5G**LOGINNOV**

D6.5

Open data publication

www.5g-loginnov.eu

Work Package 6	Project management
Task 6.5	Open data publication
Authors	Mandimby Ranaivo Rakotondrovelona (AKKA) Djibrilla Amadou Kountche (AKKA) Xavier Valente (AKKA)
Dissemination Level	Public
Status	Final
Due date	30/11/2023
Document Date	30/11/2023
Version Number	1.0

Quality Control

	Name	Organisation	Date
Editor	Mandimby RANAIVO R.	AKKA	17/11/2023
Peer review 1	Peter Schmitting	ERTICO	21/11/2023
Peer review 2	Andoni Cortés Vidal	VICOM	23/11/2023
Authorised by (Technical Coordinator)	Eusebiu Catana	ERTICO	28/11/2023
Authorised by (Quality Manager)	Eusebiu Catana	ERTICO	29/11/2023
Submitted by (Project Coordinator)	Eusebiu Catana	ERTICO	30/11/2023

Legal Disclaimer

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

List of abbreviations and acronyms

Abbreviation	Meaning
5G PPP	The 5G Infrastructure Public Private Partnership
CAD/CAM	Connected Automated Driving / Connected and Automated Mobility
CCAM	Cooperative, Connected and Automated Mobility
CCTV	Closed Circuit Television
CID	Collaborative Information Days
CSF	Critical Success Factor
DL	Downlink
DMP	Data Management Plan
DPO	Data Protection Officer
EC	European Commission
FAIR	Findable, Accessible, Interoperable and Reusable
GDPR	General Data Protection Regulation
GNSS	Global Navigation Satellite System
GPS	Global Position System
HTTPS	Hyper Text Transfer Protocol (HTTP) over Secure Socket Layer (SSL)
IPsec	Internet Protocol (IP) security
ITS-G5	Intelligent Transport Systems (ITS) 5 GHz wireless communication
KER	Key Exploitable Result
KPI	Key Performance Indicator
MCL	Maximum Coupling Loss
MEC	Mobile Edge Computing
MNO	Mobile Network Operator
NFV-MANO	Network Function Virtualisation (NFV) - Management And Network Orchestration (MANO)
OBU	On-Board-Units
ORDP	Open Research Data Pilot
POPD	Protection Of Personal Data
PU	Public
QoE	Quality of Experience
SME	Small and Medium Enterprise

TMT	Technical Management Team
TRxP	Transmission Reception Point
UE	User Equipment
UL	Uplink
VIN	Vehicle Identification Number
VSaaS	Video Surveillance as a Service
WP	Work Package

EXECUTIVE SUMMARY

The 5G-LOGINNOV project aims at showcasing the added-values of the 5G-supported technologies to the port logistic operations. As such, the primary objective of the data collection within the project is to assess the project's 5G-enabled use cases benefits. The Open access to Research Data Pilot (ORDP) aims to improve access to and re-use of research data generated by Horizon 2020 projects. 5G-LOGINNOV is fully committed to ORDP. Hence, the development of a data management plan (DMP) is required and is embodied in this deliverable which constitutes the final version.

The data generated and collected over the course of the 5G-LOGINNOV project fall into the following categories:

- Technical data include the data related to the technical development and operation of the use cases. The technical data is generated by the Living Labs components including sensors and existing platforms.
- Evaluation data concern the data used to compute the KPIs and the KPIs themselves required to perform the evaluation tasks of the project. They are collected during the trials and stored using a central data collection tool.
- Open research data contain the data and results that are published by the project to comply with ORDP requirements. They are a subset of the evaluation data which are themselves derived from the technical data.
- Internal administrative data: refer to the data generated and shared internally for administrative and management purposes.

The data management in 5G-LOGINNOV highlights the importance of the FAIR approach aiming at making the data findable, accessible, interoperable, and accessible. For that, a set of recommendations are stressed including the documentation of the data with the relevant metadata, the preference for the open data formats, the generation of digital object identifiers (DOI) for the published datasets, the choice of an open license, the use of the public data repository Zenodo.

Data security aspects are considered across the whole data lifecycle, with the most important features being the access control with authentication and authorization mechanisms, the communication channel security when transferring data, the data integrity checks and regular backup. All these features contribute to the data confidentiality, which is of utmost importance, especially when dealing with personal data. They constitute an implementation of the privacy-preserving measures ensuring a GDPR-compliant process following the recommendations from the project's ethics work package.

1   INTRODUCTION

1.1   Project overview

5G-LOGINNOV will focus on seven 5G-PPP themes and support to the emergence of a European offer for new 5G core technologies across 11 clusters of use cases. 5G-LOGINNOV's main aim is to design an innovative framework addressing integration and validation of CAD/CAM technologies related to the industry 4.0 and ports domains by creating new opportunities for LOGistics value chain INNOVation. 5G-LOGINNOV is supported by 5G technological blocks, including new generation of 5G terminals notably for future Connected and Automated Mobility, new types of Internet of Things 5G devices, data analytics, next generation traffic management and emerging 5G networks, for city ports to handle upcoming and future capacity, traffic, efficiency, and environmental challenges. 5G-LOGINNOV will deploy and trail 11 clusters of use cases beyond TRL7 including a GREEN TRUCK INITIATIVE using CAD/CAM & automatic trucks platooning based on 5G technological blocks. Thanks to the new advanced capabilities of 5G in wireless connectivity and Core Network agility, 5G-LOGINNOV ports will not only significantly optimise their operations but also minimise their environmental footprint in the city and the disturbance to the local population. 5G-LOGINNOV will be a catalyst for market opportunities built on 5G Core Technologies in the Logistics domains, thus being a pillar of economic development and business innovation and promoting local innovative high-tech SME and Start-Ups. 5G-LOGINNOV will open SMEs' and Start-Ups' door to these new markets using its three Living Labs as facilitators and ambassadors for innovation on ports. 5G-LOGINNOV's promising innovations are key for the major deep-sea European ports in view of the mega-vessel era (Hamburg, Athens), and are also relevant for medium sized ports with limited investment funds (Koper) for 5G[1].

1.2   Purpose of the deliverable

The present deliverable describes the data management approach applied over the course of the project's data lifecycle from the collection to the publication. It summarizes the purpose of data collection in the project, what are the categories of data manipulated during the project and their utility. It also outlines the approach to make 5G-LOGINNOV data FAIR i.e., findable, accessible, interoperable, and reusable by indicating what data will be generated, collected, and processed, how the research data will be preserved, and which parts of the datasets will be shared for evaluation needs and to comply with the ORDP requirements. Finally, this document provides insights in the ethical and privacy aspects as well as the recommended data security principles. This deliverable provides an update of the first version of the data management plan deliverable D6.4 (5G-LOGINNOV, 2021) and constitutes the final iteration.

This deliverable serves as an entry point to understand the project-wide data management approach in 5G-LOGINNOV. As such, it provides the big picture of data management at the project level while dealing with more focus on research data management as required by H2020 ORDP. As an entry point, it provides links towards and between more specialised deliverables, namely the D1.4 – Initial Specification of evaluation and KPIs (5G-LOGINNOV, 2022) which focuses on the evaluation data requirements, the D1.5 – Data and cyber-protection policies (5G-LOGINNOV, 2021) which provides requirements regarding data handling and cybersecurity and the D2.2 – Data collection and evaluation procedures (5G-LOGINNOV, 2022) which describes the data collection tools. Additionally, information about GDPR and personal data are given in the deliverable D1.5 (5G-LOGINNOV, 2021) and D7.2 – POPD - Requirement No. 2 (5G-LOGINNOV, 2021). This relationship is illustrated on Figure 1.

---

[1] https://5g-loginnov.eu/
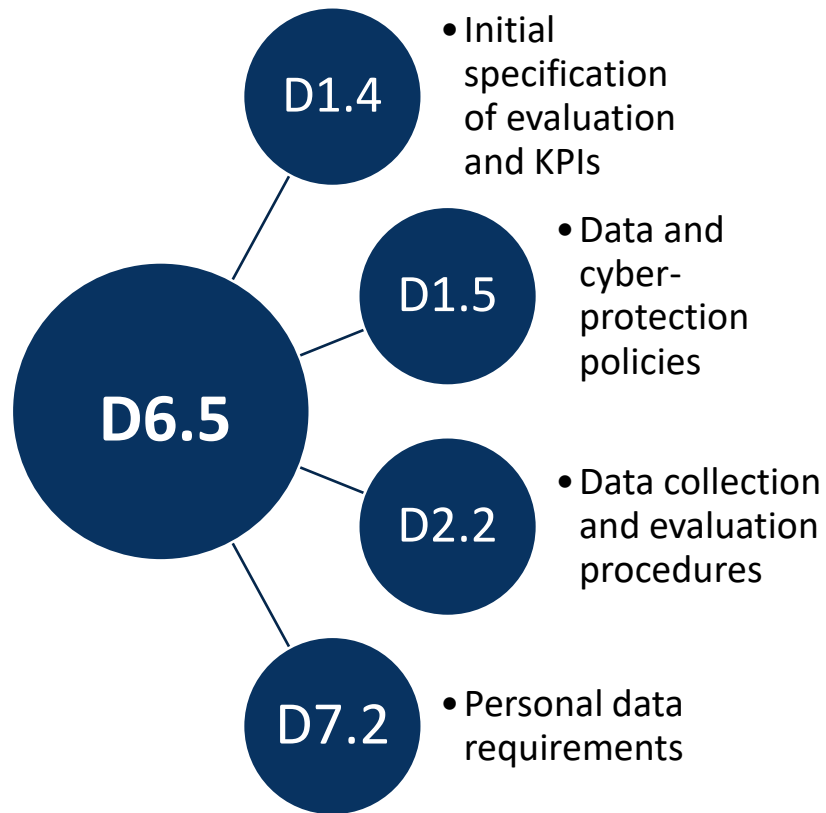
Figure 1: Link between D6.5 and data-related deliverables

1.3 Intended audience

The dissemination level of D6.5 is public, and it is available to the members of the consortium, the European Commission (EC) services and those external to the project.

This document is primarily intended to serve as an internal guideline and reference for all 5G-LOGINNOV beneficiaries, especially the governance bodies such as the General Assembly, the Steering Committee, the Technical Management Team, and the Advisory Board.

This deliverable is also of high interest for anyone willing to understand the data management procedures followed during the project as part of the process of reusing the published datasets.

2   5G-LOGINNOV DATA

2.1   Purpose of the data collection

The 5G-LOGINNOV project aims at showcasing the added-values of the 5G-supported technologies to the port logistics. As such, the primary objective of the data collection is to assess the project's 5G-enabled use cases benefits. More specifically, the data collection in 5G-LOGINNOV can be linked to three phases in the project as illustrated on Figure 2. The technical development phase corresponds to the set up and execution of the use cases at the Living Labs. The data handled during this phase corresponds to the data generated by the Living Labs data sources that are involved in the technical implementation and the use case operations. The evaluation phase corresponds to the evaluation of the use cases following the methodology described in D1.4 (5G-LOGINNOV, 2022). It requires the collection of a subset of the data generated during the technical development phase on which the evaluation is performed. Finally, the publication phase consists of the selection and the publication of a subset of the data used for evaluation as research data at the end of the project following the ORDP requirements. Only the data on which the partners agree to make openly available will be published and will constitute the open research data.
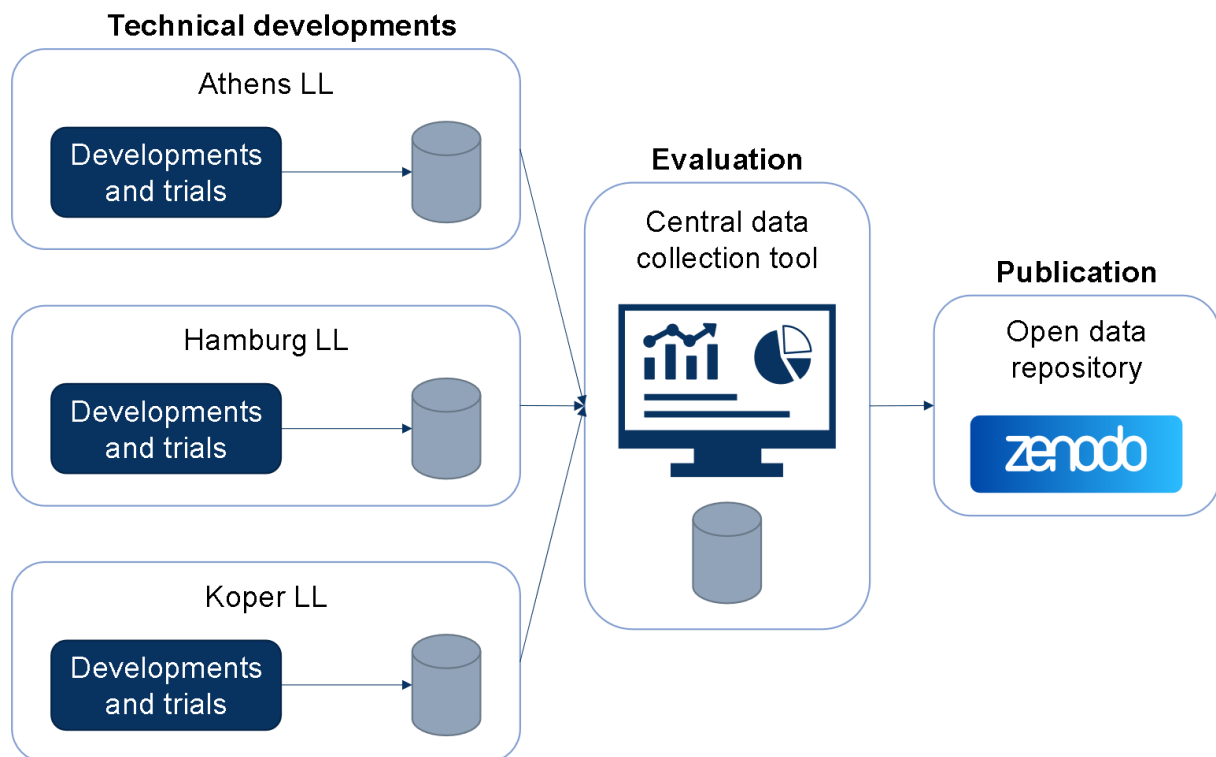
Figure 2: Data management phases in 5G-LOGINNOV

2.2   Data categories

This section focuses on the categories of data that are handled during the project which are: the *technical data*, the *evaluation data*, the *open research data*. and the *internal administrative data*. A short description of these categories is given in Table 1 while more details on the corresponding data management are provided in the upcoming subsections.

Category	Short description
Technical data	This category includes the data related to the technical development and operation of the use cases. The technical data is generated and handled by the Living Lab components including sensors and existing platforms. This category is discussed in more detail in 2.2.1.
Evaluation data	The evaluation data concerns the data used to compute the KPIs described in T1.4 for the evaluation of the project. They are collected during the trials and stored using tools developed in T2.2. This category is discussed in more detail in 2.2.2.
Open research data	This category contains the data and results that will be published by the project to comply with ORDP requirements. The open research data is a subset of the evaluation data which is itself a subset of the technical data. It is discussed in more details in 2.2.3.
Internal administrative data	This category refers to the data generated/shared internally for administrative and management purposes. It is addressed in 2.2.4.

Table 1*:* 5G-LOGINNOV data categories

2.2.1 Technical data

The technical data category includes the data related to the technical developments (see Figure 2). The technical data are generated at the Living Lab level with the data sources including the sensors, the existing platforms, the network components but also the Living Labs stakeholders. The main purpose of the technical data collection is the calculation of the KPIs necessary for evaluation. The technical data may contain personal data. In this case, the GDPR rules apply (see chapter 6). Regarding the sharing of the technical data, it is up to the Living Labs stakeholders to decide what data they accept to be shared with the evaluation team and later published under ORDP. The main categories of technical data are described on Table 2.

Categories	Sub-categories	Description
Logistics data	Vehicle dynamic	It includes the data related to the dynamic of the vehicles especially trucks such as the speed, the acceleration, the stand still time etc.
	Sensors	It includes data from sensors to collect different types of data such as the position (GPS), temperature, battery level, fuel level and consumption etc.
	Port logistics	It includes various data related to the port logistics such as the number of vehicles under maintenance, the parts in stocks, the amount of time during which trucks are parked etc.
5G-related data	Network performance	This sub-category contains data related to the network performance such as latency, bandwidth, jitter, reliability, the area traffic capacity etc.

	MEC	This sub-category contains data related to the performance of the MEC in terms of end-to-end latency and the efficiency of the network operation.
	NFV-MANO	These data are related to the performance of the NFV-MANO about the on-demand and automatic deployment, high-availability, resilience of operation of demanding logistic services and IoT-5G.
	Precise Positioning	These data are related to the accuracy of the Precise Positioning which is particularly important in the logistics domain. The accuracy and the characteristics of the GNSS and additional components such as the Ultra-Wide Band might also be logged.
Application data	AI model performance	These data help in monitoring the performance of the AI models trained and operated with the help of 5G-obtained data. Examples of indicators are the model accuracy, the model inference time etc.
	Traffic Management Application	These are data related to traffic management such as the traffic volume (number of vehicles per time) and traffic density (number of vehicles per road segment). The positions at which the volume and the density are measured are also logged.
	High-performance CCTV Surveillance Applications (including VSaaS)	These data are gathered through the 5G Closed Circuit Television (5G CCTV) which contains video streams or recorded images from surveillance cameras.
	Real-Time Tracking & Enhanced Visibility	These data are related to portable 5G tackers that collect in real time the location and condition of the trucks/goods.

Table 2. Technical data

The technical data are produced/collected/shared during the use case operations allowing the execution of the trials. Thus, they are of high importance for the evaluation phase of the project. In fact, the evaluation team relies on a subset of the technical data to conduct the evaluation processes (see section 2.2.2). Some of the technical data are confidential to the Living Labs while others will be provided to the evaluation team and will be later publicly available following the ORDP scheme.

The requirements regarding the handling of technical data and the cybersecurity aspects are discussed in detail in D1.5 (5G-LOGINNOV, 2021).

2.2.2 Evaluation data

The evaluation data category includes the technical data that will be collected during the trials. The data is selected following the evaluation methodology, the main purpose being the calculation of KPIs. The evaluation data collection and storage rely on the data collection tools described in D2.2 (5G-LOGINNOV, 2022).

The project's evaluation methodology is defined in Task 1.4 – Evaluation methodology and requirements and depicted on Figure 3.
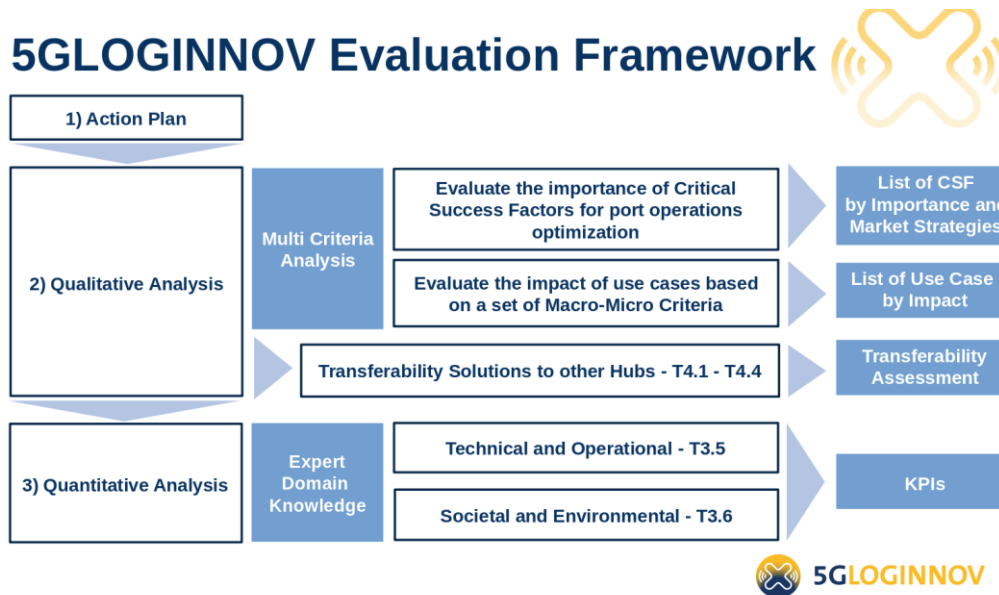

Figure 3: 5G-LOGINNOV Evaluation Methodology

This evaluation methodology uses a Multicriteria Analysis method based on two main components: i) the Key Performance Indicators (KPIs) and Macro/Micro Criteria and ii) the Critical Success Factor (CSF). The KPIs are collected at each Living Lab and used to assess the impact of the 5G-LOGINNOV Platform and of the Use Cases in the Living Lab. The Macro/Micro Criteria method is a business strategic approach that breaks down the goals of the project (i.e., Macro Criteria) into measurable objectives (i.e., Micro Criteria) that can then be evaluated using selected metrics (KPIs), while the latter extracts knowledge from port managers, employees and other stakeholders to understand whether the technologies introduced by the 5G-LOGINNOV project improved the previous operations. The complete list of KPIs is provided on Table 3 with the corresponding data structure available in D2.2 (5G-LOGINNOV, 2022).

Living Lab	KPI ID	KPI name
Athens LL	A-KPI1	Model inference time
	A-KPI2	Model accuracy/reliability
	A-KPI3	Deployment Time
	A-KPI4	Percent of empty containers runs
	A-KPI5	Mean time of container job
	A-KPI6	Time needed the device to open a network connection
	A-KPI7	CO_2 emissions
	A-KPI8	Fuel consumption
	A-KPI9	Human resource optimization
	A-KPI10	Vessel operation completion time
	A-KPI11	Model inference time

	A-KPI12	Model accuracy/reliability
	A-KPI13	Parts in stock
	A-KPI14	Vehicle breakdown
	A-KPI15	Vehicle under maintenance
	A-KPI16	Vehicles unexpected breakdown
	A-KPI17	Maintenance costs of vehicle
	A-KPI18	Assets idling
	A-KPI19	Area traffic capacity
	A-KPI20	Bandwidth
	A-KPI21	Connection density
	A-KPI22	Reliability
	A-KPI23	End-to-end latency
	A-KPI24	One-way latency
	A-KPI25	User experienced data rate
Hamburg LL	H-KPI1	Avg. truck speed single mode
	H-KPI2	Avg. acceleration activities single mode
	H-KPI3	Avg. stillstand time single mode
	H-KPI4	Truck speed profile by platoon mode
	H-KPI5	Acceleration profile by platoon mode
	H-KPI6	Stillstand time profile by platoon mode
	H-KPI7	Fuel consumption single mode
	H-KPI8	CO_2 emissions single mode
	H-KPI9	Fuel consumption platoon mode
	H-KPI10	CO_2 emissions platoon mode
	H-KPI11	Energy performance index value EPI
	H-KPI12	Acceleration performance index value API
	H-KPI13	Available 5G bandwidth on urban roads
	H-KPI14	Positioning quality on urban road networks with 5G
	H-KPI15	Latency by 5G cellular communication in urban areas
	H-KPI16	Packed Error Rate (PER) in 5G NSA production network

Koper LL	K-KPI1	Components Onboarding and Configuration Time - 5G IoT backend
	K-KPI2	Deployment Time - 5G IoT backend
	K-KPI3	Time to scale - 5G IoT backend
	K-KPI4	Service Availability - 5G IoT backend
	K-KPI5	Components Onboarding and Configuration Time - 5G IoT agent
	K-KPI6	Deployment Time - 5G IoT agent
	K-KPI7	Components Onboarding and Configuration Time - 5G CN and 5G BBU
	K-KPI8	Deployment Time - 5G CN and 5G BBU
	K-KPI9	Time to scale - 5G CN and 5G BBU
	K-KPI10	Service Availability - 5G CN and 5G BBU
	K-KPI11	Slice Reconfiguration - 5G CN and 5G BBU
	K-KPI12	5G Network Area Traffic Capacity
	K-KPI13	5G Network Availability
	K-KPI14	5G Network Bandwidth
	K-KPI15	5G Network Connection Density
	K-KPI16	5G Network Coverage Area Probability
	K-KPI17	5G Network End-to-End Latency
	K-KPI18	5G Network Reliability
	K-KPI19	Model accuracy
	K-KPI20	Model inference time
	K-KPI21	Model accuracy
	K-KPI22	Model inference time
	K-KPI23	Model accuracy
	K-KPI24	Model inference time
	K-KPI25	Time Trucks Parked in the Area
	K-KPI26	Truck speed
	K-KPI27	Truck Acceleration
	K-KPI28	Truck Stand Still Time
	K-KPI29	Fuel Consumption

Table 3. 5G-LOGINNOV KPIs as evaluation data

2.2.3 Open research data

5G-LOGINNOV has agreed to participate in the Horizon 2020 Open Research Data Pilot and follows the guidelines associated with 'open' access to ensure that the results of the project are openly available.

The open research data category includes the data that will be made openly available at the end of the project. It corresponds to a subset of the evaluation data on which the partners agree to publish. Hence, the open research datasets can be of the subcategory of technical data described in 2.2.1 or can include processed data like the KPI calculated during the evaluation (Table 3).

The list of datasets agreed to be published on the project's open data repository is provided in Annex 1.

2.2.3.1 ORDP participation

5G-LOGINNOV will ensure open access to all peer-reviewed scientific publications relating to its results and will provide access to the research data needed to validate the results presented in deposited in scientific publications. The following lists the minimum required fields of metadata that should come with a 5G-LOGINNOV project-generated scientific publication in a repository:

- The terms: "European Union (EU)", "Horizon 2020"
- Name of the action: Research and Innovation Action
- Acronym and grant number: 5G-LOGINNOV, 957400
- Publication date
- Length of embargo period if applicable
- Persistent identifier

When referencing open access data, 5G-LOGINNOV will include at a minimum the following statement demonstrating EU support (with relevant information included into the repository metadata):

"This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 957400".

2.2.4 Internal administrative data

This category refers to the data produced by the project management activities such as meeting minutes, recordings, internal reports, for historical purposes and follow-up. The data is collected by the management team including the project manager, the WP leaders and task leaders. The data is stored using a project management tool that requires the authentication of the users. The internal administrative data is confidential, only for the members of the consortium.

2.3 Data utility

The ultimate objective of the 5G-LOGINNOV data collection is the assessment of the 5G benefits on the port logistics operation. According to the data management phases in Figure 2, the technical data are required when performing the evaluation activities producing the evaluation data. The technical data and the evaluation data are then useful to the evaluation task participants to calculate and produce the KPIs necessary to evaluate the project approach.

The open research data are made available to the research community as part of the state of the art, to encourage comparison between methodologies with other approaches or projects. In particular, some of the non-confidential KPIs published at the end of the project could be used to perform an evaluation in a cross-Living Lab context or a comparison with future approaches regarding the integration of 5G-supported technologies in ports deployment.

2.4 Dataset description

This section provides guidelines on how to describe the different types of datasets to be collected and shared by 5G-LOGINNOV after the end of the project with respect to ORDP which aims to improve and maximise access to, and re-use of research data generated by Horizon 2020 projects.

The description of the different datasets should provide information on their reference, file format, standards, methodologies, metadata, and repository to be used. More details are given below. Table 4 provides the template to be used to describe 5G-LOGINNOV datasets.

Dataset Reference	Each dataset will have a reference that will be generated by the combination of the name of the project, the trial site, the use case in which it is generated and the datatype: "5G-LOGINNOV_LivingLab-Site_UC_Datatype".
Dataset Name	Name of the dataset.
Dataset Description	Each dataset has a full data description explaining the data provenance, origin, and usefulness. Reference may be made to existing data that could be reused.
Standards and metadata	The metadata attributes list to be used to find the dataset. Metadata can be split into 4 categories: • Design and execution documentation, which corresponds to a high-level description of a data collection. • Descriptive metadata, which describes each component of the dataset (including origin and quality). • Structural metadata, which describes how the data is being organised. • Administrative metadata, which set the conditions on how the data can be accessed and how this is being implemented.
File Format	Any format that defines data.
Data Sharing	Explanation of the sharing policies related to the dataset among the following options: • Open: Open for public disposal. • Embargo: It will become public when the embargo period applied by the publisher is over. In case it is categorised as embargo, the end date of the embargo period must be written in DD/MM/YYYY format. • Restricted: Only for project internal use. Each dataset must: • Have its distribution license. • Provide information about personal data. • Mention if the data is anonymised or not. • Tell if the dataset entails personal data and how this issue is considered.
Archiving and Preservation	The preservation guarantee and the data storage during and after the project (for example databases, institutional repositories, public repositories, etc.).

Table 4: Template for dataset description

3   ON MAKING 5G-LOGINNOV DATA FAIR

5G-LOGINNOV uses the FAIR (Findable, Accessible, Interoperable and Reusable) approach for the data generated during the project. The FAIR principles aim to improve the practices for data management and data curation. These principles can be applied to a wide range of data management purposes, whether it is data collection or data management of larger research projects regardless of scientific disciplines. The FAIR principles are described in the guidelines for H2020 data management (European Commission Directorate-General for Research & Innovation, 2016) and they serve as a template for data lifecycle management. They also ensure that the most important components for lifecycle are covered. 5G-LOGINNOV commits to the following actions to implement the FAIR principles.

Making data Findable, including provisions of metadata in 5G-LOGINNOV datasets:

- The datasets will have rich metadata to facilitate the findability.
- Open data formats (csv, xml) are preferred.
- All the datasets will have Digital Object Identifiers provided by a public repository.
- The standards for metadata will be provided for each dataset if applicable as presented in Annex 1.

Making data openly Accessible:

- The datasets for evaluation are described using the recommendations of metadata description from the 'Standards and Metadata' entry in Table *4*.
- The data and their associated metadata will be made available either in a public repository or in an institutional repository. ZENODO[2] has been chosen as the public repository to host the project's open research data.
- The datasets are retrievable by their identifier using a standardised, open, free, and universally implementable communications protocol.
- The protocol allows for an authentication and authorization procedure, where necessary.
- Table 4 is used to provide information on the methods or software used to access the data.

Making data Interoperable:

- The metadata vocabularies, standards and methodologies depend on the public repository and use the recommendations of metadata description from the Standards and Metadata entry in the Table 4.
- The definition of the data formats is provided in WP1 (Living Labs requirements & specifications) and WP2 (Living Labs development and deployment). The goal is to have the same formats across Living Labs which will enable the development of common data quality check tools and enable partners responsible for the evaluation dealing with the same formats across all pilot sites.

Making data Re-usable (through clarifying licenses):

- All data producers license their data to allow the widest reuse possible. The open research data are licensed under the Creative Commons Attribution 4.0 International license
- By default, the data are made available for reuse. If any constraints exist, an embargo period will be mentioned in the data sharing row of Table 4 to keep the data from reuse for a limited period of time.

---

[2] https://zenodo.org/

4 ALLOCATION OF RESOURCES

The cost to make the data FAIR in 5G-LOGINNOV shall be handled by each partner having to generate data according to the requirements expressed in the data management plan.

Akkodis (formerly AKKA) is the organization responsible for the project's data management and appoints a data manager liaising with the TMT (Technical Management Team) about the data management issues. The data manager leads the data management plan tasks and participates in the project coordination monitoring the evaluation data collection, storage, and handling, as well as the data publication as part of the ORDP.

All research data collected as part of this project are owned by the data producers or partners involved in the Living Labs. The partners in 5G-LOGINNOV take the responsibility for the collection, management, and sharing of the research data. Quality assessment falls under the responsibility of the data manager of each Living Lab.

5   DATA SECURITY

This chapter presents the overall data security requirements to be fulfilled by the project and each Living Lab. It is important to note that D1.5 (5G-LOGINNOV, 2021) details the requirements and proposed tools and methods to ensure compliancy with GDPR and cyber protection. D1.5 is confidential, only for members of the consortium, including the Commission services.

The data produced during the execution of 5G-LOGINNOV shall be stored per Living Lab in local servers if deemed necessary and in a central server corresponding to the central data collection tool described in D2.2 (5G-LOGINNOV, 2022) for the whole project. The data are made compliant with the GDPR as described in the Deliverable D7.2 (5G-LOGINNOV, 2021). The data security measures are implemented in the frame of the data management phases described in Figure 2.

Authentication and authorization: The access to 5G-LOGINNOV data is only available to the authenticated and authorised users. These categories and the rights of those users are defined and enforced. The appropriate access control policies and mechanisms (including physical access control) are identified for each trial site and also project wide to provide the authorization.

- The access to the technical data is fully controlled by the Living Lab. As such, it is up to the Living Lab to set up the corresponding authentication and authorization mechanisms.
- The access to the evaluation data relies on the authentication and authorization mechanisms provided by the central data collection tool.

Accounting: In 5G-LOGINNOV any access and modification to a resource by any user is securely logged in order to prevent users from denying that data files were accessed, altered or deleted, when auditing.

Confidentiality: The data stored in 5G-LOGINNOV servers shall be encrypted during transmission and storage.

Communication security: Access to 5G-LOGINNOV servers shall be done through encrypted communication channels such as HTTPS.

Data integrity: The data collected during 5G-LOGINNOV shall be protected from malicious and accidental modifications by any users during their transmission or their storage.

Availability: This security principle assures that the 5G-LOGINNOV servers shall be available for 5G-LOGINNOV users during the defined interval of service.

Back up: Regular backups of the data shall be programmed to mitigate any data loss due to failure from the data management tools.

6   ETHICAL AND PRIVACY ASPECTS

Ethics requirements are addressed in WP7 – Ethics requirements, and are thoroughly described in the following three deliverables. Their dissemination level is confidential, only for members of the consortium including the Commission services.

D7.1: Ethics - Requirement No.1 (5G-LOGINNOV, 2021)

This deliverable presents the procedures and criteria to be used to identify/recruit research participants. The informed consent procedures that will be implemented for the participation of humans must be submitted as a deliverable. Templates of the informed consent/assent forms and information sheets (in language and terms intelligible to the participants) are also submitted as a deliverable.

D7.2: POPD - Requirement No. 2 (5G-LOGINNOV, Due in February 2021)

It offers a description of the security measures to be implemented in order to prevent unauthorised access to personal data, or the equipment used for processing. It contains a description of the anonymization/pseudonymization techniques that need to be implemented.

D7.3 GEN – Requirement No. 3 (5G-LOGINNOV, Due in February 2021)

An independent Ethics Advisor is appointed to monitor the ethics issues involved in the project and how they are handled. The advisor must be consulted at least on the real-time video surveillance by means of body-worn cameras, portable cameras, and drone surveillance.

The implementation of the privacy protection measures is described in D1.5 (5G-LOGINNOV, 2021). The main steps consist of:

- Ensuring that only necessary data requested for the data processing are collected and processed
- Identifying the lawfulness of the data processing (e.g., informed consent)
- Providing informed consent compliant to the GDPR (Articles 12, 13, 14)
- Ensuring that the third parties know and implement the GDPR
- Providing the mechanisms to support the rights of the data subjects
- Verifying that the security mechanisms are implemented

7   CONCLUSION

This deliverable constitutes the final update of the 5G-LOGINNOV project's data management plan. It describes the data categories handled over the course of the project which are:

- The technical data referring to the data related to the technical development and operation of the use cases.
- The evaluation data including the data used to compute the KPIs calculated for the evaluation tasks of the project.
- The open research data constituting the data and results that are published by the project to comply with ORDP requirements.
- The internal administrative data corresponding to the data generated and shared internally for administrative and management purposes.

This document highlights the data management measures to ensure data security with consideration of the ethical and privacy aspects. The data management also aims at making the project's data FAIR, with the ultimate result of the publication of some of the produced datasets on the Zenodo open data repository. It allows the project to contribute to the state of the art, encouraging comparison between methodologies with other approaches or projects.

8   REFERENCES

5G-LOGINNOV. (2021). *D1.5 Data and cyber-protection policies.*

5G-LOGINNOV. (2021). *D6.4 - Data Management Plan.*

5G-LOGINNOV. (2021). *D7.1 H - Requirement No. 1.*

5G-LOGINNOV. (2021). *D7.2 POPD - Requirement No.2.*

5G-LOGINNOV. (2022). *D1.4 Initial specification of evaluation and KPIs.*

5G-LOGINNOV. (2022). *D2.2 Data collection and evaluation procedures.*

European Commission Directorate-General for Research & Innovation. (2016). *Guidelines on FAIR Data Management in Horizon 2020.*

ANNEX 1: LIST OF PUBLISHED DATASETS

Name	SeaFront - Synthetic dataset for visual container inspection
DOI	10.5281/zenodo.10204550
Description	Applying Deep Learning techniques in a supervised manner generally requires a substantial amount of labelled data. However, such labelled data is not always readily available. To address this issue, one of the most common approaches is to synthetically generate the necessary data for training the models. In the context of shipping container analysis, an automatic synthetic image generation system for containers in a port scenario has been created. This system can reproduce various visual aspects of interest, including the container itself from all sides, diverse realistic backgrounds, potential damages the container might incur during shipment, IMDG stickers on the surface and text identification codes (BIC and ISO codes). As a result, a database with automatically labelled images is obtained. This dataset has primarily two different objectives: The first one is to assist researchers in training models capable of detecting the location of the container in the image and the location and typology of the different elements that may be on the container's surface. The second one is to serve as ground-truth in evaluation tasks. We are making this dataset publicly available, comprising almost 10000 images for training and validation, along with an additional 2480 images for testing. Our aim is to provide open and free data, which is often scarce in this field.
Standards and metadata	• DataCite Metadata Schema • Related paper: https://doi.org/10.1016/j.tre.2023.103174 • Readme file included in the dataset
File Format	Image/PNG
Data Sharing	• Open access • Creative Commons Attribution 4.0 International
Archiving and Preservation	Published on Zenodo: https://zenodo.org/communities/5g-loginnov

Name	KPIs to evaluate the benefits of 5G and AI-enabled services deployment on ports logistics operations
DOI	10.5281/zenodo.10204619
Description	This dataset includes KPIs measured during 3 trials conducted at the Athens Living Lab of the 5G-LOGINNOV project. The Athens Living Lab developed a set of use cases and platforms which communicate over a private 5G NSA network with different types of end devices (5G-Trucks, 5G-Cranes, 5G-IoT, 5G UEs). The living lab use cases focused on the deployment of 5G and AI-enabled services tailored to safety/security applications as well as for improving the efficiency of daily port operations i.e., reduce costs, improve the utilization of human resources and automate logistics services. The trials context is explained in detail in the project's deliverable D3.1 – Trial methodology, planning and coordination while the data structure is described in D2.2 – Data collection and evaluation procedures. Both of the deliverables are provided in the dataset as metadata.
Standards and metadata	• DataCite Metadata Schema • Deliverable D3.1 – Trial methodology, planning and coordination • Deliverable D2.2 – Data collection and evaluation procedures
File Format	Text/CSV
Data Sharing	• Open access • Creative Commons Attribution 4.0 International
Archiving and Preservation	Published on Zenodo: https://zenodo.org/communities/5g-loginnov

Name	KPIs to evaluate the environmental impact of using 5G in data exchange for traffic management outside the port and the hinterland
DOI	10.5281/zenodo.10204637
Description	This dataset includes KPIs measured during 3 trials conducted at the Hamburg Living Lab of the 5G-LOGINNOV project. The Hamburg Living Lab addressed the potential of leveraging positive environmental impact by using 5G in data exchange for traffic management outside the port and the hinterland. The living lab deployed a methodology to capture the effect of the traffic infrastructure on regional emissions, making them comparable (standardised) by quantifying such influences under defined status of congestion and other relevant factors (driver profile, vehicle profile, loading, etc.). The trials context is explained in detail in the project's deliverable D3.1 – Trial methodology, planning and coordination while the data structure is described in D2.2 – Data collection and evaluation procedures. Both deliverables are provided in the dataset as metadata.
Standards and metadata	• DataCite Metadata Schema • Deliverable D3.1 – Trial methodology, planning and coordination • Deliverable D2.2 – Data collection and evaluation procedures
File Format	Text/CSV
Data Sharing	• Open access • Creative Commons Attribution 4.0 International
Archiving and Preservation	Published on Zenodo: https://zenodo.org/communities/5g-loginnov