5GLOGINNOV

D6.4

Data Management Plan

www.5g-loginnov.eu

This project has received funding from the
European Union's Horizon 2020 research and innovation programme
under grant agreement No. 957400 (Innovation Action)

Work Package 6	Project management
Task 6.4	Data management plan
Authors	Djibrilla Amadou Kountche (AKKA) Mandimby Ranaivo Rakotondrovelona (AKKA) Xavier Valente (AKKA)
Dissemination Level	Public
Status	Final
Due date	28/02/2021
Document Date	26/02/2021
Version Number	1.0

Quality Control

	Name	Organisation	Date
Editor	Xavier VALENTE	AKKA	08/02/2021
Peer review 1	Peter Schmitting	ERTICO	10/02/2021
Peer review 2	Estíbaliz Loyo	VICOM	15/02/2021
Authorised by (Technical Coordinator)	Eusebiu Catana	ERTICO	24/02/2021
Authorised by (Quality Manager)	Mandimby Nirina Ranaivo Rakotondravelona	AKKA	25/02/2021
Submitted by (Project Coordinator)	Eusebiu Catana	ERTICO	26/02/2021

Legal Disclaimer

5G-LOGINNOV is co-funded by the European Commission, Horizon 2020 research and innovation programme under grant agreement No. 957400. The information and views set out in this deliverable are those of the author(s) and do not necessarily reflect the official opinion of the European Union. The information in this document is provided “as is”, and no guarantee or warranty is given that the information is fit for any specific purpose. Neither the European Union institutions and bodies nor any person acting on their behalf may be held responsible for the use which may be made of the information contained therein. The 5G-LOGINNOV Consortium members shall have no liability for damages of any kind including without limitation direct, special, indirect, or consequential damages that may result from the use of these materials subject to any liability which is mandatory due to applicable law.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

List of abbreviations and acronyms

Abbreviation	Meaning
5GPPP	The 5G 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 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. As such, the development of a Data Management Plan (DMP) is required and is embodied in this deliverable.

The 5G-LOGINNOV DMP focuses on the governance of data generated and collected during the project, considering the ethics rules and GDPR. It details what data the project will generate and what datasets will be made publicly accessible for research purposes. According to the project's description of work, the DMP will be issued in two versions. This deliverable corresponds to the first version.

First, this document gives an overview of the 5G-LOGINNOV data by providing a description of the four data categories considered in the project:

- Technical data: data handled during the technical development of the Living Labs and operations of the use cases.
- Evaluation data: data selected and collected during the trials for evaluation purposes.
- Open research data: data agreed to be published under the frame of ORDP.
- Internal administrative data: data related to the internal management of the project.

Then it introduces how the FAIR approach is implemented in 5G-LOGINNOV. It also addresses the allocation of resources to data management. Additionally, it offers insight on the data management per work packages. Finally, it addresses the ethical and privacy aspects while providing recommendations regarding data security.

1 INTRODUCTION

1.1 Project overview

5G-LOGINNOV will focus on seven 5G-PPP Thematics and support to the emergence of a European offer for new 5G core technologies in 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 relating to wireless connectivity and Core Network agility, 5G-LOGINNOV ports will not only significantly optimise their operations but also minimise their environmental footprint to the city and the disturbance to the local population. 5G-LOGINNOV will be a catalyst for market opportunities build 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 lifecycle of the data to be generated, collected and processed during the project. It outlines the approach to make 5G-LOGINNOV data FAIR by indicating what data will be generated, collected and processed, what standards will be applied, how the research data will be preserved and what parts of the datasets will be shared for the evaluation needs and to comply with the ORDP requirements. This document will address also the ethical and privacy aspects as well as some data security principles. This deliverable is a living document and will be updated according to the evolution of the project:

- The initial version D6.4 (**M06**, i.e. February 2021) outlines the data management plan according to the current development of the project. It identifies a first set of data categories that will be involved in the project and proposes the data management process that will be followed in the next developments.
- The updated version D6.5 (**M36**, i.e. October 2023) will provide the final description of the datasets, the process and the tools for managing open data for research purposes.

This deliverable serves as an entry point to grasp the project-wide data management approach in 5G-LOGINNOV. As such it provides a big picture of the 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, Due in April 2021) which focuses on the evaluation data requirements, the D1.5 – Data and cyber-protection policies (5G-LOGINNOV, Due in

[1] https://5g-loginnov.eu/

April 2021) which provides requirements regarding data handling and cybersecurity and the D2.2 – Data collection and evaluation procedures (5G-LOGINNOV, Due in October 2021) which describes the data collection tools. This relationship is illustrated on Figure 1. It is important to note that these deliverables will be available in the later stage of the project with D1.5 being confidential to the consortium.

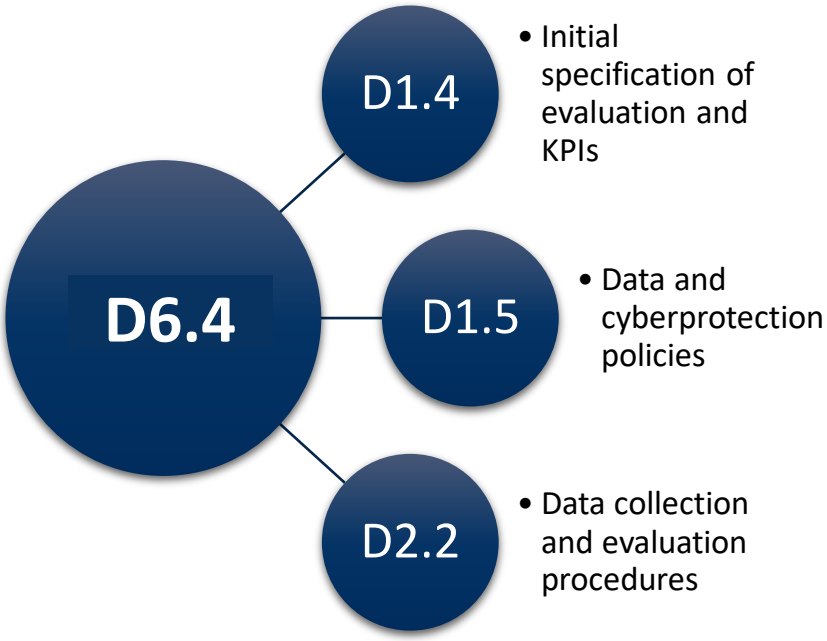

Figure 1. Relationship between D6.4 and data-related deliverables

Additionally, more information about GDPR and personal data will be given in the deliverable D1.5 – Data and cyber-protection policies (5G-LOGINNOV, Due in April 2021) and D7.2 – POPD - Requirement No. 2 (5G-LOGINNOV, Due in February 2021).

1.3 Intended audience

The dissemination level of D6.4 is 'public' (PU) 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 provides data management information to Living Lab coordinators and other technical profiles participating in WP1, WP2 and WP3.

2 5G-LOGINNOV DATA

2.1 Introduction

5G-LOGINNOV will handle different types of data which can be organised in four categories: *technical data*, *evaluation data*, *open research data* and *internal administrative data*. This chapter will describe these data categories while providing information on their management and sharing.

2.2 Data categories

This section introduces the categories of data that are handled in 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. 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. Description of 5G-LOGINNOV data categories

The data handling 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 in this phase, which falls in the technical data category, 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 a specific methodology (see D1.4 (5G-LOGINNOV, Due in April 2021)). It requires the collection of a subset of the data generated during the technical development phase (technical data) on which the evaluation is performed (evaluation data). Finally, the publication phase consists in the selection and the publication of a subset of the evaluation data as research data at the end of the project following the ORDP requirements. Only part of the evaluation data on which the partners agree to make openly available will be published and will constitute the open research data. To sum up, the open research data is a subset of the evaluation data which is itself a subset of the technical data.

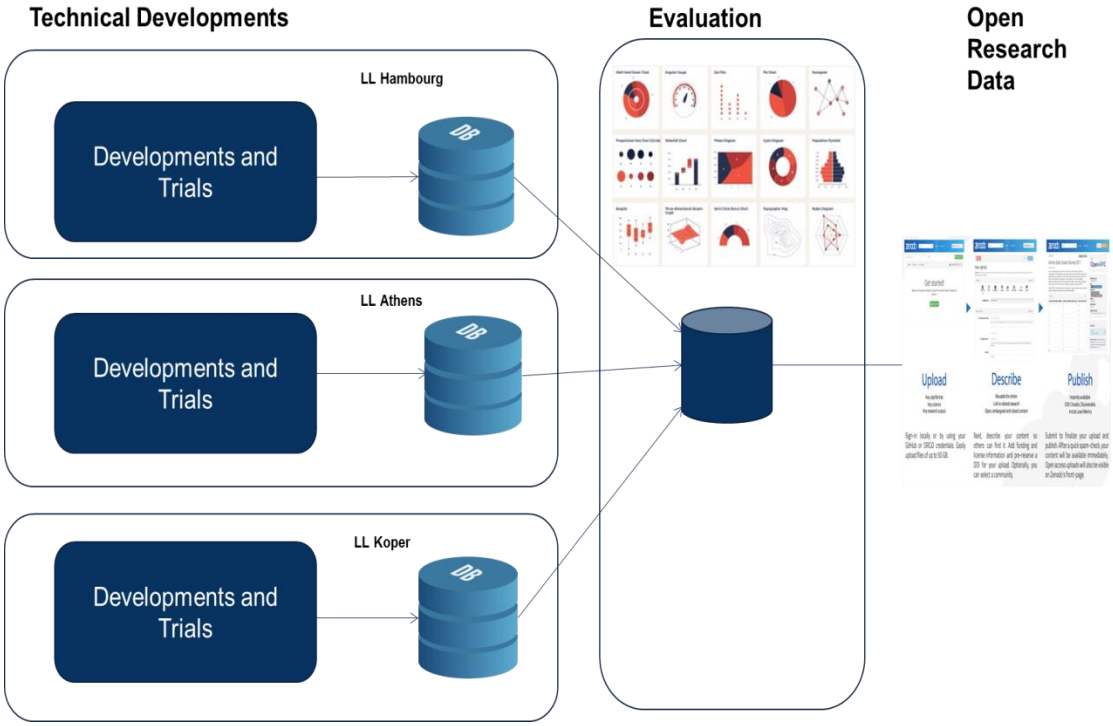

Figure 2. Data management in 5G-LOGINNOV

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 technical data may contain personal data. In this case, the GDPR rules apply (see 6.1). 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.

An early survey of the LLs data management architecture allowed the identification of the initial subcategories of technical data described on Table *2*. It will be updated over the course of the project, with a final list of subcategories provided in the second release of the Data Management Plan (in M36).

Categories	Sub-categories	Description
Driver	Background	Refers to a set of personal data collected throughout questionnaires or other means which might contain demographic, physical, cognitive, driver experiences and driving style data.
	Attitude	Includes data related to the driver's attitude towards driving and the 5G-LOGINNOV projects. This category can also contain other types of data collected by questionnaires.
Vehicle	Identification	These data are related to the vehicle identification such as the VIN.
	Sensors	5G-LOGINNOV trucks and drones embed many sensors to collect different types of data such as the position (GPS), temperature, Battery level,

		Fuel level and consumption, Oil level, Tire pressure, etc.
	Communication Units	These are the data related to characteristics of the OBUs such as the 5G, Wifi, ITS-G5, etc.
5G Technologies and their application for Logistics	Network Slicing	This sub-category contains data related to the Network Slicing related to its allocation, latencies, bandwidth, jitter, reliability and privacy. Other types of data are related to the tenant, the Network Slice Provider and the Network Slice Agent.
	MEC	This sub-category contains data related to the performances 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 measured in the Living Labs of Athens and Koper 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 component such as the Ultra Wide Band might also be logged.
	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.
Cloud Infrastructure and Services		A dedicated cloud infrastructure will be deployed for Koper to support the Use Cases 5 and 6. Data related to this infrastructure can be collected such as the server specifications, the latencies, the bandwidth, etc.
Ports	Environmental	For evaluation purposes, data related to the environmental impact of 5G-LOGINNOV will be collected using sensors deployed in the Ports.

Table 2. Illustration of Technical data

The technical data are produced/collected/shared during the use case operations allowing the execution of the trials. Thus, it is 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 LLs 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, Due in April 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, Due in October 2021).

The project's evaluation methodology is defined in Task 1.4 (Evaluation methodology and requirements) and depicted in the Figure 3. 5G-LOGINNOV Evaluation Methodology.

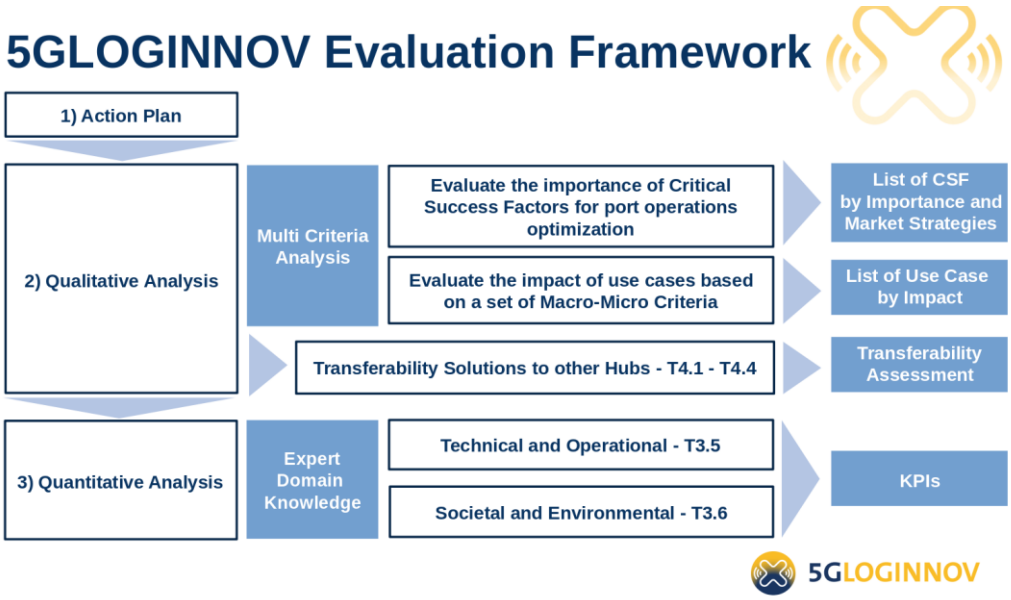

Figure 3. 5G-LOGINNOV Evaluation Methodology

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 5G-LOGINNOV Platform and of the Use Cases in the Living Lab. The Macro/Micro Criteria is a business strategic approach that break down the goals of the project (i.e. Macro Criteria) into measurable objectives (i.e. Micro Criteria) that can be then 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 Table *3* describes some of the KPIs that will be used for the evaluation. The detailed sets of data collected for each KPI per Living Lab will be provided in the next version of this deliverable (**M36**, i.e. October 2023).

Macro-criteria	Micro-Criteria	KPI	Description
Technical and	Provide accurate	Area traffic	The total traffic throughput served per

Operational	communications and recommendations for operations	capacity	geographic area.
		Coverage area probability	The area under consideration in which a service is provided by the mobile radio network to the end user in a quality (i.e., data rate, latency, packet loss rate) that is sufficient for the intended application.
		Reliability	The amount of sent packets successfully delivered to a given system node within the time constraint required by the targeted service, divided by the total number of sent network layer packets or percentage of packets properly received within the given maximum E2E latency.
	Increase number of ITC services	Availability	The amount of time a system is in condition to deliver services divided by the amount of time it is expected to deliver services in a specific area. The availability in percent is defined as the number of places (related to a predefined area unit or pixel size) where the Quality of Experience (QoE) level requested by the end-user is achieved divided by the total coverage area of a single radio cell or multi-cell area (equal to the total number of pixels) times 100.
		Bandwidth	The maximal aggregated total system bandwidth
		Coverage	The amount of Maximum Coupling Loss (MCL) in uplink (UL) and downlink (DL) between a UE and a TRxP (antenna connectors for a data rate of x bps). The coverage target should be >160 dB.
	Degree of centralization of data and information sources	OSM-VNF/Private 5G network - components onboarding and configuration	The amount of time required to complete the operation should be less than 15 min.
	Decrease traffic and incidents	Reduced density of vehicles in the yard	The density of vehicles in the case study area
Environemental	Decrease health risks for workers	Air quality	The amount of pollutants found types (e.g. NOx, HC, PM, etc.)
Societal	Improve quality of working environment	Improved yard vehicle operator safety	The amount of injuries

Table 3. A subset of 5G-LOGINNOV KPIs defined in Deliverable 1.4

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 one the subcategories of technical data described in 2.2.1 or include processed data like the KPI calculated during the evaluation.

Since the evaluation phase has not yet begun, it is not yet known what datasets will be publicly shared. This section will be enriched with such information in the second release of the DMP.

2.2.3.1 ORDP participation

5G-LOGINNOV will ensure the 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 scientific publications.The following lists the minimum 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".

More information on the ORD Pilot can be found in Annex 1 of this document.

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 Datasets description

2.3.1 General data 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.

As the nature and extent of these datasets can evolve during the project, updates of the description will be provided in the next version of the DMP.

The description of the different datasets should provide information on their reference, file format, standards, methodologies, metadata and repository to be used. More detail is given below.

2.3.2 Template used for 5G-LOGINNOV dataset description

Table 4. Template for dataset reference 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 will have 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 between the next 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 reference

2.4 Conclusion

This chapter proposes a summary of the data handled in 5G-LOGINNOV by providing a description of the corresponding data categories as well as information on the management of the data with regards to data collection and sharing. It also provides a guideline to describe the shared datasets.

3 ON MAKING 5G-LOGINNOV DATA FAIR

5G-LOGINNOV will use 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 will commit 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) will be used.
- All the datasets will have Digital Object Identifiers provided by a public repository
- The reference used for the dataset will follow a format like, for example: "5G-LOGINNOV_Living_Lab_UC_Datatype_XX" (XX: identifier to be added for similar datasets).
- The standards for metadata will be defined for each dataset as described in section 2.3.2 (standards and metadata).

Making data openly Accessible:

- The datasets for evaluation will be described using the recommendations of metadata description from the 'Standards and Metadata' entry in the Table *4*. Template for dataset reference.
- The data and their associated metadata will be made available either in a public repository or in an institutional repository. The Registry of Research Data Repositories[2] provides a useful listing of repositories that you can search to find a place of deposit. For example, ZENODO[3] has already been chosen for some EU H2020 projects and could be used in the case of the 5G-LOGINNOV project
- 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. Template for dataset reference will be used to provide information on the methods or software used to access the data.

Making data Interoperable:

- The metadata vocabularies, standards and methodologies will depend on the public repository and use the recommendations of metadata description from the Standards and Metadata entry in the Table 4. Template for dataset reference.
- The definition of the data formats will be made in the WP1 (Living Labs requirements & specifications) and WP2 (Living Labs development and deployment). The goal is to have the same format across Living Labs which will enable the development of common data quality check tools

[2] https://www.re3data.org/

[3] https://zenodo.org/

and enable partners responsible of the evaluation dealing with the same format across all pilot sites.

Making data re-usable (through clarifying licenses):

- All data producers will license their data to allow the widest reuse possible. Examples of licenses are available at https://opendefinition.org/licenses/.
- By default, the data will be 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.
- The data producers will make their data available for third parties within public repositories. The data will be reused for the scientific publication's validation purpose.

4 ALLOCATION OF RESOURCES TO DATA MANAGEMENT

The cost to make the data FAIR in 5G-LOGINNOV shall be handled by each partner who will have to generate its data according to the requirements expressed in the Data Management Plan.

In the project, Xavier VALENTE (AKKA) as Data Manager and Mandimby Nirina Ranaivo Rakotondravelona (AKKA) as Deputy Data Manager, will liaise with the TMT (Technical Management Team) about the data management issues. The Data Manager leads data management plan tasks and participates in the project coordination in terms of the evaluation data collection, storage and handling, as well as data publication as part of the ORDP.

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

5 DATA MANAGEMENT PER WORK PACKAGE

5.1 Introduction

This chapter addresses the data management approach taken in each work package. It provides insight on what kind of data is handled by the work packages. It is important to note that the data addressed in this chapter belongs to the categories defined earlier but are managed in an aggregated way (mostly as deliverables). This chapter is inspired from the deliverable D4.6 of AEOLIX (Xandra López, 2017). For each work package, the collected datasets are described using a table containing:

- Data set reference and name: Pointer to the deliverable containing the data
- Data set description: Description of the data that will be generated or collected
- Data collection procedures: Description of how the data was collected or produced
- Data privacy: Description of how data is kept confidential, if applicable
- Data sharing/confidentiality: Description of how data will be shared, e.g. publicly available or consortium internal only

5.2 WP1 Living Labs requirements and specifications

WP1	Living Labs requirements & specifications
Data set reference and name	5G-enabled logistics use cases
Data set description	Specifications for 5G-enabled advanced logistics hub and port operation use cases.
Data collection procedures	The specifications for each Living Lab are collectively defined by the concerned stakeholders.
Data privacy	Not applicable.
Data sharing/confidentiality	Publicly available as deliverable D1.1.
Data set reference and name	5G architecture and technologies for logistics use cases
Data set description	Requirements on the 5G architecture and technologies for logistics use cases.
Data collection procedures	The requirements are collectively specified by the involved partners.
Data privacy	Not applicable.
Data sharing/confidentiality	Publicly available as deliverable D1.2.
Data set reference and name	5G-enabled Living Labs infrastructure
Data set description	Requirements on the enhanced infrastructure and the components or devices using 5G technologies on the Living Labs.
Data collection procedures	The requirements are collectively specified by the involved partners.

Data privacy	Not applicable.
Data sharing/confidentiality	Publicly available as deliverable D1.3.
Data set reference and name	Initial specification of evaluation and KPIs
Data set description	Methodology and requirements, including KPIs for evaluating the impact of the 5G technologies and the new innovative devices and applications on the port operations and their use cases.
Data collection procedures	The requirements are collectively specified by the involved partners.
Data privacy	Not applicable.
Data sharing/confidentiality	Publicly available as deliverable D1.4.
Data set reference and name	Data and cyber-protection policies
Data set description	Requirements on the handling of the data to operate the use cases and the data collected for evaluation. Cyber-security policies to be applied.
Data collection procedures	The requirements are collectively specified by the involved partners.
Data privacy	Not applicable.
Data sharing/confidentiality	Available as deliverable D1.5 – confidential only for members of the consortium (including the Commission Services).

5.3 WP2 Living Labs development and deployment

WP2	**Living Labs development and deployment**
Data set reference and name	Development and deployment plan
Data set description	Process for the preparation, execution, and feedback cycles to be followed by all Living Labs.
Data collection procedures	The development and deployment plan are decided collectively by all the Living Labs. The data related to the development and deployment activities are monitored per Living Lab.
Data privacy	Not applicable.
Data sharing/confidentiality	Publicly available as deliverables D2.1 (intermediate version) and D2.3 (final version).
Data set reference and name	Data collection and evaluation procedures

Data set description	Information on the tools used for collecting and managing test data.
Data collection procedures	Not applicable.
Data privacy	The software source code will be open.
Data sharing/confidentiality	Textual description is publicly available as deliverable D2.2.

5.4 WP3 Living Labs trials and evaluation

WP3	**Living Labs trials and evaluation**
Data set reference and name	Trial methodology, planning and coordination
Data set description	Overall trial methodology including templates for test scenarios, test cases, questionnaires and description of the data collection process.
Data collection procedures	The trial methodology is collectively specified by the involved partners. The data collection process relies on T1.4 input.
Data privacy	Not applicable.
Data sharing/confidentiality	Publicly available as deliverable D3.1.
Data set reference and name	Living Labs trials preparation report
Data set description	Trial scenarios and test cases specifications per Living Lab.
Data collection procedures	The trial scenarios and the test cases are specified by each Living Lab using the methodology defined in T3.1.
Data privacy	Not applicable.
Data sharing/confidentiality	Publicly available as deliverable D3.2.
Data set reference and name	Evaluation of operation optimization
Data set description	Evaluation and assessment of LL data and LL trial results for operation optimization according to the tests scenarios defined in T3.2.
Data collection procedures	The results of evaluation are produced collectively by the involved partners.
Data privacy	GDPR must apply if any questionnaire answer will be available in the data set.
Data sharing/confidentiality	Publicly available as deliverable D3.3.
Data set reference and name	Evaluation of social, economic and environmental impacts

Data set description	Evaluation and assessment of LL data and LL trial results regarding the social, economic and environmental impacts.
Data collection procedures	The results of evaluation are produced collectively by the involved partners.
Data privacy	GDPR must apply if any questionnaire answer or interviews transcription will be available in the data set.
Data sharing/confidentiality	Publicly available as deliverable D3.4.

5.5 WP4 Marketplace and new actors

WP4	**Marketplace and new actors**
Data set reference and name	Plan for boosting marketplace and emergence of new actors
Data set description	Plan of activities to boost market and involve new actors and start-ups.
Data collection procedures	State of the art, use of questionnaires, market analysis.
Data privacy	GDPR must apply if any questionnaire answer or interviews transcription will be available in the data set.
Data sharing/confidentiality	Publicly available as deliverable D4.1.
Data set reference and name	Start-ups integration report
Data set description	Report on the process and results of the Open Call and the integration of the start-ups in the consortium.
Data collection procedures	Open Call management done collectively by the involved partners.
Data privacy	Not applicable.
Data sharing/confidentiality	Publicly available as deliverable D4.2.
Data set reference and name	Achievements with new actors and opportunities
Data set description	New business models for port maintenance and operations' efficiency, traffic management, 5G logistics corridors organizing CIDs (Collaborative Information Days).
Data collection procedures	Business models created collaboratively by the involved partners. Input from T4.2.
Data privacy	Not applicable.
Data sharing/confidentiality	Publicly available as deliverable D4.3.
Data set reference and name	Lesson learned and recommendations for stakeholders

Data set description	Recommendations to key stakeholders, to support the emergence of a European offer for new 5G core technologies enhancing next generation logistics hubs and ports in Europe and beyond.
Data collection procedures	Recommendations based on WP4 results with consideration of roadmaps for sustainable logistics.
Data privacy	Not applicable.
Data sharing/confidentiality	Publicly available as deliverable D4.4.

5.6 WP5 Dissemination and exploitation

WP5	**Dissemination and exploitation**
Data set reference and name	Communication and dissemination plan
Data set description	Communication and dissemination plans and materials (slides, flyers …).
Data collection procedures	Not applicable.
Data privacy	Not applicable.
Data sharing/confidentiality	The plans are publicly available as deliverable D5.1, D5.2 (initial version) and D5.3 (final version).
Data set reference and name	Exploitation plan and report
Data set description	Framework for the commercialization of 5G-LOGINNOV solutions.
Data collection procedures	Usage of Key Exploitable Results (KER), identification of the involved stakeholders, risk and transferability analysis.
Data privacy	Not applicable.
Data sharing/confidentiality	Publicly available as deliverable D5.4 (intermediate version) and D5.5 (final version).
Data set reference and name	Standardisation and spectrum policy report
Data set description	Recommendations towards interoperability between 5G-based service providers (with particular focus on innovative SMEs and start-ups) and other industry stakeholders (ICT vendors and MNOs).
Data collection procedures	Review of standards and other projects specifications.
Data privacy	Not applicable.
Data sharing/confidentiality	Publicly available as deliverable D5.6.

Data set reference and name	Clustering and networking results
Data set description	Clustering and networking activities especially with “sister” projects and European-wide associations and transnational initiatives with the related materials.
Data collection procedures	Information exchange between “sister” projects, joint dissemination initiatives.
Data privacy	Not applicable.
Data sharing/confidentiality	Results publicly available as deliverable D5.7. Only non-confidential materials are shared.

5.7 WP6 Project management

WP6	**Project management**
Data set reference and name	Project management plan
Data set description	Report on governance procedures and bodies, related meetings, management tools as well as the internal rules of procedures.
Data collection procedures	Produced by the management team.
Data privacy	Not applicable.
Data sharing/confidentiality	Publicly available as deliverable D6.1.
Data set reference and name	Innovation management plan
Data set description	Innovation plan and results.
Data collection procedures	Use of the CEN TS-16555 series of innovation management standards.
Data privacy	Not applicable.
Data sharing/confidentiality	Publicly available as deliverable D6.2 (plan) and D6.3 (report).
Data set reference and name	Data management plan
Data set description	List of 5G-LOGINNOV data sets including those that will be published for research purposes. Set of guidelines for data management, description of the tools to manage research data.
Data collection procedures	Produced by the data management team.
Data privacy	Not applicable.
Data sharing/confidentiality	Publicly available as deliverable D6.4 (initial version) and D6.5

	(final version).
Data set reference and name	Quality and risk management plan
Data set description	Set of guidelines and rules to efficiently manage risks and develop high quality reports and deliverables.
Data collection procedures	Produced by the management team.
Data privacy	Not applicable.
Data sharing/confidentiality	Publicly available as deliverable D6.6.

5.8 WP7 Ethics

Section 6 is dedicated to the ethical aspects of the 5G-LOGINNOV project.

5.9 Conclusion

This chapter describes the data that contributes to each of the work packages objectives and their management. The data belongs to the categories identified in the previous chapter even if it is presented in an aggregated way (mostly as deliverables).

6 ETHICAL AND PRIVACY ASPECTS

Ethics requirements are the subject of the Work Package 7: Ethics and will be thoroughly described in the following three deliverables.

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

This document will present the procedures and criteria that will 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) must be submitted as a deliverable.

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

This document will present a description of the security measures that will be implemented to prevent unauthorised access to personal data, or the equipment used for processing. It will contain description of the anonymization/pseudonymization techniques that will be implemented.

The beneficiary must evaluate the ethics risks related to the data processing activities of the project. This includes also an opinion if data protection impact assessment should be conducted under art.35 General Data Protection Regulation 2016/679.

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

An independent Ethics Advisor must be appointed to monitor the ethics issues involved in this 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. A report by the Ethics Advisor must be submitted as a deliverable at the end of each reporting period.

The dissemination level of the Deliverables from WP7 is confidential, only for members of the consortium including the Commission Services.

Some privacy aspects will be discussed in deliverable D1.5 – Data and Cyber-protection policies.

6.1 An approach to comply with GDPR in 5G-LOGINNOV

This section refers to the terminologies in the General Data Protection Regulation (GDPR: Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016) and provides an initial application of those terminologies to 5G-LOGINNOV project. The approach proposed to handle the GDPR in the 5G-LOGINNOV project consists in identifying all the concerned parties and identifies the actions they need to take in order to comply with the regulation. More specifically, this approach consists of six steps:

1. The first step consists in appointing a Data Protection Officer (DPO) and defining his mission in 5G-LOGINNOV.
2. During this step, a complete cartography of the personal data processing will be made.
3. Based on the results of the previous step, priorities need to be set about the personal data processing which may be riskier.
4. This step consists in conducting a Data Privacy Impact Assessment to evaluate the risks of the processing identified during the previous step.
5. During this step, the procedures are implemented for 5G-LOGINNOV in order to provide the protection of the personal data according to the GDPR.
6. This final step concerns the provision of a clear documentation of the whole data protection procedure that has been implemented.

Thus, this approach will enforce that personal data collected in 5G-LOGINNOV shall be (Article 5):

- processed lawfully, fairly and in a transparent manner in relation to individuals,
- collected for specified, explicit and legitimate purposes and further processed for scientific purposes,
- adequate, relevant and limited to what is necessary for the purposes for which they are processed,
- kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed,
- processed in a manner that ensures appropriate security of the personal data, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage.

Further refinement of the above-mentioned approach will be provided in next versions of this deliverable including for example a template to be filled by each data processor.

Since the project declares that the research involves personal data collection and/or processing, each data processor under the control of the data controller and the Data Protection Officer will declare how GDPR rules are fulfilled in a separate document.

7 DATA SECURITY

This chapter presents the overall security requirements to be fulfilled by the project and each Living Lab. A specific deliverable (D1.5 Data and cyber-protection policies, due on M08) will describe requirements and proposed tools and methods to ensure compliancy with GDPR and cyber protection for safe port operations. 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 for the whole project. Those data are made compliant with the GDPR as described in the Deliverable D7.2 (5G-LOGINNOV, 2021). This chapter describes some security principles that are implemented in order to protect against any type of modification. Also, a more thorough management of the servers could be made using the ISO 27001 standards, the BSI-Standards or the ENISA Good practices for Maritimes. The security principles are listed below.

Authentication: All the users wanting to get access to the 5G-LOGINNOV servers shall be authenticated. Also, proper means are used to authenticate the servers when necessary.

Authorization: The access to 5G-LOGINNOV data servers 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) shall be identified for each trial site and project wide to provide the authorization.

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 and IPsec.

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. Cryptographic mechanisms such as hash functions and digital signatures shall be used.

Availability: This security principle assures that the 5G-LOGINNOV servers shall be available for 5G-LOGINNOV users during the defined interval of service. Also, regular backups of the data should be made.

8 FINAL CONCLUSIONS

This deliverable provides an overview of the data handled in 5G-LOGINNOV. It describes the project's data categories and gives information on their management. D6.4 deliverable addresses the implementation of the FAIR principles in the project as well as the data management per work packages.

Since 5G-LOGINNOV participates in the Horizon 2020 Open Research Data Pilot, guidelines associated with 'open' access are described to ensure that the results of the project are openly available to the research community. With this document, the project commits to ensure the 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 scientific publications.

This document constitutes the first release of the Data Management Plan. In the final release, more details will be provided comparatively, specifically regarding the description of the shared datasets, the standards and methodologies, the sharing and storage methods.

It is important to note that the Data Management Plan is a living document and will be constantly updated until the end of the 5G-LOGINNOV project.

9 REFERENCES

[1] 5G-LOGINNOV. (2021). *D7.2 POPD - Requirement No.2.*

[2] 5G-LOGINNOV. (2021). *D7.3 GEN - Requirement No.3.*

[3] 5G-LOGINNOV. (Due in April 2021). *D1.4 Initial specification of evaluation and KPIs.*

[4] 5G-LOGINNOV. (Due in April 2021). *D1.5 Data and cyber-protection policies.*

[5] 5G-LOGINNOV. (Due in March 2021). *D7.1 H - Requirement No. 1.*

[6] 5G-LOGINNOV. (Due in October 2021). *D2.2 Data collection and evaluation procedures.*

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

[8] European Commission. (n.d.). *Horizon 2020 Online Manual*. Retrieved from European Commision Research and Innovation: https://ec.europa.eu/research/participants/docs/h2020-funding-guide/index_en.htm

[9] Xandra López, P. S. (2017). *Data Management Plan.* AEOLIX.

ANNEX 1: OPEN RESEARCH DATA PILOT (ORDP)

Open access refers to the online provision of scientific information that is free of charge to the end-user and reusable. This scientific information handles the peer-reviewed scientific research articles/publications and the research data underlying publications.

Under the H2020 programme, the project must also aim to deposit the research data needed to validate the results presented in the deposited scientific publications, known as “underlying data”. In order to effectively supply this data, projects need to consider at an early stage how they are going to manage and share the data they create or generate under H2020 guidelines on data management and with respect of 5G-LOGINNOV grant agreement.

> *The Commission is running a flexible pilot under Horizon 2020 called the Open Research Data Pilot (ORD pilot). The ORD pilot aims to improve and maximize 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, commercialization and Intellectual Property Rights (IPR), privacy concerns, security as well as data management and preservation questions.By extending the pilot, open access becomes the default setting for research data generated in Horizon 2020. (European Commission, s.d.)*

However, not all data can be open. Projects can therefore opt out at any stage (either before or after signing the grant) and so free themselves retroactively from the obligations associated with the conditions – if:

- Participation is incompatible with the obligation to protect results that can reasonably be expected to be commercially or industrially exploited
- Participation is incompatible with the need for confidentiality in connection with security issues
- Participation is incompatible with rules on protecting personal data
- Participation would mean that the project's main aim might not be achieved
- The project will not generate / collect any research data or
- There are other legitimate reasons (you can enter these in a free-text box at the proposal stage).”

After depositing publications beneficiaries must ensure open access to those publications via the chosen repository.

“The two main routes to open access are:

- Self-archiving / 'green' open access – the author, or a representative, archives (deposits) the published article or the final peer-reviewed manuscript in an online repository before, at the same time as, or after publication. Some publishers request that open access be granted only after an embargo period has elapsed.
- Open access publishing / 'gold' open access - an article is immediately published in open access mode. In this model, the payment of publication costs is shifted away from subscribing readers. “

In the research context, examples of data include statistics, results of experiments, measurements, observations resulting from fieldwork, survey results, interview recordings and images. The focus is on research data that is available in digital format and stored in a public repository. Normally, users can access, mine, exploit, reproduce, and disseminate openly accessible research data free of charge as explained in the following figure.

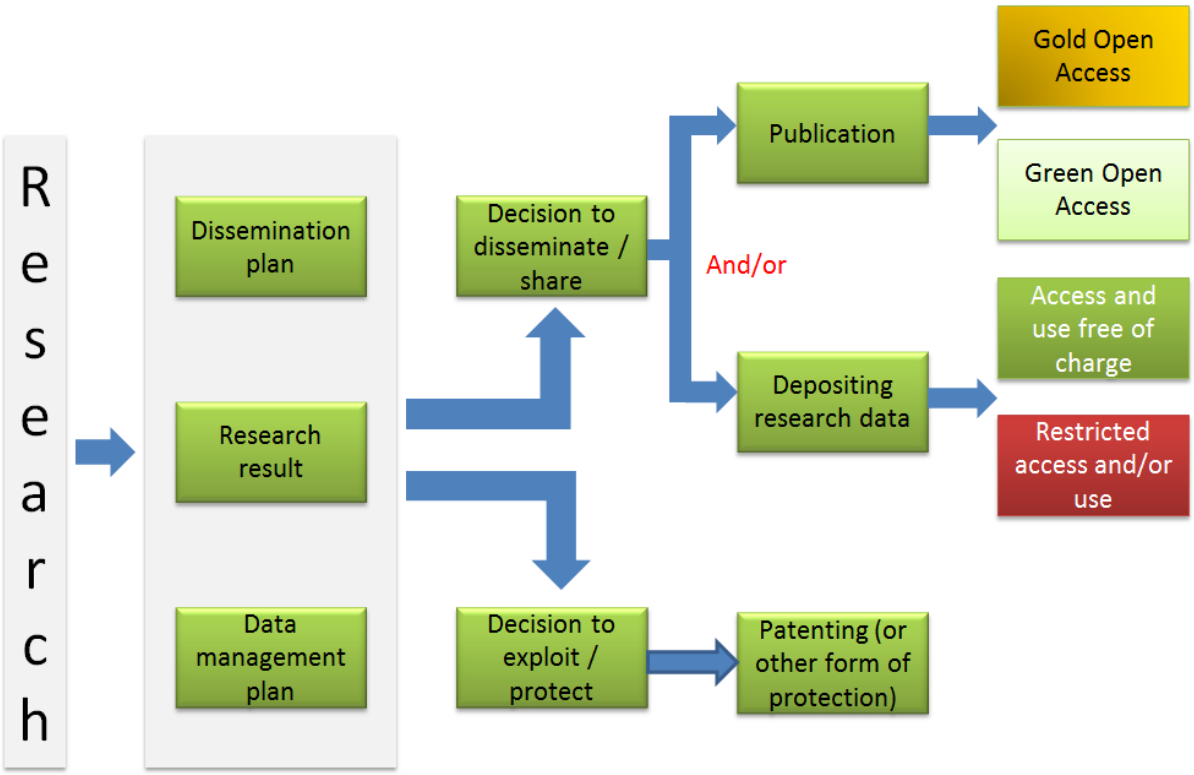

Figure 4. Principles of H2020 open access to research data