Vertical Innovations in Transport And Logistics over 5G experimentation facilities

D5.5 Data Management Report

Document Summary Information

Start Date	01/01/2021	**Duration**	36 months
Project URL	www.vital5g.eu		
Deliverable	D5.5		
Related Work Package	WP5	**Related Task**	T5.3
Contractual due date	31/10/2023	**Actual submission date**	01/11/2023
Type	ORDP	**Dissemination Level**	Public
Deliverable Editor	Ronan Frizzell (ICP)		

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No 101016567

Contributors and Peer Reviewers

Contributors
All VITAL-5G Partners
Peer Reviewers
Georgios Tsiouris (OTE), Ghazaleh Kia (SF)

Revision history (including peer reviewing & quality control)

Version	Issue Date	% Complete	Changes	Contributor(s)
V0.0	10/01/2023	0%	Initial deliverable material based on D5.4	All VITAL-5G partners
V0.1	04/09/2023	50%	Revision of deliverable text for Sections 1, 2 and 3, and update data collector. QA review at delivery minus 2 months.	ICP, EBOS
V0.2	27/09/2023	80%	Updated deliverable based on QA feedback and update of text for Sections 1-4	ICP
V0.3	02/10/2023	95%	Update of text for Sections 5 and 6, incorporate updates from data collector from project partners, and finalised text for beginning of peer review.	All partners
V0.4	27/10/23	97%	Incorporated feedback from peer review and finalised text for final QA	ICP, SF, OTE
V0.5	30/10/23	99%	Incorporated feedback from final QA	ICP, EBOS
V1.0	31/10/2023	100%	Final coordinator review	WINGS

Disclaimer

Copyright message

Table of Contents

List of Figures

List of Tables

Glossary of terms and abbreviations used

Abbreviation / Term	Description
AES	Advanced Encryption Standard
AGV	Automated Guided Vehicle
AI	Artificial Intelligence
API	Application Programming Interface
AWS	Amazon Web Services
BSS	Business Support System
CA	Consortium Agreement
CSSR	Call Setup Success Rate
DMP	Data Management Plan
DPO	Data Protection Officer
EC	European Commission
ETA	Estimated Time of Arrival
EU	European Union
FAIR	Findable, Accessible, Interoperable and Reusable
GA	Grant Agreement
GDPR	General Data Protection Regulation
GNSS	Global navigation satellite systems
GR	Greece/Greek
GUI	Graphical User Interface
HW	Hardware
IEC	International Electrotechnical Commission
IP	Intellectual Property
IPR	Intellectual Property Rights
KPI	Key Performance Index/Indicator
LEM	Legal and Ethics Manager
ML	Machine Learning
NDA	Non Disclosure Agreement
NSO	Navigation Speed Optimizer
OA	Open Access
ORD	Open Research Data
ORDP	Open Research Data Pilot
OSS	Operational Support System
RAN	Radio Access Network
RAT	Radio Access Technology
RTT	Round-Trip Time
SR	Success Rate
SW	Software
UC	Use case
WMS	Warehouse Management System
WP	Work Package

Executive Summary

This deliverable is the Data Management Report for the VITAL-5G project, which is the updated version of the Data Management Plan (DMP) that was submitted at month 12 of the project. The document provides guidelines and procedures for the VITAL-5G partners on how to manage the research data that is created, collected, and/or used within the project. The procedures outline how data should be collected, stored, and protected during the project, and how data should be handled following the conclusion of the project. The general principles of the procedures outlined in this document are intended to ensure the data is Findable, Accessible, Interoperable and Reusable, i.e., aligned with the FAIR data management principles [1].

The document supports VITAL-5G partners in contributing to the European Commission's Open Research Data Pilot (ORD pilot), the goal of which is explained as follows [1]:

"The ORD pilot aims to improve and maximise access to and re-use of research data generated by Horizon 2020 projects and takes into account the need to balance openness and protection of scientific information, commercialisation and Intellectual Property Rights (IPR), privacy concerns, security as well as data management and preservation questions."

To support partners in effective data management, this deliverable describes 1) the data lifecycle, 2) outlines the FAIR data principles, 3) includes details of the repositories in which data is stored and made accessible, 4) describes the roles of the different actors in the project, 5) the data management guidelines, and 6) outlines specific details relating to the data that will be generated or used within the project.

One of the key outcomes of this document is related to the information that partners must be aware of before sharing data publicly. Such sharing of information relates to openly releasing any research dataset, publication, or deliverable, especially those that may contain personal data. It is the goal of the project to release as much of this data as possible to support the wider scientific community and the project's interactions with third party experimenters, however, this applies only to non-sensitive data.

The issues considered cover GDPR regulations, and related legal and ethics requirements, IP management considerations, confidentiality and the processing and sharing of information from the third party experiments. The main action for all partners when considering the release of project datasets is to first assess the datasets to identify and manage any sensitive data following the guidelines outlined in this document.

Such guidelines include:

1. Ensuring all handling of personal data within the project complies with GDPR requirements, particularly around anonymisation of personal data.
2. Assessing all datasets and related publications for potential impact on IP protection activities and ensuring IP protection is given priority over any publication using approaches such as removing detailed technical information from publications or expediting IP protection activities.
3. Due to the collaborative nature of the project, partners must consider the potential that some data collected solely or jointly with other partners might be considered sensitive (for example, commercially sensitive). All data with a potential overlap with other partner activities must be discussed with those partners to ensure no issues will arise should the data be shared publicly.
4. All provisions related to the handling of sensitive data outlined in this document must be applied to the interactions with third parties. Where necessary, non-disclosure agreements (NDAs), should be signed with third parties to ensure all data sensitivities are managed appropriately.

This document presents the final version of the data management plan for VITAL-5G, which has been revised and updated as necessary compared to the initial version of the DMP (D5.4 [2], released in M12). The related activities involved all partners reassessing the datasets being generated / used in the project, ensuring they correctly captured updates that may have occurred in delivering the project objectives. Data management has also been reassessed at the project-level to ensure the processes outlined here are appropriate for the project, particularly

in terms of how data will be handled following the end of the project. A section has also been added on security considerations and one dedicated to data management in third party experiments.

Due to the advanced stage of the project, one area that was focused on in this deliverable was the mechanisms for ensuring data was available to the public after the end date of the project. To facilitate this open data sharing, which supports the project's exploitation plans and engagements with third parties both for the reminder of the project and after the project end, three main repositories have been created for public access, namely:

1. The VITAL-5G Zenodo Community repository, found at: https://zenodo.org/communities/vital-5g/. This contains the project's open access publications and public datasets. The benefits of this repository are that the data is persistent and will allow the wider community to freely access and reference VITAL-5G data for years to come.
2. The Gitlab repository, which hosts the open VITAL-5G software assets, enabling their download and reuse. This can be found at: https://gitlab.com/vital-5g/
3. The project website, https://www.vital5g.eu/, which also hosts publicly accessible data, such as project deliverables, newsletters, press-releases, use case descriptions, video clips, etc.

Overall, this document provides a comprehensive overview of the datasets being managed within the project, allows partners to understand which datasets can be made available publicly for third parties and which must be kept confidential, while providing guidance and support on how to make appropriate data available.

1 Introduction

The objective of this deliverable is to describe the final data management report for the VITAL-5G project, which supports the project partners in understanding and applying best practice processes to effectively handle project data. It is the goal of VITAL-5G to make project data openly accessible to communities outside the project, provided datasets are not commercially sensitive or otherwise restricted. This plan therefore supports VITAL-5G partners in contributing to the European Commission's Open Research Data Pilot (ORD pilot) [1], as described in the Executive Summary. The impact of this is to ensure the research conducted in the project can be built on by external parties in the spirit of continuous progress of wider research efforts.

The deliverable explains the different phases of data management, how the procedures align with the FAIR data principles (i.e., data should be Findable, Accessible, Interoperable and Reusable), and outlines the data management responsibilities of the different project actors. Of particular interest is the requirement for partners to comply with the obligations set out by the General Data Protection Regulations (GDPR) [3], which is discussed in Section 3.1. Specific information on the data generated, collected, and/or used within the project is also reported in this document. This has been gathered by polling the partners through data collection questionnaires regarding their activities in the project and the associated project data they will handle.

In order to facilitate the data collection described above and the communication of the data handling principles throughout the consortium, a point of contact was identified to act as the main data management representative within each of the partner organisations.

This document presents the final version of data management plan for VITAL-5G, which has been revised and updated as necessary compared to the initial version of the DMP (D5.4 [2], released in M12). In this version, partners reassessed the datasets being generated / used in the project, ensuring they correctly captured updates that may have occurred in delivering the project objectives. Data management has also been reassessed at the project-level to ensure the processes outlined here are appropriate for the particulars of the project, particularly in terms of how data will be handled following the end of the project. A section has been added on security considerations and one dedicated to data management in third party experiments.

In addition, due to the advanced stage of the project, one area that was focused on in this deliverable was the mechanisms for ensuring data was available to the public after the end date of the project. To facilitate this open data sharing, information and requirements related to sharing of data via the VITAL-5G Zenodo repository was updated. In addition, information is provided on the project's Gitlab repository, which hosts the VITAL-5G openly available software assets, enabling their download and reuse. Further details can be found in Section 3.3.

1.1 Mapping VITAL-5G Outputs

Purpose of this section is to map VITAL-5G's Grant Agreement commitments, both within the formal Deliverable and Task description, against the project's respective outputs and work performed.

Table 1: Adherence to VITAL-5G's GA Deliverable & Tasks Descriptions.

GA Component Title	GA Component Outline	Respective Document Sections(s)	Justification
TASKS			
T5.3	The project's DMP will address information related to all types of data that the project will generate, collect and use.	Section 3 Section 5	Section 3 covers the main considerations regarding data management in VITAL-5G, which support partners in making as

	It will also address the standards that will be used to represent the data during the project and how partners may exploit the data resulting from the project.		much of the research data as possible openly accessible outside the project. Section 5 provides details of the data to be gathered, collected and used based on assessments made by project partners.
	Data controllers and data processors will also be clearly named and documented, with the associated legislative duties of care clearly enumerated.	Section 4	Section 4 explains the approach taken to identifying data management personnel in each partner organisation, while Section 4 also outlines the responsibilities of different project actors.
	The DMP will emphasize robust data management procedures that will protect personal data collected as a result of project activities from unauthorized use or sharing, also supporting GDPR compliance as well as FAIR data principles.	Section 3	This section covers processes to ensure FAIR data principles are adhered to and information on GDPR compliance considerations is also covered.
DELIVERABLE			
D5.5 Data Management Report The deliverable contains the final DMP and formal project compliance affirmation report.			

1.2 Deliverable Overview and Report Structure

The structure of this deliverable is outlined as follows:

Section 1 is the report introduction, which outlines the deliverable's purpose and structure.

Section 2 describes high-level general principles relating to how data is handled at the different stages of its lifecycle, including stages that focus on publishing, archiving and re-use of VITAL-5G's data. This section is intended to provide introductory material to frame the more detailed information later in the document.

Section 3 describes the main data management considerations and processes that must be adhered to by project partners. It discusses key issues such as dataset sensitivities (i.e., privacy, IP, confidentiality), which are essential considerations when assessing how to gather, use or share the project's datasets. This section describes how data should be handled to ensure the FAIR data principles mentioned earlier are applied within the project, and partners understand how data can be made Findable, Accessible, Interoperable and Reusable (i.e., FAIR).

Section 4 is intended to identify roles within the project and the responsibilities of those actors in relation to data management.

Section 5 provides a detailed breakdown of data generated/collected/used within the project from the perspective of each of the partners involved. The information reported here comes from direct questionnaire feedback from partner organisations. The design of these questionnaires followed guidelines from the EC [1] on what data should be reported by the partners and covers items such as the purpose of the data, the data type, data origin, etc.

Section 6 contains the report conclusions.

2 Data Lifecycle

This section outlines the typical data lifecycle in the VITAL-5G project, each stage of which is elaborated on and key considerations for each stage are highlighted. Figure 1 gives an overview of the different stages involved in the data lifecycle and the order in which they occur.

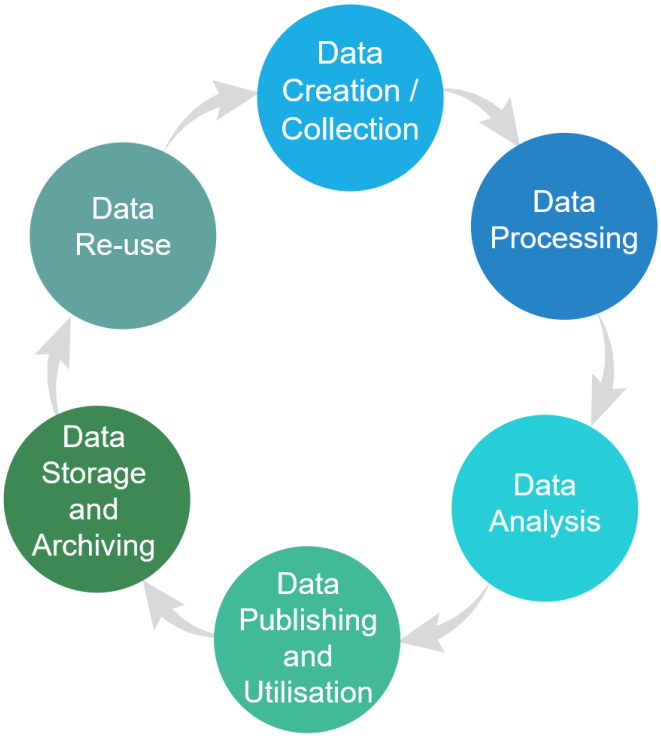

Figure 1: Data Lifecycle in VITAL-5G.

2.1 Data Creation and Collection

This is the initial point at which data is created or collected in the project and relates directly to the project's activities in the various work packages. This step must also consider the format in which data needs to be represented based on how the data will be later processed and re-used. Key considerations for this stage are therefore: recording method for data creation/collection, the format type (e.g., .xml, xls., .pdf, .docx, with a bias towards using standardised formats), data availability (location must be known and shareable, e.g., through links), and data consistency, all of which enhance the data's fitness for processing and re-use. Another key consideration here relates to research integrity and the need to inform people whose data is being collected about the scope of the data collection exercise, ensuring their consent is obtained. At all stages, robust security processes and tools (e.g., databases, communication channels) must be in place to protect the data collected.

2.2 Data Processing and Analysis

This stage covers the data processing by the project partners. In order to facilitate efficient and consistent processing, information about the data's structure and content need to be available for validation by the processing partner(s). All processing methods need to be recorded, processing tools identified, any additional data sources involved in the processing steps need to be correctly referenced and resulting output suitably and securely stored for future verification and re-use.

2.3 Data Publication and Utilisation

As discussed above, it is the intention of VITAL-5G to widely disseminate the results of the project through scientific publications and data sharing repositories, provided the data is not commercially or otherwise sensitive (see Section 3.1 for more details). This stage of the data lifecycle is therefore important from the dissemination point of view and partners are requested to consider making data available for public dissemination. This stage also includes making the data available for internal use within the consortium and is therefore the stage to notify other partners that the data is available and to ensure other interested partners have suitably controlled access to the data. Such access control is intended to ensure data security is maintained and partners sharing data must ensure that receiving partners have access permission from all data owners.

2.4 Data Storage, Archiving and Reuse

Data storage is essential for longer-term data access and subsequent reuse. A number of considerations need to be taken into account during this stage of the lifecycle, namely: 1) that the data version is the latest, 2) appropriate metadata is stored for identifying the data and understanding its origin, format, purpose, etc., 3) access control is set up and maintained for long-term data storage, and 4) expiry dates for the data should be defined, after which the data can be deleted. It is expected that data will be stored for 3 years after the end of the project, unless there are specific operational or security reasons to change this duration. Data lifetime will be discussed in more detail for specific project datasets in Section 5. Specifics regarding the repositories used in VITAL-5G are presented in Section 3.3.

3 Data Management in VITAL-5G

This section provides the main data management requirements for VITAL-5G that must be adhered to by the project partners when considering how to manage research data within the project. This applies to all data that is generated, collected, and/or used in the project. It is expected that the main types of data this plan refers to will centre on scientific and other types of publications, project deliverables, and other research data sets handled during the project. In all cases, the key considerations, which are discussed in this section, regarding intellectual property rights, legal considerations and ethics, and data archiving requirements must be adhered to.

A number of considerations were taken into account in developing this data management plan, as discussed in the following subsections. The basis of the plan was developed primarily to adhere to the FAIR data management principles [1]. These guidelines ensure that data is handled in such a way that datasets are Findable, Accessible, Interoperable and Reusable (FAIR). VITAL-5G's goal in adhering to these principles is to ensure the data is re-used as much as possible both within and outside of the project, in order to accelerate scientific research efforts and avoid duplication of effort within the scientific community or among other stakeholders interested in the project results.

According to VITAL-5G's grant agreement, examples of research data include:

1. Aggregated network data per flow/service slice
   a. provided to vertical industries, together with the relevant measured KPIs
   b. provided to orchestration systems to enable proper network functioning and KPIs matching
   c. network data gathered at the field trials
   d. network data gathered at the interworking interfaces of different site facilities
2. Specific flow/service statistics collected for debugging and/or failure management purposes
3. Detailed data traces collected for specific purposes (e.g., IoT sensor data)
4. Software source code, technical reports and manuals
5. Scientific publications
6. Project deliverables

More specific information on the details of the research datasets handled within the project will be presented in Section 5.

The following sections provide details on the key requirements and processes for VITAL-5G partners to consider when managing data as part of the project.

3.1 Dataset Sensitivities

The intention of this section is to provide partners with important information that must be considered before openly releasing any research dataset, publication, deliverable, and/or collecting/processing of personal data. It is essential that these considerations are accounted for in all data management activities within VITAL-5G. Provided none of the issues below apply, partners are encouraged, as per the project's Grant Agreement (Article 28 and 29), to actively disseminate and exploit project data. The subsections below provide guidelines on actions to take in situations where data sensitivities are identified.

3.1.1 GDPR and Related Legal and Ethics Requirements

The legal and ethics requirements of VITAL-5G are reported on in D8.1: H - Requirement No. 1 [4], which is part of WP8. Deliverable 8.1 covers the legal and ethical considerations of all activities conducted in carrying out the project goals and includes legal and ethical guidelines for data collection and usage in the project. In support of these guideline, the present deliverable (D5.5) outlines the processes and procedures to ensure robust data management within the project.

The main issues to be raised from a legal and ethical point of view in relation to data management within VITAL-5G relate to the collection and use of personal data. Personal data is defined as "[…] any information which are related to an identified or identifiable natural person". Personal data subjects "are identifiable if they can be directly or indirectly identified, especially by reference to an identifier such as a name, an identification number, location data, an online identifier or one of several special characteristics, which expresses the physical, physiological, genetic, mental, commercial, cultural or social identity of these natural persons." [5]. All handling of personal data within the project must comply with:

- The policy and legal requirements of the EU General Data Protection Regulation: EU 2016/679 (the "GDPR"), as in effect since 25 May 2018
- All other applicable national and EU regulations and guidelines on personal data processing
- Applicable regulations and best practices with regards to research projects within the EU H2020 Research Programme (e.g., The EU Cybersecurity Act, The EU Regulation on Platform-to-Business Relations)

Of particular note in this list is GDPR, which is the European Union's privacy and security law that applies to all those processing the personal data of EU citizens or residents, or those offering goods or services to these people. The regulations set out obligations for organisations to comply with in order to ensure data is sufficiently protected.

To support partners in complying with the regulations above, particularly GDPR, D8.1: H - Requirement No. 1 [4] provides procedures and guidelines to be followed when dealing with personal data. Sections 4 and 5 of D8.1 are particularly important to understand the requirements to comply with GDPR, as these sections cover:

- Gathering personal information in VITAL-5G
- Gathering personal information via cameras and sensors in VITAL-5G
- Processing personal information in VITAL-5G
- Handling personal information after VITAL-5G
- Personal information in communication
- Personal information in workshops and events
- Personal information in collaboration, interviews and communities
- Procedure related to the recruitment of participants
- Procedures to handle GDPR related compliance, including actions to take in the event of a data breach

Deliverable D8.1 also contains template consent forms that need to be filled out by participants in events and experiments related to VITAL-5G, which ensure the transparency principle of GDPR is adhered to.

It should be noted that the project has two important roles in place to support partners in compliance with all legal and ethical requirements, namely the Legal and Ethics Manager (LEM) and the Data Protection Officer (DPO). All questions relating to legal and ethics can be directed to the project's LEM or DPO, as identified in the Grant Agreement.

All research data containing personal information must be anonymised before data is shared between VITAL-5G partners and before it is disseminated to actors outside the consortium, including 3rd party experimenters.

3.1.2 IP Management Considerations

Due to the potential for commercial value associated with the protection of intellectual property (IP), all datasets and related publications must be assessed for potential impact on IP protection activities, particularly patent filings. Advice can be sought by partners from the Project Management Board and the project's innovation manager (ICP). In situations where sharing datasets has the potential to negatively impact IP protection efforts, the data will not be shared openly but will be archived in the project's data repository. If embargoes have to be placed on data, a timeline for removing those embargoes will be discussed with affected partners. It is noted

that once a patent is granted the information detailing the invention will be made public through the standard process of patent publication.

For reference, all terms and conditions concerning IP protection, ownership, foreground, background, and accessibility rights are outlined in the project's Consortium Agreement (CA).

3.1.3 Confidentiality

Before sharing data openly, partners must consider the potential that some data collected solely or jointly with other partners might be considered sensitive for a number of reasons (for example, commercially sensitive). All data with a potential overlap with other partner activities must be discussed with those partners to ensure no issues will arise should the data be shared publicly. The obligations concerning confidentiality in VITAL-5G have been outlined in the project's Consortium Agreement (CA), with specific guidance regarding confidentiality in dissemination activities given in Section 8.4.2.3 of the CA.

3.1.4 Data in Third Party Experiments

VITAL-5G has a particular focus on third party engagement and therefore necessitates the use, generation, and/or processing of data owned by third parties. As these partners are not signatories of the project's Consortium or Grant agreements, certain sensitivities could arise through these interactions. All provisions related to the handling of sensitive data outlined in this document must be applied to the interactions with third parties. Where necessary, non-disclosure agreements (NDAs), should be signed with third parties to ensure all data sensitivities are managed appropriately.

The project has made available a number of datasets for third parties to use in their experiments. Details of these datasets are available on the project website (https://www.vital5g.eu/). In addition, detailed reporting on the third party trials and the datasets used by the third parties will be provided as part of the WP4 deliverables, mainly D4.3: "VITAL-5G trial results analysis & post-processed evaluation", which is due to be released at the end of the project.

3.2 Security Considerations

In order to assure security is maintained and sensitive data is protected, partners must implement industry-standard security measures. It is the responsibility of each organisation to follow established internal processes for maintaining security of their internal IT systems. In particular, the following considerations must be accounted for to minimise any unauthorised disclosure of IP, and are required mainly due to the collaborative nature of the VITAL-5G project:

- As mentioned above, all research data containing personal information must be anonymised before data is shared between VITAL-5G partners and before it is disseminated to actors outside the consortium.
- Shared sensitive data, as discussed in Section 3.1, must be encrypted and access controls must be implemented to limit access only to authorised personnel from all involved organisations.
- Access controls to sensitive IP must be implemented, ensuring only authorised personnel have access.
- If authorised to share data external to the consortium, this must be carried out in compliance with the provisions discussed in Section 3.1.3. Of particular note is that sharing data with external parties must be governed by suitable non-disclosure agreements.
- Security measures to protect confidential / sensitive data will be maintained for at least 5 years post-project.
- Project-related data stored in partner-controlled repositories must be sufficiently protected against data loss. Routine back up of data must be implemented as part of each partner's standard data management processes to minimize the risk of data loss.

- Further relevant information can be found in Section 3.3, particularly related to security aspects of the repository used in VITAL-5G for sharing of project data.

3.3 Repositories in VITAL-5G

Digital data repositories are critical infrastructure that enable effective data management in VITAL-5G. To facilitate sharing of information between partners internal to the consortium and to parties outside of the project, two types of repositories are discussed in this section.

3.3.1 Internal Data Sharing

Archiving is essential for protecting the project's data and ensuring effective records are kept of project activities, outcomes, etc. VITAL-5G uses Teamwork [6] as the project's repository for information and file sharing during the project (including all deliverables, meeting minutes, meeting slides, working documents, etc.). Access to the project's Teamwork files is restricted to project partners, and access is controlled by EBOS, who are acting as repository administrators. Each user is assigned a unique account, which is password protected. All successful and failed login events are recorded by the Teamwork system. Certain restrictions to directory creation and manipulation have been applied to preserve the data integrity and data structure of the repository. In alignment with the project's Grant Agreement (GA), project data stored on Teamwork will be maintained for at least 2 years after the conclusion of the project.

The repository is hosted with Amazon Web Services (AWS) and the associated infrastructure is in Amazon-controlled data centres, which have strict access control measures in place. These data centres operate dedicated physical security programs (such as protecting against seismic and fire events) and carry multiple certifications, such as such as ISO 27017 for cloud security [7]. Data on Teamwork is protected with AES-256 encryption algorithms and continuously backed up to physically distinct servers. The system is constantly monitored by internal and external monitoring tools to detect any issues, including network vulnerabilities. Teamwork has achieved the internationally recognised ISO/IEC 27001:2013 standard for its information security systems and processes [8].

3.3.2 External Data Sharing

As discussed above, it is the intention of VITAL-5G to make all publications and as many research datasets as possible openly accessible to the public to comply with the FAIR data management principles. In order to do so, suitable datasets (as will be discussed in Section 3.6) will be shared through open data repositories, such as Zenodo [9] and others (where appropriate), which are part of the OpenAIRE project [10] and enable the Open Access (OA) data management policy of the EC [11]. The benefits of these types of repositories are that they provide version control, ability to add metadata, and provide digital object identifiers (DOIs) for persistent referencing of all content that is available in the repository.

The three main publicly-accessible repositories used by the project are:

4. The VITAL-5G Zenodo Community Repository, found at https://zenodo.org/communities/vital-5g
5. The Gitlab repository, which hosts the openly available VITAL-5G software assets, enabling their download and reuse. This can be found at: https://gitlab.com/vital-5g/
6. The project website, https://www.vital5g.eu/, which also hosts publicly accessible data, such as project deliverables, newsletters, press-releases, use case descriptions, video clips, etc.

3.4 Findable Data

Making data findable is essential to enhance the reuse of research datasets, both internal to the project and when sharing data publicly. This sub-section outlines guidelines which partners can apply to their datasets to ensure that data can be found efficiently and that those searching for data can quickly understand whether the data being accessed can meet their needs.

The VITAL-5G deliverables and meeting minutes are some of the key sources of data generated in the project that is used for information sharing. The project's quality assurance handbook (D7.4 Project Quality Handbook [12]) contains templates and guidelines for how these types of project documents are to be created, named and filed. This information is important to ensure data is findable, particularly in relation to the naming convention outlined. D7.4 gives further information on the naming convention of files, a selection of which is shown below in Table 2. A clear reference to document versions is included in the file names, particularly for the more complex documents, such as the deliverables and meeting minutes, where numerous partners often contribute.

The Teamwork repository (Section 3.3.1) is an important tool that ensures VITAL-5G data is findable, as it is the central location in which all key project data is stored. It is therefore a single location where project actors can access shared material. An intuitive structure of directories has been created on this site, based on the project activities and data storage/sharing needs, including directories dedicated to each of the project's work packages. The repository includes numerous searching and metadata provisioning mechanisms, including:

- document version number identification
- tagging
- texted-based file descriptions
- advanced searching capabilities
- creation of direct links to files

The combination of the repository capabilities and the standardised naming convention ensures the project data can be readily found and re-used within the project. For externally shared datasets, publications, deliverables, etc., it will be important to adhere to similarly clear naming conventions when uploading files to facilitate finding data. In addition, many of the openly accessible databases, as will be discussed in Section 3.6, have features to include metadata to better describe the file and its contents, which aids data discoverability. All VITAL-5G members sharing data in this way must ensure metadata and appropriate keywords are associated with all files shared, thus facilitating the findability of the data.

Table 2: VITAL-5G naming convention of selected document types.

Document Type	Naming Convention
Deliverables	ProjectName DeliverableNo DeliverableTitle Date Version Partner
Meeting minutes (online or physical)	ProjectName MeetingID MeetingLocation (or Online) MeetingTitle Minutes MeetingDate Version Meetings identifiers: • KOM=Kick Off Meeting • PM=Project Meeting (concerns regular General Assemblies) • TM=Technical Meeting • RM=Review Meeting • Telco=Online meeting (Skype, GoToMeeting, WebEx, etc.)
Presentations	ProjectName MeetingID PresentationTitle Partner MeetingDate

3.5 Interoperability of Data

Data interoperability is an important consideration for VITAL-5G in order to maximise the impact of the project. Details will be provided later in this document (Section 5) of the data used/generated by partners and will describe the formats used to represent that data. Standard formats will be used wherever possible to maximise interoperability of data across the different organisations involved in this project. Partners will avoid, as much as possible, the use of proprietary formats that would preclude others from using the information.

In situations where data is to be shared, partners will collaborate to ensure data formats are suitable for reuse by the receiving organisation. In the unlikely event that data can only be made available in non-standard formats, sending and receiving partners will collaborate to ensure that the data can be worked with by others in the consortium, ensuring that the structure and type of data is clearly understood and all suitable conversion codes will be created and/or shared to allow effective data-reuse. Where needed, any domain-specific terminology will be clarified to minimise barriers for interoperability, which is recognised as important for inter-disciplinary projects such as VITAL-5G. For data sharing within the consortium, commonly used file types, such as .csv, .xml, .xls(x), .odt, etc, will be used wherever possible.

For any external sharing of datasets, for example as part of the ORD pilot mentioned above, partners will make best efforts to supply data in standard formats and follow the conventions recommended by the repository being used in terms of data representation vocabularies.

3.6 Openly Accessible Data

All VITAL-5G partners are expected to make suitable research data from the project available through repositories listed on OpenAIRE (Open Access Infrastructure for Research in Europe) [10]. All efforts should be made to share available data sets, provided the following criteria are met (as discussed in more detail in Section 3.1):

- The data is not confidential and does not include personal information.
- Permission has been obtained from all relevant stakeholders that the data can be made openly accessible.
- Making the data openly accessible does not jeopardise commercial exploitation or IP protection activities.

The considerations for making the project's research data openly available for:

1. Publications
2. Project deliverables
3. Research Datasets and Assets

are discussed below in more detail to describe the processes partners must follow when releasing VITAL-5G project data openly.

3.6.1 Publications

3.6.1.1 Notification of Upcoming Publications

The process below is to be followed by all project partners to notify the consortium of upcoming publications. The process has been proposed by the project coordinator as a reasonable balance between ensuring data sensitivities are flagged and ensuring the process is realistic for partners to follow. The process has been discussed with partners and has been taken on as an active process across the project, which partners are expected to follow.

VITAL-5G Publication Release Process:

- **3-4 weeks prior to submission:**
  - **Lead author** informs the Project Coordinator (PC), and WP5 and WP6 leaders regarding their intention to write a paper, including a brief description of the paper, project data that may/will be used, the targeted venue, and all partners involved.
- **3 weeks prior to submission:**
  - **WP6 leader** informs the consortium about the intended publication, allowing partners to express interest/concern.
- **1 week prior to submission:**

- **Lead author** must share a mature, complete draft with the project coordinator, WP5, WP6 leaders and any partners that have expressed interest/concerns about the publication.
- **Concerned partners have 1 week to express immediate concerns.**
- **After the submission of the paper**
  - Partners may express concerns/requests for changes, which can be discussed among the interested parties until the day the author is notified of paper acceptance.
- **After paper acceptance and before accepting camera-ready (typeset) version:**
  - **If interested partners** have agreed to changes:
    - Authors apply these changes for the camera-ready version, which will be eventually published.
  - **If interested partners** couldn't reach an agreement:
    - In this case, the paper must be withdrawn.

3.6.1.2 Requirement for Open Access Publications

In line with the project's dissemination strategy, publications are an essential output of the project to ensure the project results reach as wide an audience as possible. This approach to publications is to conform with the open access requirements set out in the project's grant agreement (Article 29). This requires partners to ensure online access to publications is free of charge to all users, assuming none of the sensitivity considerations in Section 3.1 apply.

All VITAL-5G publications must be made open either through self-archiving ("green" Open Access) or open access publishing ("gold" Open Access), dependent on the options available to the authors from the article publishers. This helps ensure research outcome reuse is maximised. VITAL-5G recommends that partners pursue the "green" Open Access channel, where possible. It is, however, the responsibility of the partners contributing to each publication to choose the appropriate channel. The two options are described as follows:

- **Green Open Access:** this is a self-archiving option, where all publications are to be shared via OpenAIRE repositories, such as Zenodo or other openly accessible repository of their choice (as discussed in Section 3.3), in a machine-readable electronic copy (e.g., PDF), at the latest upon publication. All publications following this path must adhere to all terms and conditions of the article publisher and/or other copyright owners.
- **Gold Open Access:** This option is where articles are published in fully open access journals or in journals that allow certain articles to be openly accessible. This option typically requires the payment of a processing charge to enable open accessibility.

Also in accordance with the project's grant agreement, partners must endeavour to make all research data and tools needed to validate the results presented in each publication open. This is to be done at the same time as the publication is made openly accessible, in order to facilitate knowledge sharing and reuse of research outcomes.

Embargoes will apply to any publication or related data that is commercially sensitive or could undermine patent protection efforts. Discussions with relevant stakeholders, the project coordinator, and the project's innovation and commercialisation board will be held to determine the duration of such embargoes. Decisions on the duration of embargoes will be clearly communicated with all parties involved.

3.6.2 Project Deliverables

All VITAL-5G deliverables have been designated as either "Public" or "Confidential", where "Confidential" means that these deliverables will only be made available to consortium members (and the EC for project reporting purposes). All "Public" deliverables will be made freely available on the project website (https://www.vital5g.eu/public-deliverables/), and will be available there for up to 2 years after the completion of the project. Where possible, a publishable version of the document will be produced and shared in all other cases.

3.6.3 Openly Accessible Research Datasets and Assets

This section highlights the requirement for project partners to ensure all *non-confidential* digital research data generated/collected in the project is made available for third parties free of charge via an open access data repository.

As mentioned earlier, these datasets should include all data and tools necessary to validate the results in the project's scientific publications. In addition, all efforts must be made to ensure other sources of *non-confidential* data produced in VITAL-5G are made available in openly accessible repositories, such as source code, Network Application descriptions, field measurements, AI/ML training datasets, design files, etc. It is emphasised, however, that this requirement applies only to non-confidential datasets with no security concerns, and/or those datasets that do not contain personal data. Of particular note is the fact that sharing of datasets is not expected in cases where doing so would jeopardise commercial activities or efforts to protect intellectual property (i.e. this applies only for datasets that do not have sensitivity concerns, as outlined in Section 3.1).

Section 5 contains details of the project partners' assessment of research datasets that will be collected, generated and used within this project. That section provides information on the suitability of the datasets to be shared openly and includes information on the sensitivity of the various datasets.

In the case where data sources are to be shared publicly, particularly source code associated with open source software, the correct licensing arrangements will need to be determined depending on the dataset being released. If possible, VITAL-5G will aim to release datasets under a Creative Commons license, which encourages data reuse because source code/data can be freely worked with, provided the author/original data provider is credited and any new derivative works are released under the same licensing terms. Another option is the Apache V2.0 license[1] which allows software to be distributed on an as-is basis and does not require derivative work of the software, or modifications to the original, to be distributed using the same license. It is recognised that there are certain limitations to these approaches and the exact license to apply will have to be decided on at the point of data release.

Embargoes will apply to any data related to patent protection and/or data that is commercially sensitive. Discussions with relevant stakeholders, the project coordinator, and the innovation manager will be held to determine the duration of such embargoes. Decisions on the duration of embargoes will be clearly communicated with all parties involved.

3.6.3.1 Notification Requirements

In order to avoid any issues associated with release of confidential or otherwise sensitive information, approval must be sought by the partner wishing to release any research dataset from the project coordinator prior to making the data available openly. The coordinator will ensure the dataset have been reviewed by all data owners. This is to verify that legitimate interests regarding the dataset would not be significantly harmed by the release of the information. Provided no issues are identified, approval by email will be sent by the project coordinator to the partner concerned.

All datasets released in this way must be recorded as part of the project's dissemination tracking activities and will contribute to the European Commission's Open Research Data (ORD) pilot.

[1] https://www.apache.org/licenses/LICENSE-2.0

4 Data Management Responsibilities of Project Actors

Following robust data management processes applies across all work packages of VITAL-5G. This section aims to highlight the key data management responsibilities of actors within the project. Those actors are expected to be aware of these responsibilities and implement associated actions to ensure compliance with the guidelines set out in the data management plan. The roles and associated responsibilities are:

- **Data owners:** (i.e., those that take the lead in use, creation and/or collection of data) are responsible for complying with the processes and guidelines presented in the VITAL-5G DMP (D5.4) and this document (D5.5).
- **Work Package Leaders:** are responsible for ensuring the processes and guidelines in this document are adhered to in their respective work packages.
- **Task 5.3 Leader:** (ICP) is responsible for taking the lead to create the data management plan and report in collaboration with the other consortium partners and communicating to the partners on its content.
- The consortium has nominated a **data management main point of contact** for each partner organisation. These people are considered responsible for all actions related to the DMP, particularly in relation to the creation of the DMP document. This role also has the responsibility of ensuring all those working on VITAL-5G in their respective organisations are aware of the content of the DMP, such that the principles outlined in this document are applied while working on all VITAL-5G activities. The list of data management main points of contact is maintained within the project by ICP, to be used internally by the project and is not shared in this public deliverable for privacy reasons.
- **Project DPO/Legal & Ethical Manager:** The role of the project's DPO and LEM (IMEC) will be to provide guidance and oversight on the creation of the DMP and ensure partners are aware of their obligations under EU and national legislation.

5 VITAL-5G Data Summary

The purpose of this section is to provide a summary of the datasets that will be handled by partners in this project. The process followed helps partners assess their role in the project from a data management point of view and identify those datasets that can be publicly shared, those that are restricted, and how data will be stored, accessed, shared, etc., during the project and after its completion. Data will be handled in accordance with the guidelines presented in this document and follow the data protection processes from D8.1: H - Requirement No. 1 [4]. As discussed above, the latter document outlines legal and ethical requirements, including the obligations of partners in the event of a data breach.

5.1 Data Capture Process

Partners were requested to assess their planned work throughout the project in order to identify datasets that will be used and/or generated. The data collector template was used, as shown in Table 3, and filled out by the project partners. Where possible and appropriate, partners were requested to work on filling out the templates in groups (for example, among use case teams) for better visibility of data being considered by different partners and for more direct discussions on confidentiality concerns. The fields of the template were determined following discussions with the project's data protection officer (DPO) and the WP5 partners.

In order to support partners in assessing relevant datasets, a non-exhaustive list of potentially relevant data sets was provided to partners. Completed examples of the data collector were also provided to the partners to illustrate the required level of information expected for each dataset reported. The data collector document sent to partners is shown in Annex I.

Table 3: Data Collector Template for data management plan.

Lead partner	Partner responsible for dataset generation and/or data collection
Data Owner	Owner of data
Description of data	High-level, short description of the data (e.g., sensor data to capture ship location in port area)
Data Type	Summarise type of data e.g., measurement data, video, written reports, KPI results, etc.
Data Purpose	Short description of why the data is necessary and what will be done with the data (e.g., positional data used as input for automated route selection algorithms)
Data Origin	Specify if data is from test, partner, database, simulation, questionnaire, etc.
Which partner(s) will use the data/who has access to the data?	Other partner(s) short name(s), if appropriate, external parties, N/A otherwise
Data Format	File format used to store data (e.g., CSV, matfiles, png, .docx,.xlsx, etc.)
Dissemination level • **Public** • **Confidential** • **Personal**	Does the data contain personal information and need to be treated according to the GDPR? Is the data confidential to the consortium? Is the data confidential because it is commercially sensitive for industry partners (warehouse operators, port authorities)?
FAIR: Will the data be shared in publicly open database?	e.g., to support scientific publication on repositories such as Zenodo, Open Research Europe. If not, a reason must be given (e.g., privacy, export regulations, IP, commercially sensitive)

Data Storage (Location, Duration)	Location: e.g., Teamwork repository, company secure servers, Duration: e.g., Project duration, 5 years after project end, temporary storage on company servers only for calculations/processing
UC (Belgium, Romania, Greece)	Applicable UC

Following a review of the project's Grant Agreement and discussions with the project's DPO, the following list of datasets was provided to partners as examples of datasets to report. This list was intended as a guide for partners to assist in identifying datasets of relevance for data management plan (DMP).

Technical data to consider:

1. Sensor data/other measurement data (e.g., T2.4 "… port monitoring services, the tracking of dangerous goods, on-demand video-surveillance systems, the optimisation of order management")
2. Source code, external open source code, including AI and ML algorithms, other "intelligent applications"
3. UC Requirements
4. UC Platform requirements and specification
5. UC questionnaire feedback
6. UC design files (drawings, calculations/simulations, architecture proposals, equipment specifications, platform specification, etc.)
7. Network Application descriptions
8. UC evaluations; UC user/experimenter feedback + respondent identifying information.
9. Results from demonstrators (KPIs, related statistics, "Network Application operability" T4.2)
10. Literature review/State-of-the-art review (content ends up in deliverable and should be captured by partner documenting data used for related WP)

Partners were not requested to supply information on general project data such as deliverables, dissemination material (publications, posters, presentations, flyers, questionnaires, social media related, etc.), or general project management material (internal/EC reporting, meeting material, etc.), as this type of data will be summarised in the data management plan separately.

5.2 Details of Project Data

This section summarises the project data identified by the project partners. As this is the initial data management plan, the information contained in this section will be revised and updated as necessary as the project progresses. The sub-sections are organised according to use case, one sub-section covers the cross-use case technologies being developed as part of the project, the general project data, and the final section covers datasets needing special treatment due to the potential that it includes personal data.

Project deliverables are the core outcomes of the project and the guidelines on how deliverable documents are to be handled from a data management point of view are discussed in Section 3.6.2. Due to information already available on the project deliverables, these will not be discussed in detail in this section.

5.2.1 Use Case 1

As described in D1.1, Use Case 1: Assisted Vessel Transport will apply 5G technology in a smart port context, and showcase the benefits of 5G for enhanced port efficiency. The partners involved are: IMEC, SF, TEL and DT. For more information on this use case, the reader is referred to D1.1: Report on use case requirements [13].

The following information has been captured by the Use Case 1 partners in their assessment of relevant datasets:

Table 4: Data summary for use case 1 (Belgium) partners – table 1 of 3.

Lead partner	**SF**	**TEL**	**IMEC**	**SF**
Data Owner	SF	TEL	IMEC/SF	IMEC/SF
Description of data	It consists on the list of vessels, their schedules and their origin/destination points in a given time period.	Data about network KPIs, i.e., link quality, actual latency, throughput, resources used in the cloud, etc.,	On board sensor data	Vessel data
Data Type	Navigation and mooring schedules	KPI results	Measurement data (radar, Camera Data)	Measurement data: numerical data coordinates, speed, heading
Data Purpose	This data will be used to feed the Speed optimizer and the port Digital Twin, autonomous navigation	This data will be used to assess the right performance of the 5G network supporting the Network Applications	This data will feed the different autonomous and remote navigation algorithms, as well as the digital twin and speed optimizer	This data will feed the different autonomous and remote navigation algorithms, as well as the digital twin and speed optimizer
Data Origin	This data is gathered from SF	TEL, 5G network	Sensors on SF vessel	Vessel/GNSS sensors SF
Which partner(s) will use the data/who has access to the data?	SF, IMEC, DT	IMEC	IMEC, SF, DT	IMEC, SF, DT
Data Format	CSV, API	CSV	CSV, PNG	CSV, JSON, API
Dissemination level • **Public** • **Confidential** • **Personal**	Planning Data: Non personal, confidential, non public, Sensor Data: could be shared	Non personal, confidential, non public	Non personal, confidential, non public	Non personal, confidential, non public
FAIR: Will the data be shared in publicly open database?	no	no	no	no
Data Storage (Location, Duration)	SF, min 3 years	Stored in TEL infrastructure, min 3 years	Stored in SF / IMEC infrastructure, min 3 years	Store location SF / IMEC infrastructure, min 3 years
UC (Belgium, Romania, Greece)	UC1	UC1	UC1	UC1

Table 5: Data summary for use case 1 (Belgium) partners – table 2 of 3.

Lead partner	SF	DT	SF
Data Owner	SF	VisuRis	Port of Antwerp
Description of data	Vessels' dimensions	A tool to calculate inland navigation routes and retrieve ETAs based on given parameters	Location data of Port point of entry
Data Type	Numerical data	Processed results	Static location data
Data Purpose	This data will feed the different autonomous and remote navigation algorithms, as well as the digital twin and speed optimizer	NSO call the VisuRis API to retrieve a navigation route and ETA estimation for intermediary points	This data is necessary to define virtual areas for point entry areas. NSO uses these coordinates to generate, via geofencing, areas that trigger the NSO algorithm.
Data Origin	SF	Route calculation tool	inquiry
Which partner(s) will use the data/who has access to the data?	IMEC, SF, DT	DT	SF, DT, IMEC
Data Format	CSV	json	TBD
Dissemination level • **Public** • **Confidential** • **Personal**	Non personal, confidential, non public	Non personal, confidential, non public	Non personal, confidential, non public
FAIR: Will the data be shared in publicly open database?	no	no	no
Data Storage (Location, Duration)	Store location Seafar insfrastructure, min 3 years	Temporary storage on company servers only for calculations/processing	SF / IMEC, min 3 years
UC (Belgium, Romania, Greece)	UC1	UC1	UC1

Table 6: Data summary for use case 1 (Belgium) partners – table 3 of 3.

Lead partner	IMEC / SF/ TEL / DT	IMEC / SF/ TEL / DT
Data Owner	IMEC / SF/ TEL / DT	IMEC / SF/ TEL / DT
Description of data	Network Application software architectures	UC global software/network architectures

Data Type	Diagrams	Diagrams
Data Purpose	Describe the design and operation of Network Applications	Describe the design and operation of the whole UC
Data Origin	IMEC / SF/ TEL / DT	IMEC / SF/ TEL / DT
Which partner(s) will use the data/who has access to the data?	IMEC / SF/ TEL / DT	IMEC / SF/ TEL / DT
Data Format	PNG	PNG
Dissemination level • **Public** • **Confidential** • **Personal**	Non personal, Confidential	Non personal, Confidential
FAIR: Will the data be shared in publicly open database?	No	No
Data Storage (Location, Duration)	Teamwork repository, project duration	Teamwork repository, project duration
UC (Belgium, Romania, Greece)	Belgium	Belgium

It is noted that DigiTrans (DT) will act as a data user rather than a data owner. This was discussed in detail in a use case partner meeting on the subject. DT have confirmed that their role in data management within the project is well represented in the tables in this section.

5.2.2 Use Case 2

As described in D1.1, Use Case 2: 5G Connectivity and Data-Enabled Assisted Navigation Using IoT Sensing and Video Cameras will demonstrate the application of 5G technology in a smart port context by implementing data-assisted navigation applications using IoT sensing and video cameras installed in Galati port and on ships and barges. The partners involved are: BEIA, NAVROM, ORO, ATG and INCE. For more information on this use case, the reader is referred to D1.1: Report on use case requirements [13].

The following information has been captured by the Use Case 2 partners in their assessment of relevant datasets:

Table 7: Data summary for use case 2 (Romania) partners – table 1 of 2.

Lead partner	**INCE**	**BEIA**	**ORO**
Data Owner	INCE (for data produced in VA2)	NAVROM	ORO
Description of data	Spatio-temporally fused incoming sensor data and ML-related results.	Sensor data about air (temperature, pressure, humidity, O2-level, CO2-level), water (depth, temperature, PH level), location of ship and video surroundings	Sensor data (for QoS purposes) and KPIs from 5G infrastructure testbed (in RAN and Core components).

Data Type	Fused data: • Vessel data ○ GPS (latitude, longitude) ○ depth ○ Air quality (Temperature, pressure, humidity, luminosity, O_2, CO) • Port data: ○ depth ○ water quality (temperature, pH) ML-results • Trained model artifacts including related data, hyperparameters used, KPI performance results. Trajectory or other forecasting results displayed in UI or acquired through APIs.	Measurement data from sensors (time series), GPS, and video feeds	Data related to connected mobile or fixed devices (like video cameras) or other IoT devices. KPIs examples: accessibility - 4G CSSR, 5G Addition\ Change\ Modification SR, Traffic density, No. of users, Packet Loss, RTT, throughput, RAT data volume and session time, TCP Retransmissions, Traffic per device
Data Purpose	Its purpose is through UC2 activities: 1) to exemplify a case of data fusion from multiple sensors concurrently placed in various points in space 2) to present a working prototype that allows a user-configurable automated training of ML models and their use for forecasting (in this case related mostly on vessel trajectories). 3) to use for post-analysis and reporting.	Input for ML/AI algorithm and real-time visualization of data for ship crew.	Quantify project results
Data Origin	Sensor data acquired through tests/trials, ML-related data are produced ad hoc during trial or post- analysis	Data from UC measurements	Testbed (RAN – signalling probes, Core Modules, OSS/BSS, API)
Which partner(s) will use the data/who has access to the data?	INCE for post analysis and reporting. No other partner or 3rd parties. Availability to consortium members to be determined upon request.	UC2 partners	NAVROM, BEIA, INCE.
Data Format	Stored in RO site VMs in a Hadoop Distributed Filesystem (HDFS). Temporary backup in INCE premises.	csv, mp4	.xls/.xlsx, .doc/.docx, .pdf, .csv.

	Results may be exported ad hoc in CSV for post-analysis/reporting.		
Dissemination level • **Public** • **Confidential** • **Personal**	No personal information. Availability of ML-related results to consortium members to be considered upon request. **Confidential** (as we are not the owners of the original dataset). Only post analysis results available in public deliverables and publications are considered Public.	Confidential: it is commercially sensitive for NAVROM.	Public + Confidential
FAIR: Will the data be shared in publicly open database?	No, since: 1) Resulting data are under patent submission 2) We are not the owners of the actual sensor data, that may be commercially sensitive to other UC partners.	No. Data is commercially sensitive	If part of publication
Data Storage (Location, Duration)	Fused sensor data: Stored in RO site VMs in a Hadoop Distributed Filesystem (HDFS). Temporary storage in INCE company secure server for post-processing up to 1 year after project duration. Results/ reporting post analysis data: Stored at INCE premises indefinitely.	Data will be stored in teamwork repository for 5 years after project end	Data is stored on Local storage (ORO is owner of the infrastructure). Duration: Temporary storage on company servers only for calculations/processing.
UC (Belgium, Romania, Greece)	RO	Romania	Romania

Table 8: Data summary for use case 2 (Romania) partners – table 2 of 2

Lead partner	**ORO**	**BEIA**	**BEIA**
Data Owner	BEIA, ORANGE, INCE	NAVROM, ATG, ORANGE, BEIA, INCE	BEIA, ORANGE, INCE
Description of data	KPIs from testbed	UC2 requirements and design	Network Applications descriptions and source code
Data Type	KPIs	Reports about UC2	Source code
Data Purpose	Quantify project results; to be used by VITAL-5G Platform	Quantify project results	Support the development of Network Applications

	Components where applicable	Support Network Applications development	and the project outcome
Data Origin	Testbed	UC2 partners	UC2 Network Application developers
Which partner(s) will use the data/who has access to the data?	UC2 partners	UC2 partners	UC2 partners
Data Format	.doc	.doc, .docx, .png	plain text files (e.g., .h, .c, .java, etc.)
Dissemination level • **Public** • **Confidential** **Personal**	Public + Confidential	Public + Confidential	Public + Confidential
FAIR: Will the data be shared in publicly open database?	Public data: yes	Public data: yes	Public data: yes
Data Storage (Location, Duration)	Local storage/reported in deliverables; duration of project	Local storage/reported in deliverables; duration of project	Gitlab (https://gitlab.com/vital-5g/vital-5g_netapps) and project internal code repository
UC (Belgium, Romania, Greece)	Romania	Romania	Romania

5.2.3 Use Case 3

As described in D1.1, Use Case 3: Automation & remote operation of freight logistics (Warehouse logistics) intends to demonstrate the feasibility of applying 5G technology to optimise complex logistics operations through the use of state-of-the-art Automated Guided Vehicles (AGVs). The partners involved are: WINGS, OTE, DIA. For more information on this use case, the reader is referred to D1.1: Report on use case requirements [13].

The following information has been captured by the Use Case 3 partners in their assessment of relevant datasets:

Table 9: Data summary for use case 3 (Greece) partners – table 1 of 4

Lead partner	**WINGS**	**WINGS**	**WINGS**	**WINGS**	**WINGS**	**WINGS**
Data Owner	WINGS	WINGS	WINGS	WINGS	WINGS	WINGS
Description of data	Data from the AGV and its onboard sensors	Data from sensors distributed around the warehouse.	Data from endpoint devices (e.g., smartphones, handhelds) carried by workers	Application and Network Application requirements / specifications and Architecture	Solution detailed specs and AI/ML algorithm design and source code	Trial results KPI collection, simulations output

Data Type	AGV location, lidar readings, distance to obstacles, scanner readings, etc.	camera feed, temperature, humidity, RFID signals	Worker location, input/output to/from GUI (AGV instructions), WMS data (picking)	Requirements documents, architecture figures	Design documents, architecture figures, source code	Reporting document, KPI results
Data Purpose	Data used for the guidance of the AGV around the warehouse	Camera feed used for the identification of humans/obstacles. Temperature/humidity data used for the monitoring of the product condition. RFID data for identification of products/persons.	Used to instantiate the "follow-me" function, and as a means of interaction between the worker and the AGV. WMS data for identify effort/volume.	To be used for the implementation of the SW and HW and their integration	Implementation of the GR solution	Evaluation of trial results and validation of the GR UC
Data Origin	Live test	Live test	Live test	Partner created	Partner created	Live test, simulation
Which partner(s) will use the data/who has access to the data?	WINGS, DIA	WINGS, DIA	WINGS, DIA	WINGS	WINGS	WINGS
Data Format	.csv	.mp4, .csv	.csv, json	.docx, .jpg	.docx, .jpg, Python, Java, C, Agular files	.docx, .xlsx, Java / Python files
Dissemination level • **Public** • **Confidential** • **Personal**	Public	Personal	Personal	Public	Confidential	Public
FAIR: Will the data be shared in publicly open database?	Yes	No	No	Yes	No (commercially sensitive data)	Yes
Data Storage (Location, Duration)	Company secure server, Duration of the project Zenodo after project end here	Company secure server, Duration of the project	Company secure server, Duration of the project	Teamwork Available after project end Zenodo after project end here	Company secure server, Duration of the project	Teamwork Available after project end Zenodo after project end here

UC (Belgium, Romania, Greece)	Greece	Greece	Greece	Greece	Greece	Greece

Table 10: Data summary for use case 3 (Greece) partners – table 2 of 4.

Lead partner	**DIA**	**DIA**	**DIA**
Data Owner	DIA	DIA	DIA
Description of data	Warehouse management system data (storage locations, incoming orders, etc.) & order historical and real time data	Worker information (schedule, orders completed, time worked, etc.)	UC requirements and service level KPI definition
Data Type	Order manifest, including storage location of each pallet, customer etc.	worker schedule, KPI results	reporting document
Data Purpose	To provide the AGV with the appropriate destination of each pallet	To be used for resource optimisation, and benchmarking of trial results	Used for UC definition and solution specification and validation
Data Origin	Partner database, dummy data	Database, Live test	Partner created
Which partner(s) will use the data/who has access to the data?	DIA, WINGS	DIA, WINGS	DIA, WINGS, OTE
Data Format	.csv, .xml	.csv, .xlsx, .docx	.docx, .xlsx
Dissemination level • **Public** • **Confidential** • **Personal**	Confidential (for demo purposes some data will become publicly available)	Partially public (some sensitive information may be retracted, but most results will be published)	Public
FAIR: Will the data be shared in publicly open database?	No (commercially sensitive)	Partially	Yes
Data Storage (Location, Duration)	Company secure server, Duration of the project	Company secure server, Duration of the project Public data on Zenodo here	Teamwork Available after project end on Zenodo here
UC (Belgium, Romania, Greece)	Greece	Greece	Greece

Table 11: Data summary for use case 3 (Greece) partners – table 3 of 4.

Lead partner	**OTE**	**OTE**	**OTE**
Data Owner	OTE	OTE	OTE

Description of data	Network requirements & architectural design	Network related measurements based on network elements (RAN, core, edge) and end-devices (e.g., smartphones)	Report on network performance, calibration/settings, and measured network KPIs
Data Type	Requirements documents, architecture figures	Measurement data	reporting document, KPI results
Data Purpose	used to design and implement modules and processes	Network testing results	will be used as part of the evaluation of trial results
Data Origin	Partner created	Live test	Partner created
Which partner(s) will use the data/who has access to the data?	WINGS, OTE	WINGS, OTE	DIA, WINGS, OTE
Data Format	.docx, .xlsx	N/A	.docx, .xlsx
Dissemination level • **Public** • **Confidential** • **Personal**	Public	Confidential	Public
FAIR: Will the data be shared in publicly open database?	Yes	No	Yes
Data Storage (Location, Duration)	Teamwork Available after project end and on Zenodo after project here	company secure servers Duration of the project	Teamwork Available after project end and on Zenodo after project here
UC (Belgium, Romania, Greece)	Greece	Greece	Greece

Table 12: Data summary for use case 3 (Greece) partners – table 4 of 4.

Lead partner	**DIA**	**DIA**	**All GR partners**
Data Owner	DIA	DIA	All GR partners
Description of data	Questionnaire towards DIA workers regarding their experience	Questionnaire to third party experimenters using the GR solution	Aggregated GR trial results and GR UC validation
Data Type	experience reporting document	experience reporting document	Reporting document
Data Purpose	Evaluation of the solution from the perspective of the workers that will work with it	Evaluation of the solution from the perspective of potential additional users	Aggregated, measurements insights and conclusions from the GR trials
Data Origin	Questionnaire	Questionnaire	Partner created

Which partner(s) will use the data/who has access to the data?	DIA, WINGS, OTE	DIA, WINGS, OTE	DIA, WINGS, OTE
Data Format	.docx	.docx	.docx
Dissemination level • **Public** • **Confidential** • **Personal**	Public	Public	Public
FAIR: Will the data be shared in publicly open database?	Yes (following screening of data to ensure no personal information in free text fields of questionnaire)	Yes (following screening of data to ensure no personal information in free text fields of questionnaire)	Yes (following screening of data to ensure no personal information in free text fields of questionnaire)
Data Storage (Location, Duration)	Teamwork Available after project end on Zenodo here	Teamwork Available after project end on Zenodo here	Teamwork Available after project end on Zenodo here
UC (Belgium, Romania, Greece)	Greece	Greece	Greece

5.2.4 Cross-Use Case Datasets

This section covers partners and associated datasets that will be used, generated, and/or collected in VITAL-5G from activities that apply to all use cases. This covers information relevant to the VITAL-5G open repository, cross-use case KPIs and performance assessment datasets resulting from execution of the use cases trials.

Table 13: Data summary for partners with roles that impact several or all of the use cases.

Lead partner	**NXW**	**DHL**	**DHL**	**EBOS**
Data Owner	NXW	DHL	DHL	EBOS
Description of data	Source code and software documentation of the following VITAL-5G Platform components: - VITAL-5G Catalogue - VITAL-5G Service Lifecycle Manager - VITAL-5G Experiment Lifecycle Manager - License Manager - Multi-site Inventory	Use case and VITAL-5G platform requirements and target KPIs	Analysis and evaluation on the Network Applications, use case performance based on trial results and target KPIs	Component image/container and supporting files (images, etc.) for VITAL-5G frontend GUI interface
Data Type	Software source code Written report for the documentation	Written reports, KPI values	Network Applications results, trial results, KPI values and written reports	Software image/container

Data Purpose	The software components are part of the VITAL-5G Platform. Their release as open-source code is part of NXW exploitation strategy.	To ensure vertical industry needs are considered in use case definition and VITAL-5G platform	To evaluate, analyse and validate results from use cases	Enable VITAL-5G's front end for experimenters to interface with the VITAL-5G's portal backend functionality
Data Origin	Software development and testing.	Internal partner discussions	Data from Network Applications / test / KPIs results	EBOS
Which partner(s) will use the data/who has access to the data?	All. The software is published as open-source code.	WP1 colleagues	WP4 colleagues	EBOS will control this data; other partners can use as necessary
Data Format	Software: java files and maven project, OpenAPI, docker files. Documentation: readme files in software repository; documentation also reported in .docx format in D2.4	.docx,.xlsx	.docx,.xlsx	.xml, .json, .jpg, .png, .jsx, .js
Dissemination level • **Public** • **Confidential** • **Personal**	Public The software is published as open-source code.	Public	Public	Public
FAIR: Will the data be shared in publicly open database?	The software code is made available through a public software repository (Gitlab).	Will be shared as part of public WP1 deliverables	Will be shared as part of public WP4 D4.3 deliverable (VITAL-5G trial results analysis & post-processed evaluation)	Yes – Open/ public except in cases where confidential data is displayed
Data Storage (Location, Duration)	Location: Gitlab public repository here Duration: Project duration and at least 2 years after project end. Additional storage on NXW internal Gitlab repository for development, integration and testing activities.	Teamwork and project website	Teamwork and project website	Location: VITAL-5G Gitlab repository here Duration: Project duration and at least 2 years after project end.
UC (Belgium, Romania, Greece)	Applicable to all the UCs	All	All	All

5.2.5 VITAL-5G Platform Software Assets

Access to the V5G Platform software assets is through a Gitlab repository, the link to which is: https://gitlab.com/vital-5g/VITAL-5G_Platform

Where possible, Platform assets have been released as open-source code for third parties to extend or build new components on top of the existing platform, allowing new functionality to be implemented. This will be available for at least two years after the project end and is part of the project's exploitation strategy.

In addition, a second version of the platform is available as a set of docker-based images to provide an "as-is", pre-established version of the platform for third parties to re-use.

Table 14 lists the open-source VITAL-5G Platform components that are available for external third parties to interact with through the Gitlab repository.

Table 14: Overview of the publicly available open-source VITAL-5G Platform components.

Platform component	Lead VITAL-5G partner(s)
Portal web GUI	EBOS
Network Applications, Services & Experiments Catalogue	NXW
Service LCM	NXW
License Manager	NXW
Access Control	WINGS
Multi-site Inventory	NXW
Experiment LCM	NXW
Network Application Validator	BEIA
Intent-based requests manager	BEIA / WINGS
Slicing plugin for Nokia 5G Core	ORO
RO TB Testbed Monitoring Plugins	ORO

5.2.6 VITAL-5G Network Applications Software Assets

Access to the publicly available network applications software assets is through Gitlab at the following link: https://gitlab.com/vital-5g/vital-5g_netapps. This contains the various files to enable others to view and reuse the listed network applications, and includes readme files to facilitate reuse of the assets.

5.2.7 Other Project Datasets

The intention of this section is to summarise project datasets not covered in earlier sections of this report. These datasets relate to specific work packages that are not directly related to the implementation and validation of the project's use cases. As such, this section is segregated by work package, covering WP5 to WP8. As with the earlier sub-sections of Section 5, project deliverables will not be detailed here, instead the reader is referred to Section 3.6.2.

5.2.7.1 WP5: Commercialisation & Innovation

This sub-section is concerned with WP5 focusing on commercialisation and innovation. The most relevant datasets of interest to this work package are related to the business plan development and protection of intellectual property through patents. The work package will also include the use of questionnaires to gather feedback from consortium members and organisations external to the consortium.

The partners involved in WP5 are: ICP (WP lead), WINGS, BEIA, DHL, DIA, DT, EBOS, ATG, IMEC, INCE, NXW, TEL, ORO, OTE, SF, and all will be involved in the collection/generation of data. The data in this work package generally involves written documentation, which will be stored locally by the partners involved and on the project's Teamwork repository, generally in standard application formats (e.g. .pptx, .docx, .pdf, .jpg, etc.). Data will be preserved in accordance with data storage and archiving guidelines in Section 2.4 for the duration of the project.

Since WP5 relates to commercialisation and patent protection, some of the datasets managed within this WP will contain confidential information.

Of particular note is the use of questionnaires and feedback forms (used, for example, as part of outreach activities in WP5), any use of which must conform to the project's privacy statement (https://www.vital5g.eu/privacy-policy/) to ensure correct management of personal data.

5.2.7.2 WP6: Dissemination, Communication & Standardisation

WP6 involves the following partners: EBOS (WP lead), WINGS, BEIA, DHL, DIA, DT, ATG, IMEC, INCE, ICP, NAV, NXW, TEL, ORO, OTE, SF. Data in WP6 generally involves written documentation, such as scientific publications, social media posts, posters, presentations, flyers. These are stored locally by the partners involved and on the project's Teamwork repository, generally in standard application formats (e.g. .pptx, .docx, .pdf, .jpg, etc.).

Before the public release of any VITAL-5G related project data, partners are expected to consider the data sensitivities outlined in Section 3.1 to avoid any issues associated with personal data, confidentiality breaches, and/or interference with IP protection activities.

5.2.7.3 WP7: Project Management

The partners active in WP7 are: WINGS (WP lead), BEIA, EBOS, IMEC, ICP, NXW. The work package is concerned with general project management, technical project management, and quality assurance. The work in this WP involves creating and reviewing documents, spreadsheets and presentations to effectively and efficiently manage the project. Such files will be stored on the Teamwork repository, where necessary, for sharing with project partners. Due to the broad scope of the activities in this WP, the dataset sensitivities in Section 3.1 must be considered and confidentiality maintained, particularly in relation to personal data and financial data of project partners. Standard data formats are expected to be used (e.g. .pptx, .docx, .pdf, .jpg, etc.). Examples of the types of documents that will be manged in this work package are project deliverables, periodic reporting reports and spreadsheets, quality assurance documents (e.g., peer review records), risk assessments, and project Gantt charts.

Of particular note are the confidentiality and security considerations associated with administering access to the Teamwork repository. EBOS leads this aspect of the project management tasks and is experienced in ensuring access to the repository is well controlled.

5.2.7.4 WP8: Ethics Requirements

This work package is concerned with outlining the ethics requirements that the project must comply with and producing a deliverable describing these requirements and associated guidelines partners must implement in their day-to-day VITAL-5G activities. The main outcome of this work package is D8.1 : H - Requirement No. 1 [4], which was submitted to the European Commission in M06 and is confidential to project partners.

5.2.8 Datasets Needing Special Consideration

It is recognised that there are datasets being managed in the project on a day-to-day basis that have the potential to contain personal data. It is therefore important to directly discuss these issues and ensure policies on handling such datasets is communicated to project partners.

Table 15: Overview of datasets requiring special attention.

Dataset	Description
Video, photographic and sound recordings	Videos, photos and sound recordings can contain personal data in cases where people are captured in-frame or are recorded. This can occur inadvertently or otherwise during the course of an experiment depending on its nature. Partners are responsible for handling such occurrences accordingly to protect the people involved by anonymising data before it is shared internally within VITAL-5G. To do so, all personally identifiable information, including faces, must be obscured or removed from video and photographs, and any sound recordings must be edited to remove or mask any segments containing recordings of people, ensuring the recordings are anonymised.
VITAL-5G contact lists	This is an Excel spreadsheet containing contact information of the project partners and identifies what aspects of the project these are involved in. This file must be treated as confidential information and not shared outside the consortium. It is stored securely on the project's Teamwork repository and is to be used by partners only for project-related communications.
Meeting material	This refers to all documents created for project meetings (including minutes, agendas, attendance lists, presentations, etc.). These can potentially contain personal and other types of sensitive data and must be handled accordingly to protect the organisations and people referenced.
Workshops and training material	This refers to agendas, participation lists, presentations, etc associated with workshops/training sessions. If material is to be shared outside the consortium, it must be anonymised to remove all personal data, and checked for other confidential information, as discussed in Section 3.1.
Usage of cookies	If cookies are needed for the VITAL-5G website, or other web applications, a notification must be sent to the user regarding the use of cookies and the management of personal information.

Project related research data	All research data containing personal information must be anonymised before data is shared internally within VITAL-5G.
Personal data related to external stakeholders	For personal information gathered through outreach activities from external stakeholders, an opt-in/opt-out option must be provided regarding further communications from the VITAL-5G project, and all data must be stored securely in compliance with GDPR requirements (Section 3.1.1).

6 Conclusions

This deliverable is the Data Management Report for the VITAL-5G project and outlines the processes and guidelines for project partners to follow to ensure robust data management within the project and beyond the project's end date. The document highlights that it is the aim of VITAL-5G to publicly share as many of the research datasets generated/collected during the project as possible. It is also the goal of the project to publish research results in open access formats, to encourage wider reuse of project outcomes. Such sharing of data is subject to confidentiality concerns, the presence of personal data, security provisions, and IP protection considerations (as discussed in Sections 3.1 and 3.2).

This document also outlines how the FAIR data management principles were used as a basis to guide the development of the plan in this document. These guidelines ensure that data is handled in such a way that datasets are Findable, Accessible, Interoperable and Reusable (i.e. FAIR). To facilitate sharing of data, publications, software assets within the consortium and external to the project, different repositories are discussed.

The responsibilities of different project actors in terms of this data management plan are outlined to ensure those actors understand their role in maintaining compliance with the guidelines set out in this document. The document then summarised the datasets that will be handled as part of VITAL-5G, information on which was collated following a cross-project data collection exercise, where partners were requested to assess their roles in the project and the types of data they would need to generate, collect, and/or access. The data collector used to gather this information included specific fields and descriptions to assist partners in assessing those datasets across a number of criteria (for example, data purpose, data formats, data ownership, restrictions on data use, inclusion of personal data, etc.). This information provides a comprehensive overview of the datasets being managed within the project, allowing partners to understand which datasets can be made available publicly and which must be kept confidential.

This document presents the final version of data management planning for VITAL-5G, which has been revised and updated as necessary compared to the initial version of the DMP (D5.4, released in M12). This was to ensure all partners reassessed the datasets being generated / used in the project, ensuring they correctly captured updates that may have occurred in delivering the project objectives. Data management has also been reassessed at the project-level to ensure the processes outlined here are appropriate for the particulars of the project, particularly in terms of how data will be handled following the end of the project. To facilitate this sharing of project assets with interested third parties, various repositories have been highlighted in this document where publicly accessible data can be found.

7 References

[1] European Commission (2016). Guidelines on FAIR Data Management in Horizon 2020. Retrieved from https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

[2] VITAL-5G, Deliverable D5.4, "Data Management Plan". Online. Available:

[3] Complete guide to GDPR compliance. Online. Available: https://gdpr.eu/

[4] VITAL-5G, Deliverable D8.1, "H - Requirements No. 1". Project confidential document.

[5] Intersoft Consulting, GDPR Personal Data, Retrieved from https://gdpr-info.eu/issues/personal-data/

[6] www.teamwork.com

[7] https://aws.amazon.com/compliance/data-privacy/

[8] https://www.teamwork.com/security/iso/

[9] https://zenodo.org/

[10] https://www.openaire.eu/

[11] European Commission, Open access & Data management, Retrieved from https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-dissemination_en.htm

[12] VITAL-5G, Deliverable D7.4, "Project Quality Handbook", Project confidential document.

[13] VITAL-5G, Deliverable D1.1, "Report on use case requirements". Online. Available: https://www.vital5g.eu/wp-content/uploads/2022/05/VITAL5G-D1.1_Report-on-Use-case-requirements-v2.0.pdf

Annex I: Data Management Plan Data Collector

The format of the Data Collector used to record the data that will be collected from each trial site, based on VITAL-5G partners' input, is presented in this Annex.

Vital-5G DMP Data Collector

Instructions:

1. Please think through planned work in each UC (from requirements generation, through to execution of the UC experiments, including feedback from users) to identify datasets that will be used and/or generated during the project.
2. Please use the "Data Collector" table to provide a high-level overview of those datasets.
3. Please group all similar datasets together in one column of the data collector to minimize reporting effort by partners.
4. Please use a copy of the table for each different dataset being reported.
5. The "Technical UC data" list below is intended to provide a non-exhaustive list of examples of datasets that could potentially be reported on. Please use this as a guide to assist with the exercise.
6. The completed tables at the end of this document provide an understanding of the required level of information expected for each dataset.

Thank you for your assistance with this exercise.

Examples of datasets to report:

The list below is intended as a guide to help identify datasets to consider for the data management plan (DMP)

Technical UC data

1. Sensor data/other measurement data (e.g. T2.4 "... port monitoring services, the tracking of dangerous goods, on-demand video-surveillance systems, the optimization of order management")
2. Source code, external open source code, including AI and ML algorithms, other "intelligent applications"
3. UC Requirements
4. UC Platform requirements and specification
5. UC questionnaire feedback
6. UC design files (drawings, calculations/simulations, architecture proposals, equipment specifications, platform specification, etc.)
7. Network Application descriptions
8. UC evaluations; UC user/experimenter feedback + respondent identifying information.
9. Results from demonstrators (KPIs, related statistics, "Network Application operability" T4.2)
10. Literature review/State-of-the-art review (content ends up in deliverable and should be captured by partner documenting data used for related WP)

Datasets **not** to report

This section shows the types of data not to be reported on by partners. These types of information sources will be summarised in a separate section of the DMP.

General project data:

1. Deliverables
2. Dissemination material (publications, posters, presentations, flyers, questionnaires, social media related, etc.)
3. Project management material (internal/EC reporting, meeting material (minutes, agenda, presentations etc), other supportive documents (incl. internal documents)

Day-to-day data that includes personal information:

1. Project contact lists
2. Meeting related material (agendas, presentations, signature lists, minutes)
3. Workshops/Conferences and Training sessions
4. Reporting documents
5. Deliverables, internal documents and other project reports
6. Publications
7. List of stakeholders (external to project)

Data Collector

Lead partner	Partner responsible for dataset generation and/or data collection	
Data Owner	Owner of data	
Description of data	High-level, short description of the data (e.g. sensor data to capture ship location in port area)	
Data Type	Summarise type of data e.g. measurement data, video, written reports, KPI results, etc.	
Data Purpose	Short description of why the data is necessary and what will be done with the data (e.g. positional data used as input for automated route selection algorithms)	
Data Origin	Specify if data is from test, partner, database, simulation, questionnaire, etc.	
Which partner(s) will use the data/who has access to the data?	Other partner(s) short name(s), if appropriate, external parties, N/A otherwise	
Data Format	File format used to store data (e.g. CSV, matfiles, png, .docx,.xlsx, etc.)	
Dissemination level • **Public** • **Confidential** • **Personal**	Does the data contain personal information and need to be treated according to the GDPR? Is the data confidential to the consortium? Is the data confidential because it is commercially sensitive for industry partners (warehouse operators, port authorities)?	
FAIR: Will the data be shared in publicly open database?	e.g. to support scientific publication on repositories such as Zenodo, Open Research Europe. If not, a reason must be given (e.g. privacy, export regulations, IP, commercially sensitive)	
Data Storage (Location, Duration)	Location: e.g. Teamwork repository, company secure servers, Duration: e.g. Project duration, 5 years after project end, temporary storage on company servers only for calculations/processing	
UC (Belgium, Romania, Greece)	Applicable UC	

Please repeat table for additional datasets

Examples of completed data collectors

Lead partner	Partner responsible for dataset generation and/or data collection	Partner 1	Partner 2	Partner 3
Data Owner	Owner of data	Partner 1	Partner 2	Partner 3
Description of data	High-level, short description of the data (e.g. sensor data to capture ship location in port area)	KPIs from testbed	Data for positioning reference signals	Collecting sensor data for Tx,y
Data Type	Summarise type of data e.g. measurement data, video, written reports, KPI results, etc.	KPIs	Numerical data; coordinates	Speed, temperature, camera footage
Data Purpose	Short description of why the data is necessary and what will be done with the data (e.g. positional data used as input for automated route selection algorithms)	Quantify project results	Project needs to develop algorithms	Input to the AI/ML algorithms
Data Origin	Specify if data is from test, partner, database, simulation, questionnaire, etc.	Testbed	Simulated data for device positioning	UC measurements
Which partner(s) will use the data/who has access to the data?	Other partner(s) short name(s), if appropriate, external parties, N/A otherwise	Project partners involved in UC	Other partners who will need the data	Consortium partners after anonymisation techniques applied
Data Format	File format used to store data (e.g. CSV, matfiles, png, .docx,.xlsx, etc.)	.doc	CSV, matfiles, png	Text, csv
Dissemination level • **Public** • **Confidential** • **Personal**	Does the data contain personal information and need to be treated according to the GDPR? Is the data confidential to the consortium? Is the data confidential because it is commercially sensitive for industry partners (warehouse operators, port authorities)?	Public + Confidential	Public + Confidential	Confidential
FAIR: Will the data be shared in publicly open database?	e.g. to support scientific publication on repositories such as Zenodo, Open Research Europe. If not, a reason must be given (e.g. privacy, export regulations, IP, commercially sensitive)	Public data: yes	If part of publication	No - commercial
Data Storage (Location, Duration)	Location: e.g. Teamwork repository, company secure servers, Duration: e.g. Project duration, 5 years after project end, temporary storage on company servers only for calculations/processing	Local storage/reported in deliverables; duration of project	Partner network/ repository, Project duration	Temporary storage only for calculations /processing; data deleted afterwards
UC (Belgium, Romania, Greece)	Applicable UC			