

# EUROPEAN COMMISSION

Horizon Europe Framework Programme  
HORIZON- CL4-2022-DIGITAL-EMERGING-01-03  
Advanced multi-sensing systems

GA No. 101093008

**Integrated, Modular, Multisensing,  
Mid- and Near- IR sensing platform**



**M3NIR - Deliverable report**

**D1.2 - Data Management Plan**

## Disclaimer / Acknowledgment



Funded by the  
European Union

Funded by the European Union (under grant agreement No. 101093008). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

## About M3NIR

M3NIR plans to boost photonics-based sensing technology in terms of performance, reduction of footprint, energy consumption and costs. The value chain to be implemented by the M3NIR project includes the optimization of the manufacturing approaches of mid-infrared photonic devices, the development of relevant electronics and the high-level integration of other components like microfluidics for the realization of high-performance sensors. With the extra focus towards miniaturized configuration, robustness, maintenance-free operation, low power consumption and cost-effective sensing, M3NIR aims to create a modular, highly adaptable and efficient multi-sensing platform, which will be validated in the domains of environmental monitoring and healthcare. The project makes use of novel schemes for component integration and packaging, enabling both scientific and industrial breakthroughs.

## M3NIR consortium members



## Document information

<b>Deliverable No.</b>	D1.2
<b>Related WP</b>	WP1
<b>Deliverable Title</b>	Data Management Plan
<b>Deliverable Date</b>	30 – 06 – 2023
<b>Deliverable Type<sup>i</sup></b>	Report
<b>Lead Author</b>	Dalius Kontrimavičius (METIS)
<b>Co-Author(s)</b>	-

## Document history

Date	Version	Prepared by	Approved by	Description
20/03/2023	V0.1	Dalius Kontrimavičius	-	First draft
29/03/2023	V1.0	Dalius Kontrimavičius	Geert Van Steenberge	First approved draft
09/06/2023	V1.1	Dalius Kontrimavičius	Geert Van Steenberge	Revised draft
30/06/2023	V2.0	Dalius Kontrimavičius	Geert Van Steenberge	Final

## Dissemination level<sup>ii</sup>

PU	Public	X
SEN	Sensitive, only for members of the consortium (including the Commission Services)	

---

## Publishable summary

The purpose of the Data Management Plan (hereinafter - DMP) is to technically describe the way data during the project life cycle will be handled. Like all the Horizon Europe community, **M3NIR** project is also highly committed to the open science requirements, thus actual purpose of DMP is to ensure the availability and utility of the project's research data. The plan outlines the measures that will be taken in order to maximize access and re-use of the data for further purposes and applications. Also, it should be noted, that plan is constructed following the DMP template structure and typical questions provided by the Funding and Tenders portal of the European Commission<sup>1</sup>.

Data summary section of the DMP describes the data that will be used in the project to the extent which is available at the initial project implementation stage. Summary includes the type and format, the purpose, the size, the origin of the data, basic info, whether already existing data will be re-used throughout the project. Data management should also consider the potential of the data to be used outside of the project.

DMP also should describe the measures to ensure, that data will be findable, accessible, interoperable, reusable (*FAIR<sup>2</sup> data* principle). Thus, document brings insights towards the practice that will optimize the potential of data finding, data deposition in the repositories, potential restrictions to the free access, metadata generation and availability aspects. Application of various uncommon standards, formats, variables and methodologies might interfere with the data interoperability and reusability, thus project-specific practise should be explained and justified in relevant cases. Measures towards the other research outputs except data should be foreseen as far as issues over the project results are not covered by other documents, Consortium Agreement, for instance.

Data management in compliance with FAIR principles might be associated with the allocation of extra resources and costs. That's relevant in particular while dealing with data security matters including provisions for data storage and recovery. Since **M3NIR** project will rely on open data repositories and their trusted services, extra costs for data processing are not expected. **M3NIR** doesn't consider development of own databases for research data storage, sharing except dedicated project website, which will function as a reference towards other trusted repositories mainly.

Project doesn't consider collection, processing and sharing of sensitive personal data and there are no ethical reasons that could impact data distribution and sharing foreseen. Still users of the published data should be cautioned via disclaimer to consider carefully the nature of the datasets before using them to remove legal liability from the data owner and the data publisher.

Finally, current **M3NIR** DMP is a living document, it's an interim version for internal use prior to the validation activities in the project. Updates over the data management will follow in line with the project developments.

---

<sup>1</sup> <https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/how-to-participate/reference-documents;programCode=HORIZON>

<sup>2</sup> <https://www.nature.com/articles/sdata201618>

---

## Contents

1. Introduction.....	7
2. Data summary.....	8
3. FAIR data .....	9
3.1. Making data findable, including provisions for metadata.....	9
3.2. Making data accessible.....	9
3.3. Making data interoperable .....	11
3.4. Increase data re-use .....	12
4. Other research outputs .....	13
5. Allocation of resources.....	14
6. Data security.....	15
7. Ethics .....	16
8. Other issues .....	17
9. Conclusions.....	18

---

## List of acronyms, abbreviations and definitions

<b>Abbreviation</b>	<b>Definitions</b>
<b>DMP</b>	Data Management Plan
<b>DOI</b>	Digital Object Identifier
<b>GDPR</b>	General Data Protection Regulation
<b>PDK</b>	Process Design Kit
<b>TRL</b>	Technology Readiness Level
<b>WP</b>	Work Package

# 1. Introduction

A DMP is a document that outlines from the start of the project, how research data will be handled both during the project and after project implementation phase will be over. In the **M3NIR** project DMP is scheduled as an integral part of WP1 – Project Management and task T1.3 – Data Management Plan.

Having in mind duration of the overall **M3NIR** project, which is far above 12 months, DMP has to be submitted as a deliverable. Deliverable D1.2 – Data Management Plan comes in the form of a report and describes the way data is generated, collected, and handled during the project. This is a living document, the interim version for internal use and submission at M6. The updated version at M29 following the main development phase and prior to the validation activities, and the final updated version will be delivered at the end of the project at M42.

*Table 1: Overview of deliverable D1.2.*

<b>Deliverable Number</b>	<b>Short deliverable name</b>	<b>Lead beneficiary</b>	<b>Type</b>	<b>Dissemination level</b>	<b>Due date</b>
D1.2	Data Management Plan (interim version)	METIS	R	PU	M6

## 2. Data summary

Will you re-use any existing data and what will you re-use it for? State the reasons if re-use of any existing data has been considered but discarded.

It's possible that data will be re-used. This data will likely come from similar projects, completed among project consortium partners, on the topic of photonics development and mid-IR sensing solutions for various applications. Cumulatively, the **M3NIR** partners have participated in more than 300 research projects during the last 5 years, many of which were developing photonics and mid-IR sensing solutions capabilities. Preliminary it is planned, that the experience – making use of advanced photonics subsystems, development of a novel laser-based sensing configuration, photonic integration technologies in various material platforms, early detection of diseases via high-accuracy, low-cost and non-invasive screening, etc. – will be exploited in the **M3NIR** project as well.

What types and formats of data will the project generate or re-use?

Types of data will include mainly fabrication recipes/methodologies/designs (e.g., CAD designs), diagrams, images, connection diagrams, plots, detailed PDKs, measurement data, characterization results, etc. For the photonic integrated circuits' design Graphic Database System (GDS) files will be used. PDKs will be generated in different formats – libraries of building blocks in GDS, Design Rule Checking (DRC) files as layouts, Ruby scripts and Python scripts for implementation of layouts.

What is the purpose of the data generation or re-use and its relation to the objectives of the project?

Data generation or re-use serves the creation of harmonized datasets. The data will be re-used for benchmarking, to compare it with the data within the new developments to allow new valuable insights.

What is the expected size of the data that you intend to generate or re-use?

This is a living document, it's interim version for internal use prior to the validation activities in the project. Presumably, generated data are expected to stay up till few hundreds of GBs. Updates over the expected size of the data will follow in line with the project developments.

What is the origin/provenance of the data, either generated or re-used?

The generated and re-used data come from past experience of project partners and empirical proficiency throughout the project life cycle.

To whom might your data be useful ('data utility'), outside your project?

The **M3NIR** projects sets itself multiple scientific and technological objectives up till TRL5, thus data generated during the project might be utilised among the other players in the area of multi-sensing spectroscopic sensor platforms, microfluidics, integration and packaging schemes, data acquisition and analysis toolkit development.



## 3. FAIR data

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

*Will data be identified by a persistent identifier?*

Data will be held as easy to find as possible using a persistent identifier. This will include keywords which are consistent within the relevant area. Specified repositories (for instance, [Zenodo](#)) assign a globally unique and persistent identifiers (DOI).

Project documents/deliverables, dissemination materials will follow clear and harmonized naming convention (D1.1.) for easier traceability over the data structure and different versions.

*Will rich metadata be provided to allow discovery? What metadata will be created? What disciplinary or general standards will be followed? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.*

The minimum set of metadata will be created: publishing date, author and co-authors, the institution(s) that the author(s) are working in, title of the article/data set, dataset description or abstract, place of the publication (an link to it). If repository will have no dedicated fields for the information like grant project name, acronym and number, this information should be noted in the dataset description/abstract.

If repositories outside the project website support such a functionality (for instance, DOI), supporting enriched metadata will be generated.

Orientation towards the descriptive metadata creation standard will help to ensure metadata compatability among different disciplines, which collaborate in the **M3NIR**.

*Will search keywords be provided in the metadata to optimize the possibility for discovery and then potential re-use?*

Specific keywords will be provided. They include specific terms related to photonics, sensing applications, platforms, environment to ensure data discovery through web services and search engines.

*Will metadata be offered in such a way that it can be harvested and indexed?*

Data will be offered in an open way for harvesting and indexing. Metadata of each record within the repositories is indexed and searchable directly in internal search engines immediately after publishing. Metadata for individual records as well as record collections are harvestable by the record identifier and the collection name.

### 3.2. Making data accessible

**Repository:**

*Will the data be deposited in a trusted repository?*

Data will be deposited in an accessible way as much as possible. The exception might occur due to specific information being identified as business confidential. In this case – accessibility might depend on such type of information and have corresponding limitations.

Data will be deposited using trusted repositories like [Zenodo](#), [B2SHARE \(eudat.eu\)](#), [ResearchGate](#) and

will be accessible through the dedicated project website ([menir-project.eu](http://menir-project.eu)).

Following the latest developments of the European Open Science Cloud ([EOSC](#)) and availability of fully operational production-grade infrastructure, use of EOSC and data depositing there might be considered at the later stages of **M3NIR** project.

**Have you explored appropriate arrangements with the identified repository where your data will be deposited?**

Different storage solutions were evaluated, and various options are still available. Information will be specified upon the maturity of the Project. The current solution – project website and open access repositories like [Zenodo](#), [B2SHARE \(eudat.eu\)](#), [ResearchGate](#) – were chosen as the best option in terms of costs/benefits.

All project deliverables and working documents will be available to authorized consortium members of the project through the internal project management tool and document repository – MS Teams workspace. Public project deliverables, the executive summaries of non-public deliverables will be published in the project website and will carry references to the public repositories.

**Does the repository ensure that the data is assigned an identifier? Will the repository resolve the identifier to a digital object?**

DOI will be applied to data sets whenever possible. [Zenodo](#) provides DOI and it's versioning to keep records updated and findable.

#### **Data:**

**Will all data be made openly available? If certain datasets cannot be shared (or need to be shared under restricted access conditions), explain why clearly separating legal and contractual reasons from intentional restrictions. Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if opening their data goes against their legitimate interests or other constraints as per the Grant Agreement.**

Not all the data will be made openly available. Specific data which will be identified as business confidential, will not be shared outwardly. For this kind of data **M3NIR** project Consortium Agreement Section 10 “Non-disclosure of information” provisions will apply.

Evaluation of data among partners will be carried out to sample certain variations of a material so that publications can be made, and data can be made as available as possible.

**If an embargo is applied to give time to publish or seek the protection of the intellectual property (e.g. patents), specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.**

It is found to be challenging to predict how long embargos status for certain types of data might be applied at the Project's initial stage. Therefore, it will be considered on a case-by-case basis. This kind of information is supposed to be clearer at the later stages of the project implementation. Specific data in case of embargos status for certain types of data and/or data on protection of the intellectual property will be updated in DMP during the lifecycle of the project.

However, it is foreseen that embargos status won't be required except the cases described above. Still some external repositories allow content to be deposited under an embargo status with the different end

date. The repository will restrict access to the data for a limited time period and the content will become publically available automatically after the embargo status is cancelled.

Will the data be accessible through a free and standardized access protocol?

Data will be made accessible through an open, free and standartized universally implementable access protocol.

If there are restrictions on use, how will access be provided to the data, both during and after the end of the project?

A general principle will be applied which ensures that there should not be any long-term restrictions to the accessibility and availability of public data. It is considered to be possible to access data independently, without any additional requests and/or costs, etc.

How will the identity of the person accessing the data be ascertained?

No personal data will be required for accessing the public data and no exhaustive evaluation of identities will be carried out. However, if *cookies are* applied for certain depositories, web-site, etc. – user data will be stored according to GDPR data privacy rules or any other legal standards that protects personal data of European Union citizens and affects any organization that stores or processes their personal data.

Is there a need for a data access committee (e.g., to evaluate/approve access requests to personal/sensitive data)?

No, data access committee will not be required for public data. No authorization is necessary to retrieve it.

### Metadata:

Will metadata be made openly available and licensed under a public domain dedication CC0, as per the Grant Agreement? If not, please clarify why. Will metadata contain information to enable the user to access the data?

Metadata will be made openly available and licensed under public domain. Metadata at [Zenodo](#) is licensed under CC0, except for email addresses.

How long will the data remain available and findable? Will metadata be guaranteed to remain available after data is no longer available?

There are no foreseen restrictions for availability and findability both for data and metadata. However, it can be updated throughout the lifecycle of the project considering the status of business confidential information and its availability.

Will documentation or reference about any software be needed to access or read the data be included? Will it be possible to include the relevant software (e.g., in open-source code)?

Data will be published in a widely applicable formats and should not cause any difficulties to access data under normal circumstances. In case any specific format under specific software will be generated, reference to this software will be noted accordingly.

## 3.3. Making data interoperable

What data and metadata vocabularies, standards, formats or methodologies will you follow to make your data interoperable to allow data exchange and reuse within and across disciplines? Will you follow community-endorsed

---

interoperability best practices? Which ones?

No specific vocabularies and standards for the project's data is foreseen, but every effort will be made to publish data in a format which is the most accessible and formatted in a computer readable way.

Metadata will be created in a descriptive way to enable identification, location and retrieval of information resources by users.

In case it is unavoidable that you use uncommon or generate project-specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies? Will you openly publish the generated ontologies or vocabularies to allow reusing, refining or extending them?

No uncommon or project-specific ontologies are foreseen in the project.

Will your data include qualified references to other data (e.g., other data from your project, or datasets from previous research)?

If additional datasets are needed to complete the data, or if complementary information are stored in a different dataset, qualified references and scientific links among these datasets will be described.

However, updates over the expected data sets and their links with data from previous research will follow in line with the project developments.

### 3.4. Increase data re-use

How will you provide the documentation needed to validate data analysis and facilitate data re-use (e.g., readme files with information on methodology, codebooks, data cleaning, analyses, variable definitions, units of measurement, etc.)?

All public datasets will have comprehensive documentation addressing the data structure, the definition of variables, and the units of measurement.

Will your data be made freely available in the public domain to permit the widest re-use possible? Will your data be licensed using standard reuse licenses, in line with the obligations set out in the Grant Agreement?

All published datasets will be made freely available and licensed under the standard reuse licenses.

Will the data produced in the project be useable by third parties, in particular after the end of the project?

Data, which doesn't fall under the scope of Consortium Agreement Section 10 "Non-disclosure of information" provisions will be reusable by third parties.

Will the provenance of the data be thoroughly documented using the appropriate standards?

The documentation and metadata of each dataset recognize the data provenance through proper citation of the source of information using the formats widely accepted by the scientific community.

Describe all relevant data quality assurance processes.

Data generated throughout the project will pass scrutiny following the M3NIR project deliverable "D1.1. - Risk Identification Management & Quality Assurance Plan". Deliverable 1.1 sets project internal data quality assurance procedures (non-public).

---

## 4. Other research outputs

In addition to the management of data, beneficiaries should also consider and plan for the management of other research outputs that may be generated or re-used throughout their projects. Such outputs can be either digital (e.g., software, workflows, protocols, models, etc.) or physical (e.g., new materials, antibodies, reagents, samples, etc.).

When other research outputs are generated, the management of those outputs is shared with their creators. Specific details towards the project results are described in the Consortium Agreement Section 8 “Background - Results”.

Beneficiaries should consider which of the questions pertaining to FAIR data above, can apply to the management of other research outputs, and should strive to provide sufficient detail on how their research outputs will be managed and shared, or made available for re-use, in line with the FAIR principles.

When other research outputs are generated, compliance with the FAIR principles shall be managed by their creators. Specific details towards the project results are described in the Consortium Agreement Section 8 “Background - Results”.

---

## 5. Allocation of resources

What will the costs be for making data or other research outputs FAIR in your project (e.g., direct and indirect costs related to storage, archiving, re-use, security, etc.)?

At the current stage, there are no extra costs being foreseen for the factors mentioned above. Project website is envisaged as a routine tool for project communication activities and marginal costs for extra functionalities (data repository, reference, for instance) are not considered as significant. Other repositories are available free of charge and indirect costs like for data deployment will be borne by a corresponding project partner.

This is a living document, it's interim version for internal use prior to the validation activities in the project. Updates over the expected data making FAIR costs will follow in line with the project developments.

How will these be covered? Note that costs related to research data/output management are eligible as part of the Horizon Europe grant (if compliant with the Grant Agreement conditions)

Data storage, archiving, re-use, and security costs associated with the project website are sought through the funding of project partners.

Who will be responsible for data management in your project?

METIS is responsible for the data management of all web services published on the M3NIR project website. Each published dataset has its own data manager(s) specified in the dataset itself or metadata. For the consistency of publicly available data, M3NIR project partners are requested to note METIS ([dalius@metisbaltic.lt](mailto:dalius@metisbaltic.lt)) their intention to deposit data/publications in any repositories outside the project website.

How will long term preservation be ensured? Discuss the necessary resources to accomplish this (costs and potential value, who decides and how, what data will be kept and for how long)?

The data will be retained for the lifetime of the repositories. Multiple repositories and their declared policies suppose, that data will be retained indefinitely.

---

## 6. Data security

What provisions are or will be in place for data security (including data recovery as well as secure storage/archiving and transfer of sensitive data)?

Website data security will be ensured by the hosting service provider.

External repositories have their own reliable data security policies, dealing with server management, data storage, back ups, physical access to the equipment security measures. Data files and metadata within the repositories are backed up nightly and replicated into multiple copies in the online systems.

No sensitive and/or personal data will be stored externally.

All project deliverables and working documents will be available to authorized consortium members of the project through the internal project management tool and document repository, MS Teams workspace.

---

## 7. Ethics

Are there, or could there be, any ethics or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA).

There are no ethical reasons that could impact data distribution and sharing foreseen in **M3NIR**. A disclaimer is associated with each dataset to remove legal liability from the data owner and the data publisher. Users are also cautioned to consider carefully the nature of the datasets before using them.

Will informed consent for data sharing and long-term preservation be included in questionnaires dealing with personal data?

No questionnaires dealing with personal data will be created throughout the project.



---

## 8. Other issues

Do you, or will you, make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones (please list and briefly describe them)?

Currently there are no plans foreseen to use another data management procedure(s), except data storage, data quality assurance measures, covered by **M3NIR** project deliverable “D1.1. - Risk Identification Management & Quality Assurance Plan”. Aforementioned deliverable sets project internal data quality assurance procedures, defines project file management platform, internal structure, data sharing guidelines.

---

## 9. Conclusions

The **M3NIR** project acknowledges, that development of DMP is one of the key elements to manage research data generated in the action responsibly and in line with the FAIR principles. Well elaborated DMP eliminates contingencies, provides guidelines towards data safeguard measures to be put in place, shows practical approach of the project to keep data valuable both for scientific community and for public in general.

This is a living document, it's interim version for internal use. The updated version at M29 following the main development phase and prior to the validation activities, and the final updated version will be delivered at the end of the project at M42.

---

### <sup>i</sup> **Deliverable Type**

Please indicate the type of the deliverable using one of the following codes:

R Document, report

DEM Demonstrator, pilot, prototype

### <sup>ii</sup> **Dissemination level**

Please indicate the dissemination level using one of the following codes:

PU Public

SEN Sensitive, only for members of the consortium (including the Commission Services)