Integrating Network Digital Twinning into Future AI-based 6G Systems

D7.2

Data management plan

Document Information	
Grant Agreement N°	101136314
Authors and institutions	Inès Elisabeth KAHIN (R2M Solution) ; Sébastien FAYE (LIST)
Date	28 June 2024
Related WP	WP7 | Project Coordination
Dissemination level	PU | Public

6G SNS

D7.2 | Data Management Plan

Document change history

Version	Date	Author	Description
V0.1	27/05/2024	Inès Elisabeth KAHIN (R2M Solution)	First draft of the DMP
V0.2	21/06/2024	Sébastien FAYE (LIST)	Review of the DMP
V0.3	25/06/2924	All partners	Review of the DMP
V1	28/06/2924	Inès Elisabeth KAHIN (R2M Solution)	Final version of the DMP

Quality control

	Name	Organisation	Date
Editor:	Inès Elisabeth KAHIN	R2M Solution	28/06/2024
Peer review 1:	All partners		21/06/2024
Peer review 2:	Sébastien Faye	LIST	21/06/2024
Authorised and submitted by (Project Coordinator):	Sébastien FAYE	LIST	28/06/2024

Legal disclaimer

6G SNS

D7.2 | Data Management Plan

Executive Summary

This Data Management Plan (DMP) outlines how research data generated or collected during the course of the project will be managed, stored, and shared. It serves as a roadmap for handling data throughout the project lifecycle, ensuring compliance with Horizon Europe's requirements and promoting good data management practices in the 6G-TWIN project. It clarifies which data will be openly accessible and which will remain confidential within the consortium.

The project's data constitutes of a master dataset for RAN, CN and UE of specific scenarios to be used as a baseline in the WPs, a dataset for basic models, a dataset for functional models, federation interface reference implementations, federated simulation framework prototype, a use-case 1 dataset on teleoperated driving, a use-case 2 dataset on energy savings and data generated from the testbeds and demonstrators.

Following Horizon Europe guidelines, the report emphasises the importance of Open Access (OA), defined as the free online provision of reusable scientific information. The rationale behind OA is underscored, highlighting the benefits of openly sharing publicly funded research data with the research community, commercial sector, and civil society.

While certain collected data will be openly disseminated, certain work packages may generate internal-use data primarily intended for consortium management and communication.
Although most project datasets will be openly accessible, exceptions may arise due to privacy concerns, administrative burden, or limited utility to external users. Nevertheless, the project endeavours to openly disseminate its research findings, safeguarding against exploitable results or privacy infringements.

Specific disclaimer:
● The structure and the approach of this deliverable is common for data management deliverables written by R2M Solution in European projects and thus contains similar content as other Data Management Plan (DMP) deliverables.

D7.2 | Data Management Plan

Abbreviations and acronyms

Abbreviations and acronyms	
CC	Creative Commons
CN	Core Network
DMP	Data Management Plan
DS	Dataset
GA	Grant Agreement
NDT	Network Digital Twins
PII	Personal Identifiable Information
RAN	Radio Access Network
UE	User Equipment
URLLC	Ultra Reliable Low Latency Communications
WP	Work Package

D7.2 | Data Management Plan

Table of Contents

6GSNS

D7.2 | Data Management Plan

Table of tables

1. INTRODUCTION

1.1. Aims and objectives

6G-TWIN aims to develop an AI-native reference architecture for 6G systems, incorporating Network Digital Twins (NDT) for real-time optimisation and management of complex network scenarios. This involves creating and managing multi-layered virtual representations of 6G systems across various domains and technologies. The project's solutions will be tested in two use cases focusing on mobility and energy efficiency.

The Data Management Plan (DMP) for the 6G-TWIN project outlines the systematic organisation, storage, and dissemination of data generated or collected throughout the project's duration. It delineates the procedures for handling both open access and confidential data within the consortium, adhering to relevant regulations and guidelines. The DMP ensures transparency and accountability in data management practices, facilitating collaboration and knowledge sharing among project partners. It also addresses data security and privacy considerations, safeguarding sensitive information while promoting open access to research outcomes wherever feasible.

1.2. Intended audience

This report serves two primary audiences. Firstly, it is targeted internally towards the consortium partners' organisations involved in the project's implementation across Luxembourg, Belgium, Italy, France, Germany, Portugal, Belgium, Türkiye and Ireland. The Data Management Plan (DMP) aims to establish consistent data handling practices among these partners to enhance efficiency and reliability throughout the project's execution.

Secondly, the report is directed towards a broader audience interested in an AI-native reference architecture for 6G systems that incorporates Network Digital Twins (NDT) as a core mechanism for the end-to-end, real-time optimisation, management, and control of highly dynamic and complex network scenarios. The DMP will delineate standard data formats, meaningful metadata, and open repositories to facilitate data sharing and enable widespread replication of the project's outcomes.

1.3. Relation to other activities in the project

As part of Work Package 7, the Data Management Plan (DMP) has been developed under the coordination of R2M, with significant input from other project partners. Work Package 7 is directly involved in executing the DMP, with each partner responsible for their respective activities. The most crucial datasets for validating the project's impact and drawing conclusions in scientific publications are data generated from the testbeds and demonstrators (WP4), with a domino effect from WP1 and WP2 (datasets from models) and WP3 (interface

6G SNS

D7.2 | Data Management Plan

implementations, simulation framework prototype). The evaluation and use of these datasets for impact assessment will be done in WP5.

This deliverable D7.2 Data Management Plan is complementary to the deliverable D2.1 Data governance, privacy, and harmonization, which proposes and defines the necessary standards, architectures, and mechanisms required to manage data efficiently and securely in a federated environment, facilitating interoperability and compliance with regulations.

1.4. Report structure

The report begins by outlining its purpose, intended audiences and the process for ongoing development. Section 2 outlines the concepts of open access publishing and open access data in scientific research.

Related issues such as a classification for project datasets and an overview of copyright licensing for open access are also discussed. Each project dataset is then described in detail in Section 3, using a standardised template.

2. APPROACH TO DATA MANAGEMENT

2.1. Open access

Open access (OA) is understood as the free, online provision of re-useable scientific information to other users. There are many good reasons to make the data and findings from publicly funded research openly available to the research community, the commercial sector and civil society.

Data availability is therefore categorised at this stage in one of three ways:
- Open Data that is shared for re-use or that underpins a scientific publication.
- Consortium Confidential data that is accessible to all consortium partners and the Local Implementation Groups (under certain specific conditions described in the project Consortium Agreement) but retained within the consortium.
- Private Data that is maintained by an individual partner for his own purposes.

Not all project work packages will produce datasets that are intended for public dissemination; much of the data created and stored during the project is for internal management and communication within the consortium only. It is therefore likely that much of the data is categorised as Consortium level data. However, the project will seek to openly disseminate its research findings, except in cases where there are defined exploitable outcomes, privacy concerns or there will be a high administrative burden for a dataset or limited worth to other users. The two main aspects of this dissemination approach are open access to scientific publications and open access to research data. Each is considered in the following sections.

6GSNS

D7.2 | Data Management Plan

2.2. Open data

There are four main aspects of open data summarised in the acronym FAIR:

- Findable- data has a unique, persistent ID, located in a searchable resource, and documented with meaningful metadata.
- Accessible- data is readily and freely retrievable using common methods and protocols, metadata is accessible even if the data is not.
- Interoperable- data is presented in broadly recognised standard formats, vocabularies and languages.
- Re-useable- data has clear licences, and accurate meaningful metadata conforming to relevant community standards and identifying its content and provenance.

The data management plan establishes how this approach will be realised in practice with the initial plan presenting an overview and greater detail provided in the interim and final reports as the work packages proceed.

2.3. Data storage and sharing

The project has four main data storage and sharing facilities according to the type of data and its intended accessibility.

- Private data: Stored locally on partner storage systems, subject to institutional back up practices.
- Consortium data: Uploaded to the project collaboration platform Microsoft Teams, a cloud storage system for simple and secure access for all partners from a web browser. Data is maintained with regular offsite backups.
- Open data: Two facilities will be used during the project:
  - The project website, managed by R2M, will be the first point of contact for public dissemination. It will host project technical reports and other materials such as events listings, blog articles, images, videos, links to partner organisations and related projects.
  - Large, re-useable data sets will be deposited in an open data repository [e.g. Zenodo: https://zenodo.org/communities/6g-twin/] selected by the task leaders during the delivery of the relevant work packages.
  - OpenAIRE, to which all datasets related to the project will be linked.

2.4. Copyright licences

When material is widely shared, copyright licences protect the authors of work and grant specific rights to publishers and others to use this work. The European Commission encourages authors to retain their copyright whilst disseminating it as open access. For documents and media, the Creative Commons (CC, ref 5) license family provides legal tools to enable open access in these circumstances, with CC-BY (licensees may copy, distribute, display and perform the work and make derivative works and remixes based on it only if they

6GSNS

give the author or licensor the credits in the manner specified by these or by any reasonable other manner) and CC0 (freeing content globally without restrictions) enabling re-use by third parties. Similarly, for software and data, licenses such as the GNU and BSD license families are in widespread use.

Where research findings are published in a journal or other scientific outlet, there should be consideration of the copyright agreement with the publishers, which may involve an embargo period. Submission in the 6G-TWIN repository requires the author to agree to a non-exclusive distribution licence, and a Creative Commons licence may be added at this stage.

At this initial stage, it is not possible to define the copyright arrangement for each project dataset. The most appropriate licensing arrangements for each of the project datasets will be investigated as they are better characterised by their respective work packages. The final data management plan will be updated to that effect.

3.   DESCRIPTION OF PROJECT DATASETS

A dataset is a collection of related data, collected or created, including any existing data or third-party sources, which will be used or reused during and beyond the lifetime of the 6G-TWIN project. The description of the datasets below is based upon the Horizon Europe DMP template.

The template is a set of questions that each Data Manager responsible for a specific dataset should answer with a level of detail appropriate to the project. The DMP is intended to be a living document in which information can be made available on a finer level of granularity through updates as the implementation of the project progresses and when significant changes occur.

3.1.   Collected datasets

This first version of the Data Management Plan mostly addresses the most important data that will be handled in the 6G-TWIN project. A list of the most important datasets that have been identified as of June 2024 in 6G-TWIN is available below. It should be noted that this list is likely to evolve depending on the evolution of the various activities foreseen in the project.

The datasets are numbered according to their primary work package and task number, as laid out in the 6G-TWIN DoA.

Table 1: List of collected datasets

Dataset name	Task	Data contributors	Data manager
Master dataset for RAN, CN and UE of specific scenarios to be used transversely/as a baseline in the WPs	All tasks	All partners	LIST

Dataset for basic models	Task 2.2: Network topology and components modelling (basic models)	LIST, IMEC, POLIBA, UBOU, UBI, VIAVI	LIST
Dataset for functional models	Task 2.3: Network dynamics modelling (functional models)	UBOU, LIST, IMEC, POLIBA, VIAVI	UBOU
Federation interface reference implementations	Task 3.2: Develop a secure federated simulation environment with an abstract basis and well-defined interfaces	TUD, ACC, VIAVI	TUD
Federated simulation framework prototype	Task 3.2: Develop a secure federated simulation environment with an abstract basis and well-defined interfaces	TUD, ACC, VIAVI	TUD
Use-case 1 dataset on teleoperated driving	Task 4.2: Vertical application use case 1: Teleoperated driving	LIST, POLIBA, UBOU, ACC, PX, VIAVI	LIST
Use-case 2 dataset on energy savings	Task 4.3: Vertical application use case 2: Energy Savings	IMEC, VIAVI, ACC	IMEC
Data generated from the testbeds and demonstrators	Task 4.4: Collection and analytics of data generated from the testbeds and demonstrators	VIAVI, ACC, PX	VIAVI

3.2.   The form - template

Information about each dataset has been collated by Task Leaders in the format presented below.

Number and name of the dataset - TEMPLATE	
1. DATA SUMMARY	
Task & Data Manager	Work Package and Task numbers related to the dataset, and the Data Manager who takes responsibility?
Availability	Private, Consortium or Open, as defined in section 2.3

6GSNS

D7.2 | Data Management Plan

Purpose of the data	What is the purpose of the data collection/generation and its relation to the objectives of the project?
Origin of the data	What is the origin of the data? Produced in the framework of the project or reused?
Storage	Where will the data be stored? (e.g: open data repository like Zenodo, 6G-TWIN website for public deliverables, Microsoft Teams) as defined in section 2.3
Data nature and size	What types and formats of data will the project generate/collect? What is the expected size of the data?
Data utility	To whom might your data be useful, at Consortium level or external to the project?
Dataset Specific Metadata	Keyword(s) that categorise data to make it linked/searchable
2. FAIR DATA	
Making data findable	Will the data be accessible through a free and standardised access protocol? What information is needed for the data to be read and interpreted in the future? how this documentation and metadata will be captured / created? what metadata standards will be used and why?
Making data openly accessible	Will all data be made openly available? If certain datasets cannot be shared (or need to be shared under restricted access conditions), explain why, clearly separating legal and contractual reasons from intentional restrictions. Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if opening their data goes against their legitimate interests or other constraints as per the Grant Agreement.
Making data interoperable	In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?

6GSNS

D7.2 | Data Management Plan

Increase data reuse	How will you provide documentation needed to validate data analysis and facilitate data re-use (e.g. readme files with information on methodology, codebooks, data cleaning, analyses, variable definitions, units of measurement, etc.)? Will the provenance of the data be thoroughly documented using the appropriate standards?
3. ALLOCATION OF RESOURCES	
Costs	What are the costs for making data FAIR in your project? How will these be covered? Note that costs related to open access to research data are eligible as part of the Horizon Europe grant (if compliant with the Grant Agreement conditions).
Long term preservation	Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)?
4. DATA SECURITY	
Data security, recovery and storage	What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)? Is the data safely stored in certified repositories for long term preservation and curation?
5. ETHICAL ASPECTS	
Ethical and legal issues on data sharing	Are there any ethical or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review. How will you protect the identity of participants if required? e.g. via anonymisation
Informed consent	Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data?
6. OTHER ISSUES	
Other national/funder/sectorial/departmental procedures	Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones?

6GSNS

D7.2 | Data Management Plan

3.3. Master Dataset

Master dataset for AN, CN and UE of specific scenarios to be used transversely/as a baseline in the WPs	
1. DATA SUMMARY	
Task & Data Manager	This dataset will be transverse to most of the work package. It might be initiated in WP1, and then populated incrementally in WP2, WP3 and WP4. Data Manager: LIST
Availability	Consortium
Purpose of the data	The models and assets created in the project will receive input data from various access networks, core networks, and user equipment. This data may be collected or generated specifically for each task, as described below. However, to avoid replicating the same process multiple times, we may also have a baseline dataset. This dataset will be GDPR-compliant and will include real data, profiles, and artificially generated data, to be used across the various activities of the project.
Origin of the data	To be fully determined. The data will most probably be a mix of data collected in the reality, through a real network or the partners' emulators; or synthetically generated through the partners' tools or Generative AI techniques.
Storage	On a LIST server accessible to all the partners.
Data nature and size	The data would follow the Smart Data Model defined in Deliverable 2.1. The size will highly depend on the scenarios to be considered.
Data utility	At the consortium level to give a baseline to the partners needing data for model training.

6GSNS

D7.2 | Data Management Plan

Dataset Specific Metadata	Access network, core network, user equipment datasets.
2. FAIR DATA	
Making data findable	The data would follow the Smart Data Model defined in Deliverable 2.1.
Making data openly accessible	Most likely not.
Making data interoperable	The data will be interoperable. It would follow the Smart Data Model defined in Deliverable 2.1.
Increase data reuse	The data would follow the Smart Data Model defined in Deliverable 2.1.
3. ALLOCATION OF RESOURCES	
Costs	/
Long term preservation	This will be discussed at a later stage.
4. DATA SECURITY	
Data security, recovery and storage	The data will be stored in a private server, only accessible through a VPN connection.
5. ETHICAL ASPECTS	
Ethical and legal issues on data sharing	Not applicable.
Informed consent	Not applicable.
6. OTHER ISSUES	
Other national/funder/sectorial/departmental procedures	/

3.4. Data collection in WP2

2.2 Dataset for basic models	
1. DATA SUMMARY	
Task & Data Manager	WP2 – Task 2.2, Data manager: LIST
Availability	The basic model (data schema) will be available as open source (smart data model). Data will be private to the consortium members.

D7.2 | Data Management Plan

Purpose of the data	The data will be used to build the structure of a digital replicate of real network (digital twin) to perform the following operations: monitoring, simulation, prediction, enhancement.
Origin of the data	Produced in the framework of the project (emulation, simulation).
Storage	Smart data model will be available under open-source repository.
Data nature and size	The data will be structured as a typed graph (for the basic model) principally following the NSGI-LD format (Smart Data Model). Size < 1Go (data schema)
Data utility	Each member of the consortium that will handle data should follow the defined structure. External members that need to have a reference model for storing, sharing network related data.
Dataset Specific Metadata	Keyword(s) that categorise data to make it linked/searchable
2. FAIR DATA	
Making data findable	The basic model, the typed graph structure will be necessary to read and interpret the data. A documentation should also explain each entity, relation and attributes of the typed graph (at least under the form of a glossary). The data will be described in NSGI-LD following the standard defined by the Smart Data Model.
Making data openly accessible	Open access to the basic model (GitHub repository)
Making data interoperable	In the project we will define some bridges from and to relevant other format (e.g., DTDL from Microsoft) if needed.
Increase data reuse	/
3. ALLOCATION OF RESOURCES	
Costs	/

6GSNS

D7.2 | Data Management Plan

Long term preservation	Similar as other DS
4. DATA SECURITY	
Data security, recovery and storage	Similar as other DS
5. ETHICAL ASPECTS	
Ethical and legal issues on data sharing	The basic model defines only a data-schema and will not contain any sensitive information.
Informed consent	/
6. OTHER ISSUES	
Other national/funder/sectorial/departmental procedures	/

2.3 Dataset for functional models	
1. DATA SUMMARY	
Task & Data Manager	UBOU
Availability	Open
Purpose of the data	The datasets provided in this WP consist of pre-processed, cleaned and aggregated data from the raw datasets generated in 3 and 4.X. The delivered data is fed into the various AI/ML-based functional models that will be proposed in this WP.
Origin of the data	- Datasets generated in 6G-TWIN DSs 3, 1.4, 4.1, 4.2, and 4.4. - Can use open datasets/generated from other projects and frameworks.
Storage	Similar as other Datasets
Data nature and size	Nature of data: mainly tabular (.csv) but also other types stored in .txt files. Expected size: 1-10 GB.
Data utility	The data generated here will be used by the consortium and by external parties to develop data-driven functional models.

6GSNS

Dataset Specific Metadata	
2. FAIR DATA	
Making data findable	The datasets will be findable on the project website. Metadata for the datasets will be provided, including title, abstract, author, keywords, and details on when and how it was created.
Making data openly accessible	Will all data be made openly available? Yes
Making data interoperable	/
Increase data reuse	In addition to the metadata, the datasets will be accompanied by a detailed report on how they were generated, a README file with technical metadata, and a notebook (.ipynb) file containing some statistics and graphs on the data.
3. ALLOCATION OF RESOURCES	
Costs	Similar as other DS
Long term preservation	Similar as other DS
4. DATA SECURITY	
Data security, recovery and storage	Similar as other DS
5. ETHICAL ASPECTS	
Ethical and legal issues on data sharing	The dataset is preprocessed and does not contain private data.
Informed consent	/
6. OTHER ISSUES	
Other national/funder/sectorial/departmental procedures	/

3.5. Data collection in WP3

3.1 Federation interface reference implementations	
1. DATA SUMMARY	
Task & Data Manager	Task 3.2: Develop a secure federated simulation environment with an abstract basis and well-defined interfaces Data Manager: TUD
Availability	Open
Purpose of the data	KPI3.4: Provide a reference implementation of the federation interface as Open Source software.
Origin of the data	Produced in the project
Storage	Open Data Repository
Data nature and size	Source Code, ≤1GB
Data utility	Interested researchers and engineers both within and beyond the project
Dataset Specific Metadata	6G-TWIN, simulation
2. FAIR DATA	
Making data findable	Accessible through direct source code download, potentially also from version control.
Making data openly accessible	Openly available under a copyleft license.
Making data interoperable	Uses typical software build system.
Increase data reuse	Documented in-line where possible, using separate documentation files where required.
3. ALLOCATION OF RESOURCES	
Costs	None (hosted at open data repository)
Long term preservation	Handled by open data repository
4. DATA SECURITY	
Data security, recovery and storage	Handled by open data repository
5. ETHICAL ASPECTS	
Ethical and legal issues on data sharing	None
Informed consent	Not applicable
6. OTHER ISSUES	
Other national/funder/sectorial/departmental procedures	None

3.1 Federation interface reference implementations

3.2 Federated simulation framework prototype	
1. DATA SUMMARY	
Task & Data Manager	Task 3.2: Develop a secure federated simulation environment with an abstract basis and well-defined interfaces Data Manager: TUD
Availability	Open
Purpose of the data	KPI3.1: Design a platform-independent solution that allows the integration of new frameworks by implementing an abstract interface.
Origin of the data	Produced in the project
Storage	Open Data Repository
Data nature and size	Source Code, ≤10GB
Data utility	Interested researchers and engineers both within and beyond the project
Dataset Specific Metadata	6G-TWIN, simulation
2. FAIR DATA	
Making data findable	Accessible through direct source code download, potentially also from version control.
Making data openly accessible	Openly available under a copyleft license.
Making data interoperable	Uses typical software build system.
Increase data reuse	Documented in-line where possible, using separate documentation files where required.
3. ALLOCATION OF RESOURCES	
Costs	None (hosted at open data repository)
Long term preservation	Handled by open data repository
4. DATA SECURITY	
Data security, recovery and storage	Handled by open data repository
5. ETHICAL ASPECTS	
Ethical and legal issues on data sharing	None
Informed consent	Not applicable
6. OTHER ISSUES	
Other national/funder/sectorial/departmental procedures	None

3.6. Data collection in WP4

4.2 Use-case 1 dataset on teleoperated driving	
1. DATA SUMMARY	
Task & Data Manager	Task 4.2 Vertical application use case 1: Teleoperated Driving. Data manager: LIST
Availability	Open
Purpose of the data	The data are collected from the teleoperated driving testbed and may include activity of teleoperated driving agents, behaviour of the network infrastructure components and network performance KPIs achieved.
Origin of the data	Data is generated within the project from the teleoperated driving testbed.
Storage	Open Data Repository
Data nature and size	Formats will likely be CSV, JSON or similar. Size < 100GB
Data utility	Interested researchers and engineers both within and beyond the project. Particular interest expected for those conducting T4.4.
Dataset Specific Metadata	Digital twin, teleoperated driving, ultra-reliable, low latency, URLLC, critical system, simulation.
2. FAIR DATA	
Making data findable	The datasets will be findable on the project website. Associated information for the datasets will be provided, including title, abstract, author, keywords, and details on when and how it was created, such that it can be related to a particular experiment.
Making data openly accessible	The data will be made openly available unless prevented by the provisions of the Grant Agreement or other governing agreements.

D7.2 | Data Management Plan

Making data interoperable	Documentation of the various fields contained in the data will be provided. Mappings to research outputs where applicable and possible will be provided.
Increase data reuse	Documentation, descriptions of experimental steps, and other information that aids reuse will where possible be embedded with the data, either as documents and files, or as links to other artefacts.
3. ALLOCATION OF RESOURCES	
Costs	Data will be hosted at an open data repository to minimise or avoid cost.
Long term preservation	Handled by open data repository
4. DATA SECURITY	
Data security, recovery and storage	The data will be stored in a certified repository for long term preservation and curation.
5. ETHICAL ASPECTS	
Ethical and legal issues on data sharing	The data concerns entities in a testbed and will not contain any PII. No ethical or legal issues are anticipated on data sharing.
Informed consent	There is no PII and thus informed consent for data sharing and long term preservation is not required.
6. OTHER ISSUES	
Other national/funder/sectorial/departmental procedures	No

4.3 Use-case 2 dataset on energy savings	
1. DATA SUMMARY	
Task & Data Manager	Task 4.3 Vertical application use case 2: Energy Savings. Data manager: IMEC
Availability	Open
Purpose of the data	The data are collected from the energy saving testbed and may include activity of teleoperated driving agents, behaviour of the

6GSNS

	RAN and core network infrastructure components, network performance KPIs and energy consumption achieved.
Origin of the data	Data is generated within the project from the energy saving testbed.
Storage	Open Data Repository
Data nature and size	Formats will likely be CSV, JSON or similar. Size < 100GB
Data utility	Interested researchers and engineers both within and beyond the project. Particular interest expected for those conducting T4.4.
Dataset Specific Metadata	Digital twin, energy saving, low energy, RAN and core, simulation.
2. FAIR DATA	
Making data findable	The datasets will be findable on the project website. Associated information for the datasets will be provided, including title, abstract, author, keywords, and details on when and how it was created, such that it can be related to a particular experiment.
Making data openly accessible	The data will be made openly available unless prevented by the provisions of the Grant Agreement or other governing agreements.
Making data interoperable	Documentation of the various fields contained in the data will be provided. Mappings to research outputs where applicable and possible will be provided.
Increase data reuse	Documentation, descriptions of experimental steps, and other information that aids reuse will where possible be embedded with the data, either as documents and files, or as links to other artefacts.
3. ALLOCATION OF RESOURCES	
Costs	Data will be hosted at an open data repository to minimise or avoid cost.
Long term preservation	Handled by open data repository

6GSNS

D7.2 | Data Management Plan

4. DATA SECURITY	
Data security, recovery and storage	The data will be stored in a certified repository for long term preservation and curation.
5. ETHICAL ASPECTS	
Ethical and legal issues on data sharing	The data concerns entities in a testbed and will not contain any PII. No ethical or legal issues are anticipated on data sharing.
Informed consent	There is no PII and thus informed consent for data sharing and long term preservation is not required.
6. OTHER ISSUES	
Other national/funder/sectorial/departmental procedures	No

4.4 Data generated from the testbeds and demonstrators	
1. DATA SUMMARY	
Task & Data Manager	Task 4.4 Collection and analytics of data generated from the testbeds and demonstrators Data manager: VIAVI
Availability	Open
Purpose of the data	The data will be the output of analytics performed on the data generated in T4.2 and T4.3.
Origin of the data	Data is generated within the project and derived from the teleoperated driving and energy saving testbeds.
Storage	Open Data Repository
Data nature and size	Formats will likely be CSV, JSON or similar. Size < 1GB
Data utility	Interested researchers and engineers both within and beyond the project.
Dataset Specific Metadata	Digital twin, teleoperated driving, ultra-reliable, low latency, URLLC, critical system,

6GSNS

	energy saving, low energy, RAN and core, simulation.
2. FAIR DATA	
Making data findable	The datasets will be findable on the project website. Associated information for the datasets will be provided, including title, abstract, author, keywords, and details on when and how it was created, such that it can be related to a particular experiment.
Making data openly accessible	The data will be made openly available unless prevented by the provisions of the Grant Agreement or other governing agreements.
Making data interoperable	Documentation of the various fields contained in the data will be provided. Mappings to research outputs where applicable and possible will be provided.
Increase data reuse	Documentation, descriptions of experimental steps, and other information that aids reuse will where possible be embedded with the data, either as documents and files, or as links to other artefacts.
3. ALLOCATION OF RESOURCES	
Costs	Data will be hosted at an open data repository to minimise or avoid cost.
Long term preservation	Handled by open data repository
4. DATA SECURITY	
Data security, recovery and storage	The data will be stored in a certified repository for long term preservation and curation.
5. ETHICAL ASPECTS	
Ethical and legal issues on data sharing	The data concerns entities in a testbed and will not contain any PII. No ethical or legal issues are anticipated on data sharing.
Informed consent	There is no PII and thus informed consent for data sharing and long term preservation is not required.
6. OTHER ISSUES	

Other national/funder/sectorial/departmental procedures	No

4. CONCLUSIONS

The present document represents the first version of the 6G-TWIN Data Management Plan established in June 2024 (Month 6). It sets a benchmark to identify the actions that need to be implemented by the 6G-TWIN partners in order to fulfil the European Commission's requirements in terms of data management and accessibility of the research data.

As 6G-TWIN performs research on data models, datasets compiled in this document are subject to evolve in the upcoming months. This is why the DMP is a living document and will be updated and further developed during the project's lifetime. A first revision of the DMP is planned at M18.