



ICON

HORIZON-HADEA

WP1 Project Management
Deliverable D1.1 Data Management Plan

Main Author(s):

Aleksandra Kaszubowska-Anandarajah, TCD

Kaida Kaeval, TalTech

Lukas Bernhardt, FAU

Elisa Morón, CyRIC

Due date of deliverable: June 30, 2025

Actual submission date:

Deliverable type: DMP

Dissemination level: PU

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. The European Union cannot be held responsible for them. The project is supported by the SNS-JU.



Funded by
the European Union

Project information

Project Name:	Intent and Context-Aware Optical Networks
Grant Agreement:	101189703
Project Duration:	1 January 2025 – 31 December 2027

List of Contributors

Partner number	Partner acronym	Partner name	Name of contact
1	TCD	Trinity College Dublin	Nadia Bolshakova
1	TCD	Trinity College Dublin	Aleksandra Kaszubowska-Anandarajah
1	TCD	Trinity College Dublin	Laura Galaviz
2	ADTRAN	Adtran Networks SE	André Sandmann
4	CyRIC	Cyprus Research and Innovation Centre	Andreas Papadopoulos
4	CyRIC	Cyprus Research and Innovation Centre	Elisa Morón
7	TalTech	Tallinn University of Technology	Kaida Kaeval

Document control

Document version	Date	Author / Organisation	Change
V1.0	16/06/2025	Nadia Bolshakova/Trinity College Dublin	First draft
V2.0	23/06/2025	Lukas Bernhardt/FAU	Second version
V3.0	25/06/2025	Tomáš Horváth/BUT	Third version
V4.0	29/06/25	Aleksandra Kaszubowska-Anandarajah	Final version

Table of contents

Project information	2
List of Contributors	2
Table of contents	3
Executive Summary	4
Introduction	5
1. Data Management Plan	5
1.1. Data summary	5
1.1.1. Externally provided data	5
1.1.2. Project generated data	6
1.1.3. Data types and formatting	6
1.2. Data storage	6
1.3. Reference name	6
1.4. Data set description and metadata standards	7
1.5. Data sharing and preservation overview	7
1.6. Procedures for sharing and preserving data	8
1.7. Long-term storage and data preservation	8
2. FAIR data	9
2.1. Making data findable, including provisions for metadata	9
2.2. Making data openly accessible	9
2.3. Making data interoperable	10
2.4. Making data re-usable	10
3. Data governance	10
4. Data security	11
4.1. Data ethics	11
Conclusions	12
References	12

Executive Summary

This deliverable presents the Data Management Plan (DMP) of the project. It gives an overview of the data generated, collected, and stored within ICON and how the project and its partners will provide findable, accessible, interoperable, and reusable (FAIR) data.

ICON consortium abides by the European Commission's stand that the information gathered as part of an EU-funded project should not be paid for again each time it is accessed or used by the research community. In addition, this information should benefit European companies and citizens. Thus, ICON strives to make the scientific data produced within the project available online and no extra cost to European researchers, innovative industries and citizens. At the same time, careful consideration will be given to ensuring that no intellectual property rights (IPR) and commercial interests of participants are compromised. The consortium will also ensure the long-term preservation of the information.

The collection, management and sharing of sensing data, especially from live networks, evokes several technical, privacy, and security challenges. The legal aspects of data ownership vary drastically between different EU countries, further complicating the issue. In addition, there is no standard describing the format for sharing sensing data and part of the work within ICON is aimed at developing such a format. Thus, this DMP intends to be a living document and will be reviewed and updated periodically, as the project's implementation progresses and when significant changes occur. There will be formal reviews of the DMP in mid-project (M18) and then again in M30 (6 months before the end).

Introduction

Deliverable 1.1 presents the Data Management Plan (DMP) for the ICON project, which will support quality, efficiency and consistency in data management across all project partners. The main objective of this deliverable is to describe the data management lifecycle for the information to be collected, generated, processed, and shared (open access where possible) to ensure FAIR principles during and after the lifetime of the ICON project.

In consultation with the project partners, this DMP describes the procedures of data handling, the data categories used within the project, how data will be exploited or shared and made accessible for verification and re-use, as well as ethical aspects and data protection requirements. All data will be handled following European Union (EU) legislation, adhering to relevant national and international regulations.

As part of creating research data that is findable, accessible, interoperable, and reusable (FAIR), the DMP incorporates considerations on:

- the identification of the data types to be created, processed, and/or generated
- the identification of the methodologies and standards to be utilised
- the metadata specifications that will make the data traceable and comprehensible
- defining criteria concerning access, open access, and sharing of the data
- outlining the ways that the data will be conserved during and after the end of the project
- the indications regarding the archiving and preservation of the project's open datasets
- whether data will be disseminated/made open access.

By adhering to the guidelines and best practices outlined in this DMP, the ICON project aims to enhance its operational efficiency, drive innovation, and achieve its goals effectively.

1. Data Management Plan

As the DMP describes the data handling involved in the ICON project, partners' consultation is fundamental to identifying the types and formats of data that will be generated and collected by them and their provisions to comply with the FAIR principles and ensure open access to them (wherever possible).

To fully comply with this DMP, the Project Coordinator has requested the relevant information from the ICON partners and will continue to collect details once the working groups make substantial progress.

1.1. Data summary

This section addresses the origin, types and formatting of data expected to be generated during the execution of the ICON project. So far, two categories have been identified in the first part of the project: external data provided to the project and data generated within the ICON.

1.1.1. Externally provided data

Externally provided data refers to the data that may be available through external sources rather than being generated/firsthand by the ICON team. Provided data can come from public databases, commercial sources, collaborative partners, and historical records, among others, such as pre-existing data, public datasets, and collaborative data by consortia/networks that will be used to complete and achieve tasks, deliverables, and objectives from the ICON project.

When confidential or sensitive data is contributed to the ICON project by external sources/parties, it will be governed by a specific Non-Disclosure Agreement (NDA) which will specify the restrictions and/or conditions for its implementation or exploitation, between the external party and the Consortium

members involved in the work package, in which the data will be used. If such data are publicly available, the consortium may use them only if their license is not violated.

1.1.2. Project-generated data

ICON will generate qualitative and quantitative research data (measurement data, calculation data, figures, software tools, hardware/experimental details) arising from the research programme and research dissemination activities.

Therefore, project data refers to all the information and datasets generated, collected, processed during the lifetime of the ICON project including, but not limited to, any produced directly by hardware and/or algorithm other than personal data, all output of measurements, raw data, refined data, analyses data, geodata, evaluation data and electronical data and/or data in writing produced by the ICON project's activities.

1.1.3. Data types and formatting

We will investigate different approaches to data formatting and decide on the most appropriate one based on the work carried out in T2.3 and T7.4.

ICON will deal with different classes of data (e.g., raw data, pre-processed data), with specific requirements regarding their description and metadata, archiving, and potential availability with open access. In addition, each data source (e.g. sensing interrogator or transceiver) provides data in different formats. To effectively combine the data by, e.g. means of correlation, the data structure needs to be aligned or converted, and the timestamps need to be synchronised or generated with the same format and accuracy.

Additional data types may be included during the ICON project execution as they are identified and defined. The following sections present how these aspects will be managed within the project.

1.2. Data storage

The project data will be shared across the team and organised using standardised naming conventions. All the documents generated by the consortium will be saved and shared through the institutional Microsoft SharePoint, managed and maintained by the project coordinator. Also, version control will be managed within this system. The raw sensing data, on the other hand, due to their large volume, would not be stored in one location. Instead, each partner will hold their data locally, while publishable open data will be stored at Zenodo. Finally, processed data will be uploaded to the cloud platform deployed as part of work in WP4 (T4.3).

1.3. Reference name

For the documents and other types of data files generated during the ICON project, it is highly recommended that the consortium follow a common naming scheme: ICON_

[Type]_[Name]_[Date]_[Partner]_[Version], where:

- [Type] is the data type (e.g., raw or pre-processed data)
- [Name] is a short and distinct name for the data,
- [Date] is the date when the data was produced/shared with format YYYYMMDD,
- [Partner] is the short name of the partner that generated the data,
- [Version] is an integer value starting with the number 1 and incremented by one for every new version.

1.4. Data set description and metadata standards

For the purpose of this DMP, a dataset can either be a single file or a set of correlated files with their associated metadata. Since a dataset may contain files of different formats, the metadata is used to describe and index the data in the dataset. Basic metadata will be used in front-page information to facilitate the efficient recall and retrieval of information by project partners and external referees. The metadata must contain at least the following information:

- Identifier(s): one or a set of unique identifiers (e.g., DOI, deliverable number);
- Creator(s)/author(s): name(s) and affiliation(s) of the partners that generated the data
- Title: a distinct name for the data set
- Publisher: the entity publishing the dataset
- Date: the year and month of publication, optionally the day and time can also be included
- Type: the type of the data, as defined in the section 1.1.3
- Format: the format(s) of the data included in the data set
- Keywords/subjects: terms that describe the main topics or themes of the dataset
- Description/abstract: a comprehensive description of the data, including its collection scenario (in the case of experimental data), or the statistical assumptions (in the case of statistical data)
- Language: the language(s) in which the dataset is available
- Rights/usage terms: information about the rights and usage terms associated with the dataset
- Availability: open/upon request/unavailable
- Embargo date: date at which the data will be openly available.
- Deletion date: if the data is stored temporarily for the usage in upcoming WPs, for verification purposes of a concept, test set up, or for comparisons with other datasets, and not for publishing. The initial deletion date for such data is the 31st of December 2027, but its storage can be extended in 2-year periods, when requested by project partners. The decisions to extend the storage duration of any data used or generated during the ICON project can be made by the Project Coordinator.

The mandatory information described above will be used to enable the indexing and harvesting of the data sets produced by the project, allowing for a quicker and more contextual search. Additionally, the data sets will be accompanied by a detailed description of the data collection and/or analysis of their variables, including descriptive variable names and a short description of their unit. Finally, data labelling will also be included, in case the data contains known events. This labelling could include the location, timestamp (e.g., using a Unix Epoch timestamp (milliseconds/ microseconds), and type of event.

1.5. Data sharing and preservation overview

All data with Consortium access level generated in the ICON project will be accessible during the project to all partners.

The data deposited on the sharing platform will be stored and preserved by the Project Coordinator and will be available only to the project consortium. The sharing and preservation of the data generated within the ICON project are highly dependent on the data type. A priori, each consortium partner is responsible for maintaining and preserving the data generated from their activities. This platform will be maintained by the project coordinator, who will continue to support and host the platform at least 24 months after the project end, unless otherwise agreed with the partners. It is worth pointing out that security standards are in place to protect the shared folder from unauthorised access and disclosure of confidential information. Regardless, it should be noted that sensitive data should not be shared through the shared folder via Microsoft SharePoint and should be securely maintained on the internal servers of project partners.

Online sharing platforms will be used for open and controlled data access. In particular, Zenodo has been identified as the platform of choice as it allows storage of 50 GB of data per file with 100 files per data set. As a result, the ICON community was created in the Zenodo platform for open sharing of data

https://zenodo.org/communities/icon_eu_project/settings. All project-related (open) publications and datasets will be added there. Other material, such as leaflets and brochures, will also be included there for completeness.

Additional sharing and development tools can be set up for specific tasks if needed.

All public deliverables, after acceptance, will be available on the ICON project website <https://iconproject.eu/>.

1.6. Procedures for sharing and preserving data

The experimental data will be gathered from lab tests and testbeds owned by academic institutions as well as from field trials, organised on top of the infrastructure owned by the project or external partners. As a result, the sharing and dissemination of the data can vary significantly depending on its origin. While we expect all the lab-based data to be freely shareable, the field-trial data will be subject to careful consideration to ensure that it does not contain any sensitive information, e.g., position of military vessels in case of interrogator-based sensing systems, or business-sensitive quality of transmission data in case of transceiver-based sensing data. Thus, each data set will be initially classified to determine if any sharing restrictions need to be placed on it.

Part of the work in WP2 (Task 2.3) is concerned with the investigation of the standard method of presenting and sharing the sensing data for use in telecommunication and non-telecommunication related applications. This includes the investigation of the necessary data pre-processing to remove the security or sensitive information from the data before sharing it. The preliminary workflow showing the blocks of data collection, transmission, storage, and sharing, is presented on Fig. 1. This workflow can change during the project timeframe, to reflect the changes in the sensing technologies, changing legal restrictions or requirements, or to reflect the possible automation of data collection, aimed at achieving uniform appearance and structure of the collected internal or sourced external data.

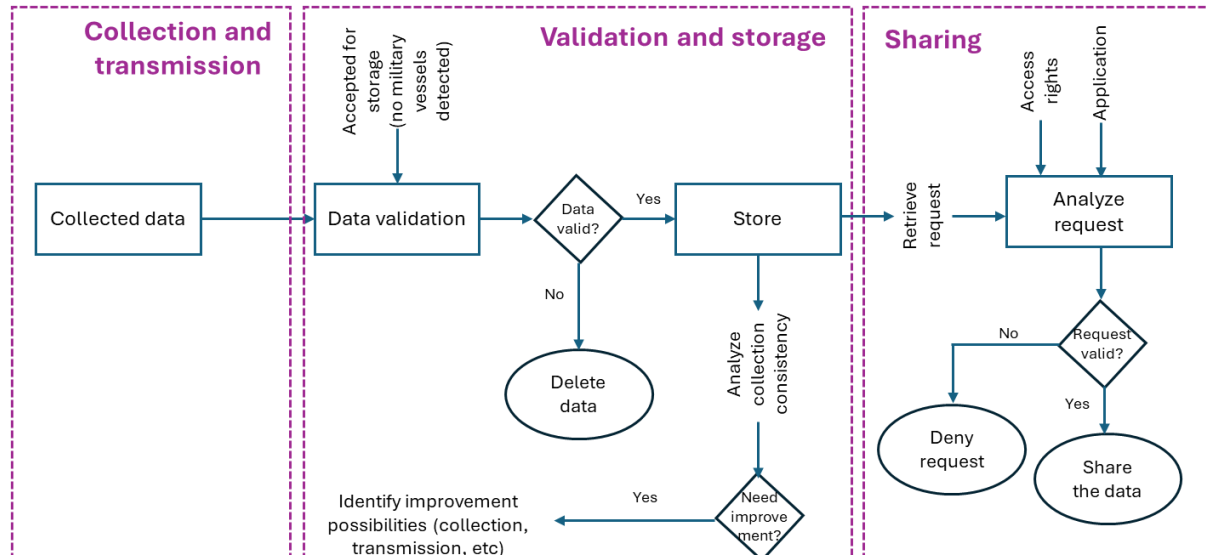


Fig.1: Preliminary blocks in the data management workflow

1.7. Long-term storage and data preservation

Scientific and technical papers, posters, videos, and other publications. Data from this category is meant for sharing with the research community or even the general public. All scientific publications will be made following the green or golden “open access” model. Non-scientific publications will be shared freely with every interested party. This data is related to the work carried out within T7.1, which runs throughout the project. Information on the planned publications is included in the Dissemination Plan (D7.8). Newsletter registrations and sign-ups will be performed via the project’s website. For

signing up, only basic information is requested, e.g., name, surname, and email address, to allow proper subscription to the newsletter list and automated dissemination of newsletters via email sent directly to the registered address. This process complies with EU GDPR for basic information collected and acknowledged by the subscriber, with required acceptance of the ICON project privacy policy. Publications will be shared in the form of PDF, video or photo files. The possible use or re-use by the scientific community is provided by adhering to the terms and conditions set out in the Grant Agreement, as well as to the terms of the publisher.

Information about tools and instruments: At this stage, it is not foreseen to require specialised tools. Any datasets used in scientific papers will also be accessible through the same location.

Accessibility: Once a publication is made available on the project repository, it will be freely available to the community in the way described in the appropriate section below. No embargo period is foreseen.

All shareable data will be published and hosted as per individual availability on the online repository: https://zenodo.org/communities/icon_eu_project/. Zenodo is free of charge. The community is curated by CyRIC, and new data might be added even after the project ends, as explained in the previous section.

2. FAIR data

The ICON consortium is committed to embracing and following the FAIR principles for data management throughout the project lifecycle. In this way, the ICON project will ensure that all outcomes are available for consultation, research, innovation, and development purposes and adequate communication and dissemination of the ICON project results.

This section describes how the project ensures its generated data is findable, accessible, interoperable and reusable.

2.1. Making data findable, including provisions for metadata

The ICON consortium has established standard data set names and a list of mandatory descriptive metadata that will be provided for each data set generated within the project. The complete list of required metadata is provided section 1.3 Dataset description and metadata standards. Despite the diversity in specific fields, many metadata standards share common elements, which the ICON project has observed, as seen in the list of mandatory metadata stipulated above.

2.2. Making data openly accessible

Project data produced for deliverables with a public dissemination level will be open to the general public. They will also follow Open Research Data Pilot (ORDP) principles (an initiative under the Horizon 2020 and Horizon Europe programs of the European Union) since the principles of the ORDP align closely with the FAIR data principles. This data could be used for scientific publications and made available through publication in specific open-access repositories such as <https://data.mendeley.com>.

Additionally, the website will consolidate and index or link to all the openly available data generated by the ICON project. Through the project's social networks, the research community will be notified about the availability of recently opened data. In cases where large amounts of data are stored on online sharing platforms, these platforms are responsible for indexing and preserving the data beyond the duration of the project.

2.3. Making data interoperable

The ICON project data seeks to be compatible with the European Open Science Cloud (EOSC). To this end, the ICON project data will be generated or collected in standard file formats, making them readable to anyone with the corresponding software (see section 1.6), when the access rights to the data are granted.

To ensure data interoperability, the ICON consortium will use an applicable standard glossary/vocabulary from recognised literature in the field. Abbreviations and acronyms will be fully spelt out and explained upon their first use, with detailed explanations provided in a dedicated chapter of this DMP.

Additionally, each ICON partner is committed to documenting the project data produced comprehensively, including methodologies, definitions, and context, to ensure its clarity and understanding.

2.4. Making data re-usable

The measures defined in the previous sections contribute to making the data shared by the ICON consortium highly reusable.

The generation of detailed metadata permits the data generated by the project to be used by other instances as industrial and academic entities. The project data produced with open access will be preserved by the online sharing platforms, therefore making the data reusable beyond the duration of the project. Also, the data shared as plain text or accompanied by detailed versioning of programming languages will make it possible to open and use the data in the future.

In addition to the aforementioned measures, the datasets will be accompanied by a detailed description of the data collection and/or analysis of its variables, including descriptive variable names and a short description of its unit (see section 1.4). The license, under which the data is shared, will be carefully analysed, always selecting the most permissive one, while preserving the interests of the generating partner.

To ensure the quality assurance process, data sets will be reviewed by at least two partners other than the generating partner(s). The review process will be triggered by the partner generating the dataset and will involve analysing whether the naming, metadata, and accompanying source code meet the expected quality established in this document. Review rounds can be established if needed. The detailed review reports will be private to the consortium; however, the reviewers will be acknowledged in the public version of the datasets as reviewers.

The consortium's obligations for reusability will be addressed in the Consortium Agreement. Research data will be licensed under the latest version of the Creative Commons Attribution International Public Licence (CC BY) or Creative Commons Public Domain Dedication (CC 0). The accessibility of data/research outputs will adhere to all legal and contractual obligations.

3. Data governance

Effective data management relies on clearly defined roles and responsibilities to ensure the responsible handling and use of data throughout the project lifecycle.

The following roles and responsibilities play a crucial role in the data management of the ICON project:

- Project Coordinator (TCD) is responsible for overseeing and coordinating the data management activities. The Project Coordinator ensures that the data management plan is effectively implemented and that all data-related tasks are carried out in accordance with the relevant policies and guidelines. The Project Coordinator works closely with the WP Leaders and Task Leaders to ensure seamless communication and coordination between team members in relation to data

management. The Project Coordinator serves as a central point of contact for data-related matters, facilitates data sharing and collaboration, and ensures that data management responsibilities are assigned and fulfilled by the respective partners.

- Work Package Leaders are responsible for overseeing a specific work package
- Task Leaders are responsible for overseeing specific tasks.
- Data processors are the testbed owners of the project that will undertake to collect / process data for the specific purpose of the activity in which it has been collected/generated within the framework of the project.

4. Data security

Data is shared among the ICON project partners via the project sharing and collaboration platform (MS Teams/SharePoint), which adopts industry-standard security tools such as two-factor authentication. The platform is maintained by the Project Coordinator. Additionally, a backup instance of all data will be maintained on the coordinator's account to ensure redundancy and data availability. These tools are used in addition to the specific institutional security tools adopted by each partner.

The data, datasets, documents, and files of each partner organisation must be stored either on their respective drives or the ICON collaboration platform. Each partner is responsible for ensuring proper data backups and secure storage, with access restricted to team members only for any data held locally at their organisation. This guarantees that only individuals involved in the implementation of the ICON project have access to the project and the provided data throughout the project's duration.

In combination, these tools provide comprehensive protection of data stored within the ICON project sharing platform. To protect the privacy of all involved parties, data sharing processes must anonymise all shared data. This means that any partner-related information will be excluded from filenames and metadata, ensuring that sensitive information remains confidential and secure.

4.1. Data ethics

The ICON consortium is committed to conducting data management in accordance with the highest ethical standards and all relevant EU, international, and national laws regarding ethical principles. It is important to highlight that our datasets are collected from testbeds. This ensures that the datasets do not contain any personal information about users. Instead, they focus on performance measurements and network conditions in a controlled environment.

The consortium will continuously monitor for any ethical or privacy concerns and implement a comprehensive strategy to address such issues. Specifically, the project will adhere to the GDPR regulations concerning the protection of individuals in relation to the processing of personal data and the free movement of such data.

Conclusions

The deliverable D1.1: Data Management Plan outlines the activities necessary to ensure the highest quality of project outputs through effective management of project and externally provided data. This document specifies data types, formats, naming, document formatting, and data management practices. It also details the FAIR principles adopted by the ICON project, as well as the data ethics and security measures required to meet the project's objectives.

This Data Management Plan is designed to guide all ICON project partners in handling, storing, and sharing data and datasets to achieve top-quality project outputs. The DMP will be adhered to throughout the project's duration to ensure the successful delivery of deliverables and planned results. It will be updated as needed to create new versions.

References

(GDPR 2018) Art. 4 GDPR – Definitions - General Data Protection Regulation (GDPR), March 2018,
URL: <https://gdpr-info.eu/art-4-gdpr/>