9 'First Interim Report on IPR & Standardisation Activities', D9.10 'Second Interim Report on IPR & Standardisation Activities' and D9.11 'Final Report on IPR & Standardisation Activities'.

In general, the project consortium aims to apply a license such as e.g. CC-BY-SA 3.0⁵ as a general rule to all publications and content on the platform in order to facilitate the widest re-use as possible.

The consortium aims for all the open research data to be made available for re-use without any data embargo, meaning that all data will be made openly available and free to re-use upon their publication. However some research data may also be considered as personal data, as such may not be made available in open access. The consortium will dedicate an internal workshop to identify potential restricted data and

⁵ Attribution-ShareAlike 3.0 Unported

option for their potential re-use. The finding of this workshop will be integrated in a later version of the DMP

The data will remain reusable (labelled accordingly with the applicable licenses, e.g. Creative Commons) forever.

Quality assurance will take place during the project according to the procedures set out by the project coordinator.

Anonymized annotated image data and metadata files, i.e. research data, will be available for at least five years after the project is completed. Once the project is completed, if an open archive solution fails to support such task, we may use NTUA's desktops for hosting the data processed and or created by NTUA. Data that is stored at EXUS's servers (anonymized data) will be kept for five years after the project ends.

4 Allocation of resources

The publication of anonymized annotated data, i.e. images and corresponding masks is not expected to result in additional costs.

As described in the Description of Action, all project publications will be published as open access papers either following the gold or green model of the open access Horizon 2020 requirements. The reasons for this choice relate to the free nature and the online availability of data used for all activities, as the project is (i) built on previous research results (improved quality of results), (ii) fosters collaboration and avoids duplication of efforts (greater efficiency), and (iii) involves all types of end-users of research results (improved transparency of the scientific process).

As defined by the Grant Agreement, the partner responsible for data management is the project coordinator, UAB Metis.

Costs for open access fees have already been allocated when drafting the project budget at the proposal stage. These costs are included in the overall budget of the project coordinator.

The costs for long-term preservation of the data have been taken into account by the coordinator. There will be no further costs related to the preservation of data beyond the end of the funding period.

5 Data security

The PHOOTONICS project uses trusted third-party service providers in order to securely store and transfer research data in the course of the project.

The project coordinator has reviewed the terms and conditions of the third-party service providers to ensure that storage and transfer comply with the security standards required for the protection of personal data.

The Description of Action states that the PHOOTONICS project will use state of the art technologies for secure storage, delivery and access of personal data, as well as managing the rights of the users. In this way, there is complete guarantee that the accessed, delivered, stored and transmitted content will be managed by authorised persons, with well-defined rights, at the right time.

In addition, the consortium confirms to comply with the following guidelines in order to ensure the security of the data:

- Keep pseudonymised data and personal data of participants separated,
- Encrypt data if it is deemed necessary by the responsible project partner,
- Store data in at least two separate locations to avoid loss of data,
- Limit the use of USB drives and personal folders,
- Save digital files in one of the preferred formats (as set out above), and
- Label files in a systematically structured way in order to ensure the coherence of the final data set.

For a successful and secure data exchange between the systems and in order to ensure CIA triad systems are using security protocols to accomplice:

- **Confidentiality:** As mentioned above personal data are removed from the images (even the ones that are embedded, in case of that kind of data) in order to accomplish anonymization. Matching between an image (or other kind of medical data) and a person is done only when absolute necessary and this is strictly restricted to non-authorized personnel. Additionally, every user will have his/her personal account with the corresponding access.
- **Integrity:** In order to accomplish the integrity of the data, special hashing algorithms are used to ensure that the data will be exchanged without any unauthorized modification and in any transaction the receiver is able to check if the files are the ones that are expected.
- **Availability:** Data are stored in more than one places and by automated backup processes, they are all synchronized. In case of a malfunction of a database, the system will connect to another with identical data.

6 Ethical aspects

Ethical aspects of the project are dealt with in the ethical reports (D2.1, 2.2 and 2.3) and will be further elaborated in the forthcoming initial and final ethics reports.

7 Conclusion

The purpose of this Data Management Plan (DMP) was to set out the first data management plan for the PHOOTONICS project.

It provides information on:

- What categories of research data will be collected, processed and/or generated by the PHOOTONICS project,
- Which research data will be made publicly accessible and by which open access model,
- Which methods will be used in respect of the FAIR data principles for handling research data during and after the project,
- Which additional safeguards will be implemented to ensure respect of the FAIR data principles in terms of allocations of resources, data security and research ethics.

References

European Commission, Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020, 21 March 2017.