



Project no.:	101095738		
Project full title:	6G SHort range extreme communication IN Entities		
Project Acronym:	6G-SHINE		
Project start date:	01/03/2023	Duration	30 months

D1.2 – DATA MANAGEMENT PLAN – FIRST VERSION

Due date	31/08/2023 (M6)	Delivery date	30/08/2023 (M6)
Work package	WP1		
Responsible author(s)	Berit Hvidberg Christensen (AAU)		
Contributor(s)	Enrico Maria Vitucci (CNIT) Baldomero Coll Perales (UMH) Gilberto Berardinelli (AAU) Ramoni Ojekunle Adeogun (AAU)		
Version	V1.0		
Reviewer(s)	Frank Burkhardt (Fraunhofer IIS)		
Dissemination level	Public (PU)		

VERSION AND AMENDMENT HISTORY

Version	Date (DD/MM/YYYY)	Created/Amended by	Changes
0.1	07/07/2023	Berit H. Christensen	Structure and basic content
0.2	17/07/2023	Berit H. Christensen	First complete draft based on COO input
0.3	02/08/2023	Berit H. Christensen	Updates based on contributors' input
0.8	16/08/2023	Berit H. Christensen	Updates based on reviewers' comments
0.9	23/08/2023	Berit H. Christensen	Minor refinements

TABLE OF CONTENTS

FIGURES	3
TABLES	3
ABBREVIATIONS AND ACRONYMS	3
EXECUTIVE SUMMARY	4
1 INTRODUCTION	5
1.1 GENERAL	5
1.2 THE DMP TEMPLATE	6
1.3 WHY IS THE DMP IMPORTANT?	6
1.4 WHO USES THE DMP?	6
1.5 FAIR PRINCIPLES AND OPEN RESEARCH DATA	6
2 DATA SUMMARY	7
2.1 IS EXISTING DATA BEING RE-USED?	7
2.2 TYPES AND FORMATS OF DATA GENERATED	7
2.3 PURPOSE OF THE DATA COLLECTION	9
2.4 EXPECTED SIZE OF THE DATA	9
2.5 ORIGIN OF THE DATA	10
2.6 DATA UTILITY	10
3 FAIR DATA	10
3.1 MAKING DATA FINDABLE, INCLUDING PROVISIONS FOR METADATA	10
3.2 MAKING DATA ACCESSIBLE	11
3.3 MAKING DATA INTEROPERABLE	13
3.4 INCREASE DATA RE-USE	14
4 ALLOCATION OF RESSOURCES	15
4.1 COST ESTIMATION AND RESOURCES FOR MAKING THE DATA FAIR	15
4.2 DATA MANAGEMENT RESPONSIBILITIES	15
4.3 COSTS AND POTENTIAL VALUE OF LONG-TERM DATA PRESERVATION	15
5 DATA SECURITY	15
6 ETHICAL ASPECTS	15
7 OTHER ISSUES	16
8 CONCLUSIONS	16
REFERENCES	17

FIGURES

Figure 1 The Creative Commons License Options	13
---	----

TABLES

Table 1 Data Management Plan Overview	5
Table 2 Summary of 6G-SHINE’s planned generated datasets	8
Table 3 Repositories used in the project.....	12

ABBREVIATIONS AND ACRONYMS

.mat	MAT-files are binary MATLAB® files that store workspace variables
.py	Python Software
ADAS	Advanced Driver-Assistance System
CARLA	Integrated High-Fidelity 3D Simulation Platform
CC-BY-NC	Creative Common Attribution-NonCommercial
CC 0	Creative Common Public Domain Dedication
CAVS	Connected and autonomous vehicles
DOI	Digital Object Identifier
DMP	Data Management Plan
EC	European Commission
EOSC	European Open Science Cloud
EU	European Union
GA	Grant Agreement
GNSS	Global Navigation Satellite System
HTTPS	Hypertext Transfer Protocol Secure
FAIR DATA	Making data findable, accessible, interoperable, and reusable
ISBN	International Standard Book Number
ISRC	International Standard Recording Code
IPR	Intellectual property rights
PoC	Proof of concept
LOS/NLOS	Line of Sight/Non Line of Sight
SQL	Structures Query Language
UTF-8	Unicode Transformation Format 8-bit
3GPP	The 3rd Generation Partnership Project
URL	Uniform Resource Locator
UTF-8	Unicode Transformation Format 8-bit
WP	Work Package

EXECUTIVE SUMMARY

The Data Management Plan (DMP) for the 6G-SHINE project identifies and categorises the different types of data that are expected to be collected, processed, generated, and disseminated in the project, namely: public deliverables, scientific publications, contributions to standards, datasets, Software Code etc.

Furthermore, it describes the methods to ensure the generation of FAIR data, making it findable, accessible, interoperable, and reusable. The DMP describes the life cycle of the data, how it will be handled during and after the end of project. In addition, it describes how the data will be selected for usage and presentation according to the project's data collection methods and standards, and finally how the data will be organised and stored.

The DMP for the 6G-SHINE project is a “living document”, which will be updated and enhanced as the project evolves as described in Work Package 1. (First, second and third versions) Any reissue of an updated DMP will be defined and agreed upon by the consortium.

In the current version of the DMP, a summary of the data that will be collected and generated from the project, including the purpose of data collection/generation as well as the origin, type, and format of data, is presented.

1 INTRODUCTION

1.1 GENERAL

Efficient data management planning is a part of research best practice, aiming at improving and maximising access to and the re-use of research data. The Horizon Europe Model Grant Agreement requires that a Data Management Plan (DMP) is established and regularly updated. This report constitutes the deliverable D1.2 Data Management Plan (DMP) produced under WP1, Task 1.4, of the 6G-SHINE project. It is developed in the sixth month after the start of the project as a first report on the subject, and it will be followed by subsequent editions, as described in the Grant Agreement.

The DMP for the 6G-SHINE project represents the collaborative efforts of the coordinator and the input and support of all the project partners. It includes detailed information on how data will be collected, processed, and generated in the Work Packages (WP) during the 30 months of the project’s lifespan. The DMP follows the template provided by the European Commission for the Horizon Europe project [1]. In short, the DMP defines:

- the handling of research data during and after the end of the project,
- what data will be collected, processed and/or generated,
- which methodology and standards will be applied,
- whether data will be shared/made open access.
- how data will be curated and preserved (including after the end of the project).

As mentioned, the DMP will evolve during the lifetime of the project as new data and methods will emerge. The DMP will be reviewed and updated two times during the lifetime of the project and will reflect significant changes in policies, organisation, or other influencing factors that occur to the consortium and/or the project.

Table 1: Data Management Plan Overview

Deliverable No.	Deliverable Name Work	Due Date (month)
D1.2	Data Management Plan first version	M6
D1.3	Data Management Plan second version	M12
D1.4	Data Management Plan third version	M30

1.2 THE DMP TEMPLATE

The European Commission has issued a [template](#) to assist with the completion of the DMP. The template contains a set of questions to be answered according to the project. These questions have helped to form the structure for the following chapters, where the main sections and proposed content of the template are listed, including 6G-SHINE's responses.

1.3 WHY IS THE DMP IMPORTANT?

The DMP provides an overview of the research data generated by the project, the types, and formats of this data, and how this data is processed and stored to make them findable, accessible, interoperable and reusable, according to the principles of FAIR data management. The purpose of the DMP is to guide and ensure the appropriate data management throughout the project's lifetime. The DMP also includes:

- Any commercial and security issues of the data, the definition of the project data that will be shared/made for open access and information on the project's IPR,
- The definition of the repository to handle the data during and after the end of the project,
- A description of the methodology used to handle the data collected, processed and/or generated.

1.4 WHO USES THE DMP?

All project participants who are involved with data collection, processing and storage must respect the procedures for data handling and the privacy notice. All other project participants may consult this document to acquire an overview of the category and of the usage of data in the project.

1.5 FAIR PRINCIPLES AND OPEN RESEARCH DATA

6G-SHINE will follow the FAIR guidelines for open research data to make its data Findable, Accessible, Interoperable and Reusable. These guidelines concern the data handling, evaluation, sharing and reuse of the manually and automatically generated input.

FAIR Data Management provides for:

- Summary of the data collected.
- Methods to ensure that the data is FAIR.
- Resources to be allocated.
- Data security.

These principles precede the implementation choices and do not necessarily suggest any specific technology, standard, or solution for implementation.

2 DATA SUMMARY

2.1 IS EXISTING DATA BEING RE-USED?

Data from other projects and existing work done previously by the project partners might be re-used at their discretion and permission.

2.2 TYPES AND FORMATS OF DATA GENERATED

A preliminary overview of the data types expected to be collected and generated in the 6G-SHINE project is presented below and listed in Table 1:

- WP1/2/3/4/5/6: Public deliverables
- WP2/3/4/5: Scientific publications (Mainly scientific papers created by the consortium members)
- WP3/5: datasets (Training and testing algorithms, including AI solutions, Proof-of-Concepts)
- WP2/4: Software Code (Open-source software, released in public repositories)
- WP2/3/4/5: Contributions to standard
- WP1/2/3/4/5/6: other publications (promotion material, press releases, white papers, etc.)

According to the project's scope, the data types are labelled following the expected method of usage:

- O - Open; entries which are candidates for open research data.
- L – Limited; parts of data or post-processed versions shall be evaluated to see if it can be published as open research data, while the complete, unprocessed data are expected not be made open.
- C – Closed; no part of this data is expected to be made accessible as open research data:

Datasets listed as 'Limited' or 'Closed' are restricted because making those openly available has been flagged as being against the beneficiaries' legitimate interests, as industrial partners expect commercial exploitation from these data.

PUBLIC DELIVERABLES

All project deliverables will be listed on our website (www.6gshine.eu) and be available for download after the submission to EC. Deliverables will present the project advances in the multiple technical fields. Deliverables will be provided in PDF format.

SCIENTIFIC PUBLICATIONS

The scientific publications generated by the consortium partners will be made available to the public and the scientific community in open source, also using well-known open-science repositories such as [arXiv](https://arxiv.org/) or Zenodo. (Table 3, repositories used in the project) The generated publications will also be listed on the [6G-SHINE website](http://www.6gshine.eu).

CONTRIBUTIONS TO STANDARDS

6G-SHINE aims at pioneering the main technology components for short-range communications in in-X subnetworks. Given the timeline of the project, which is expected to end earlier than the beginning of 6G standardization, the focus will be on pre-standardization. As described in deliverable D6.1 "Dissemination and communication plan" [2], the project aims at contributing to the activities in major

international standardization bodies, as well as to major industry fora. Potential contributions to pre-standardization are also presented in D6.1. Once a contribution to standardization is in preparation, this will be discussed within the consortium. Its potential availability in open access will also be discussed among the consortium members.

OTHER PUBLICATIONS

6G-SHINE targets to reach a broad audience via the generation of different types of publications besides scientific publications. Those include flyers, posters, press releases, whitepapers, and information articles in national newspapers. There will be information on the website and on 6G-SHINE social media channels about the release of these publications. Non-confidential publications will also be uploaded on the project website or linked to the websites where the publications are made available.

DATASETS AND SOFTWARE CODE

The datasets and Software Code listed below will be made available in open access during the project.

Table 2 Summary of 6G-SHINE’s planned generated datasets

WP (Partner)	Name of Dataset	Short description and format	Expected file size	Label
WP2/ Task 2.1/ Subtask 2.1b (UMH)	In-vehicle data traffic traces	Realistic data sets of raw and processed data traffic of in-vehicle networks that accurately represent the in-vehicle data traffic, especially the one related to the ADAS (Advanced Driver-Assistance System) domain of connected and autonomous vehicles (CAVs). The data sets differentiate the dataflow generated by different in-vehicle mounted sensors (cameras, radar, lidar, IMU, GNSS, etc.) and control functions by means of realistic traffic patterns (packet rate and size, traffic bursts, time correlation, etc.). The data sets are generated using an integrated high-fidelity 3D simulation platform (CARLA) with an automated driving stack (Autoware) available at UMH. (Software Code: Open-source software)	12,5 MB	O
WP2 Task 2.2. (CNIT)	Simulated and measured radio propagation data	The design of the 6G-SHINE radio system requires a deep understanding of the radio propagation characteristics in the scenarios where the subnetworks are anticipated to operate. The radio propagation characteristics could be affected by a high clutter density, obstructed line-of-sight (e.g., engine in a vehicle or robotic parts in industrial environments affecting the LOS/NLOS conditions) and from being enclosed by a metal housing. The dataset consists of digital maps of the studied environments (in the form of point clouds or 3D vector data) and Matlab code to post-process measured and simulated data. Simulated data are generated with ray-tracing prediction tools and made available in the form of tables and/or Matlab structures. The post-processing code is used to produce propagation parameters, such as Path Loss, Delay Spread, Angle dispersion, cross-polarization ratios, etc., that can be used for modelling purposes and for the design of PHY and MAC enablers to achieve the desired service requirements. Also, basic Matlab implementations of macroscopic models for Reconfigurable	Between 1 GB and 10 GB	O

WP (Partner)	Name of Dataset	Short description and format	Expected file size	Label
		Intelligent Surfaces (RIS), that can be used to assess realistic performance of RIS and integrated into ray tracing simulations, are made available in the dataset (Software Code: Open-source software)		
WP4/ Task 4.1 (AAU)	Simulation code for radio resource management in hyperdense subnetwork scenarios.	The simulation code for mobile subnetwork scenarios is developed to model and perform the system-level evaluation of solutions in dense deployments of mobile subnetworks. The code base includes functionalities for creating deployments, simulating subnetwork mobility, modelling propagation, integrating algorithms for radio resource management and performing system level analysis. (Software Code: Open-source software)	Around 100 MB.	O
WP4/ Task 4.1 (AAU)	Synthetic data traces for interference in dense mobile subnetwork scenarios	Simulated data sets of interference power in mobile subnetworks with different radio resource management techniques. The datasets capture the statistics of interference power in frequency, time, and spatial dimensions. The frequency dimension corresponds to different frequency sub-bands, and we capture the spatial dimension with the location of subnetworks inside simulated factory halls. The collected data sets contain time traces of aggregated interference power measured over each frequency sub-band at all subnetwork locations with the following dynamic sub-band allocations: random (fixed) sub-band selection, greedy channel selection, and centralised graph colouring. The data sets are generated using the in-X subnetworks simulator developed at AAU. The simulator implements functionality for system-level evaluation of mobile subnetworks and utilises channel models defined for indoor factory environments by 3GPP. (Software Code: Open-source software)	Between 10 GB and 20 GB.	O

2.3 PURPOSE OF THE DATA COLLECTION/GENERATION AND RELATION TO THE OBJECTIVES OF THE PROJECT

The overall objective of the project is to pioneer the main technology components for in-X wireless subnetworks, short range low power radio cells to be installed in a wide set of vertical and consumer entities like robots, vehicles, production modules, and classrooms for the sake of supporting extreme communication requirements in terms of latency, reliability, or data rates. With a TRL 2-4, 6G-SHINE will leverage the opportunities offered by the peculiar deployment characteristics of such short-range subnetworks for a highly performant yet cost-efficient radio design that allows bringing wireless connectivity to a level of pervasiveness which has never been experienced earlier.

2.4 EXPECTED SIZE OF THE DATA

The expected volume of the data is presently unknown, but it is foreseen that it will stay small enough to be stored on the standard available servers used in the project. There is no indication of a need to budget for repository storage outside of the scope of what repositories accept within their free limit quotas.

2.5 ORIGIN OF THE DATA

The 6G-SHINE partners' research activities will generate the data. Partners may contribute additional data to the project. This data will be handled in accordance with agreements of usage (Consortium Agreement)[3] in the project and will also, as far as possible, be described in Table 1.

2.6 DATA UTILITY

The data will be useful for subsequent research carried out by the partners, benefit the scientific community, and act as a basis for any subsequent projects.

3 FAIR DATA

3.1 MAKING DATA FINDABLE, INCLUDING PROVISIONS FOR METADATA

Identification by a persistent identifier

All data deposited in ArXiv, Zenodo or other relevant repositories will be identified with a Digital Object Identifier (DOI) whenever possible to increase findability and citability.

Standards for metadata

The project will adhere to industry standards and will respect GA provisions regarding metadata: "Metadata of deposited data must be open under a Creative Commons Public Domain Dedication (CC 0) or equivalent (to the extent legitimate interests or constraints are safeguarded), in line with the FAIR principles (in particular machine-actionable) and provide information at least about the following: datasets (description, date of deposit, author(s), venue and embargo); Horizon Europe funding; grant project name, acronym and number; licensing terms; persistent identifiers for the dataset, the authors involved in the action, and, if possible, for their organisations and the grant. Where applicable, the metadata must include persistent identifiers for related publications and other research outputs."

Each dataset has metadata associated with it, for descriptive, structural, and administrative issues. The embedded metadata in data will be part of the data collection and documentation process, which are elaborated before the publication of the metadata, both in terms of sanitising and for adding metadata. Some metadata will be created during the publication procedure for a given repository and follow the (metadata) standard(s), for the specific repository. As for the specific metadata format, Zenodo utilises Data Cite [4], a de facto standard for describing datasets. All partners are requested to include the following acknowledgement of funding in their metadata, which is supported by Data Cite in the metadata standard in the grant section, linking to the grant number 101095738, or mention the name of the project and grant agreement in case of metadata standards that do not have a separate grant section.

Search keywords

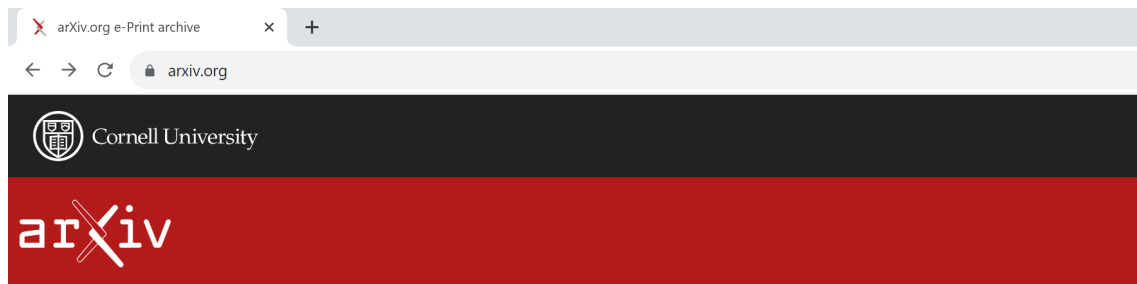
For generic repositories, the search keywords will be extracted from the metadata and provided as part of the descriptive metadata for the dataset. If subject-specific repositories are used, the search keywords will follow the standard utilised by the given repository.

Metadata harvest and indexation

Metadata harvest and indexation will be available depending on the repositories' protocols. As for Zenodo, this is harvested into the most common search engines for research data like OpenAIRE [5].

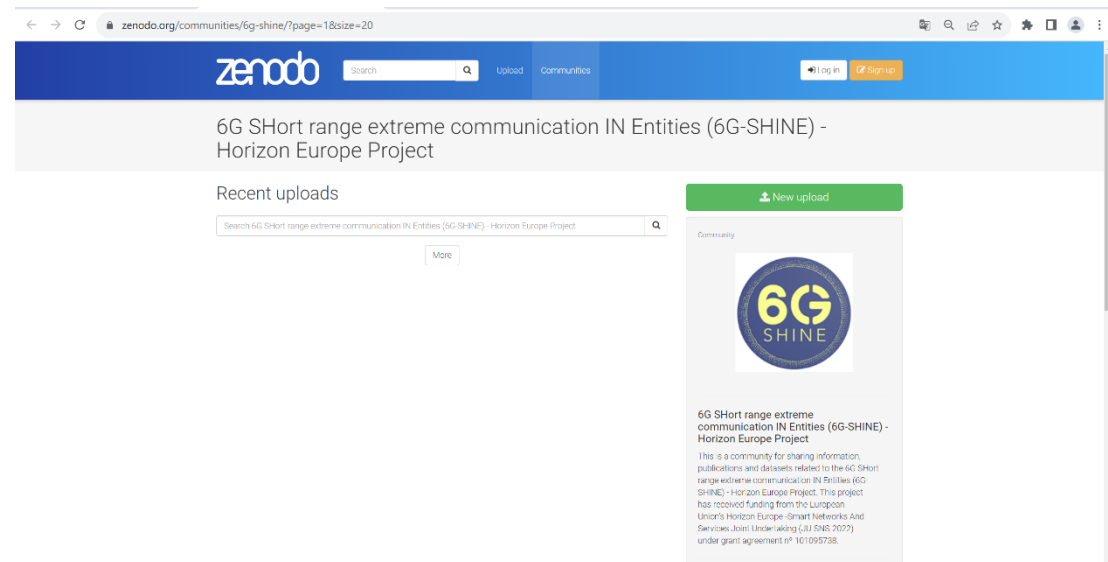
3.2 MAKING DATA ACCESSIBLE

A repositories community has been created in [ArXiv](#) and [Zenodo](#) all partners have been advised to use ArXiv, Zenodo or their institutional repository.



arXiv is a free distribution service and an open-access archive for 2,315,529 scholarly articles in the fields of physics, mathematics, computer science, quantitative biology, quantitative finance, statistics, electrical engineering and systems science, and economics. Materials on this site are not peer-reviewed by arXiv.

Subject search and browse:
Physics



Datasets deposited in Zenodo are provided with a DOI. Some of the institutional repositories that will be used, e.g., VBN [6], can provide a DOI upon demand. If this option is unavailable, partners are instructed to deposit in Zenodo.

When restricted data is stored on internal servers, partners are expected to include a note on Zenodo regarding the existence of the dataset. Access will be decided by project partners on a case-to-case basis. Partners will strive to describe a procedure for screening requests. This will follow the CA and GA. Repositories partners are planning to use (non-exhaustive list):

Table 3 Repositories used in the project.

Partner	Repository
All	Zenodo are recommended by default 6G-SHINE Zenodo Community . For scientific publications arXiv is also widely used.
AAU	As producer of the code, AAU will take responsibility for the storage and preservation of the code at VBN, Aalborg University's Research Portal GitHub (for codes). The general project information, relevant publications, and datasets will be presented on VBN, Aalborg University's Research Portal . The code for model description and some further Python scripts will be published on GitHub public repositories and possibly harvested to Zenodo to make the code citable with a DOI.
UMH	UMH, as a producer of the datasets, will take responsibility for the storage and preservation of the traces. DOI will be assigned, and data traces will be available on GitHub.
CNIT	CNIT, as a producer of the datasets, will take responsibility for the storage and preservation of the traces. DOI will be automatically assigned to each data set via Zenodo or Figshare [7]. Data traces will also be available on European Open Science Cloud (EOSC). This ensures open access availability of all data for at least ten years and creates permalinks for all datasets, which can be used for referencing in publications.

Data

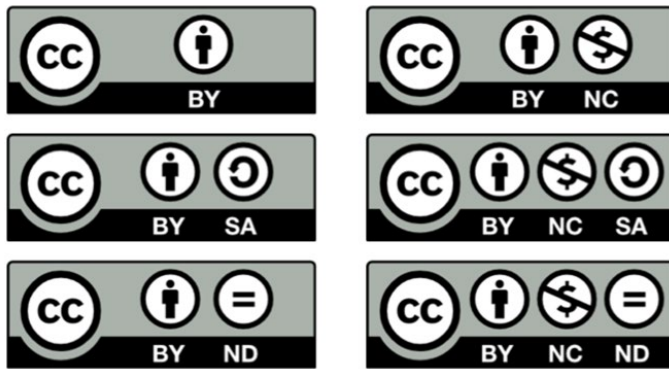
Data(sets) which can be made publicly available are identified in **Table 2** (datasets marked as O – Open). The data will be available in various file formats, with some being open, whilst others will be proprietary file formats due to the software used during the collection and analysis phase and to avoid the possible loss of important information if data is converted. The protocol for access will be determined by the repository used, most often a HTTPS protocol.

For the project members, a common project platform is available in Teams, where the different datasets to be shared can be stored for internal use. Further, data only used by individual partners will be stored on their platforms and according to their internal rules.

Public data will be deposited in Zenodo or other relevant repositories depending on each partner's preferences. There will not be an active data curation procedure for the specific repository. Still, data should remain available and findable for five years after the end of the project (the Zenodo website mentions that the data will be available for at least 20 years). A need for a data access committee, e.g., to evaluate/approve access requests to personal and sensitive data, has not been identified.

In the 6G- SHINE options, CC BY-NC is chosen. This license allows re-users to distribute, remix, adapt, and build upon the material in any medium or format for non-commercial purposes only, and only so long as attribution is given to the creator.

Figure 1 The Creative Commons License Options



CC BY-NC includes the following elements:

BY ⓘ Credit must be given to the creator

NC Ⓢ Only non-commercial uses of the work are permitted.

SA ⓘ (Share Alike) Adaptations must be shared under the same terms.

ND ⓘ (No Derivatives) No derivatives or adaptations of the work are permitted.

@ <https://creativecommons.org/about/ccllicenses/>

Metadata

Whenever possible, metadata will be made openly available and licenced under a CC 0, as per the Grant Agreement. Metadata for datasets listed as limited or closed are also CC 0, but because making those data openly available has been flagged as being against a beneficiary’s legitimate interests, the access to data will be restricted. The metadata will include information about software needed to access or read the data, and, when possible, a link to the relevant software will be provided.

3.3 MAKING DATA INTEROPERABLE

Version control

Each data owner is responsible for deciding on the file naming conventions and including a clear version-control guideline. Version control will follow the file naming conventions that include a version control in the file name.

Standards to be followed

All openly accessible data will be uploaded in a commonly accessible format and encoding (such as UTF-8).

The file formats are listed in Table 1. We are aware that some formats are considered proprietary. Where possible and feasible, we will convert files to non-proprietary formats. However, some data are best represented in the proprietary file formats, as some capabilities of the formats are not matched by their open alternatives.

Common vocabulary for data types

Data will be collected from many different partners; therefore, data types might differ depending on the provider. Therefore, the partners of the 6G-SHINE project intend to give the data relevant and representative names and use shared codes for representation in data, which clearly explain the content of the data type, including, e.g., the actual units in which they are measured, if relevant. This way, the partners can easily compare relevant data of the same type. The practice for this is yet to be determined when writing the DMP.

Qualified references to other data

Where possible, the documentation of the data will include descriptions with references to other datasets. Furthermore, if needed, the metadata in Zenodo will relate to other datasets, e.g., pointing to related datasets, papers etc.

3.4 INCREASE DATA RE-USE

Documentation is needed to validate data analysis and facilitate data reuse

Partners are encouraged to provide documents such as, e.g., readme files with information on methodology, codebooks, data cleaning, analyses, variable definitions, and units of measurement, whenever possible. The data and publications will be interlinked using metadata in Zenodo or other repositories where possible.

How the data will be licenced to permit the most comprehensive reuse possible

For publicly published data, an appropriate license will be selected, i.e., CC-BY-NC license or equivalent. The exact license is to be decided by the data owner at upload time.

When the data will be made available for re-use?

Whenever it is possible to provide the data, it will be made available as soon as possible after its generation. The publication will follow the publisher's guidelines if the data is supplementary to an article.

Usability by third parties of data produced in the project, in particular after the end of the project

Parts of data will not be made publicly available and cannot be used by third parties. The data that is made publicly available can be used by other interested parties, with proper credit to creators and according to the specified license.

Documentation of data provenance

The current data provenance details are not decided. All researchers obey good research handling practice, which includes (described within each discipline) a way of ensuring that documentation of data provenance is ensured, aiming to make the results reproducible.

Data quality assurance procedures

Data Quality Assurance requires a plan to reduce error and bias, data processing and validation errors, and to increase reliability and integrity. In this respect, data quality is acceptable if the collection methods are stable and protected from bias or manipulation, sufficiently documented and detailed, and cross-checked.

Quality control of the data in 6G-SHINE is the responsibility of the partner generating it, whilst quality assurance is a horizontal management activity of the project, led by the Project Coordinator. Details on 6G-SHINE's quality control procedures, which include internal review for deliverables' production, are set out in the Project Handbook [8], which partners should follow.

At this point (Version 1.0), there are no additional quality assurance measures at the partner's levels.

4 ALLOCATION OF RESOURCES

4.1 COST ESTIMATION AND RESOURCES FOR MAKING THE DATA FAIR

Providing FAIR data is the responsibility of each partner, and a budget for open access fees has been included in the project budget for the partners who requested it. As for making data available through Zenodo, we expect to remain within the free limits of Zenodo. Currently, we do not expect extensive data handling procedures for making data available, like converting data to new files, cleaning data or the like before publication.

4.2 DATA MANAGEMENT RESPONSIBILITIES

The project's General Assembly will handle potential issues regarding data management. Project partners have the overall responsibility for the generated data, however, in case of disputes, the project coordinator will be responsible for resolving these, aided by legal advisors, as necessary.

4.3 COSTS AND POTENTIAL VALUE OF LONG-TERM DATA PRESERVATION

No additional cost is foreseen for long-term data preservation. All costs are included in the project budget.

5 DATA SECURITY

Each partner is required to follow the data security standards of their own institution, including guidelines for information security. This will also include local guidelines for backup and restore procedures, as well as for ensuring, e.g., the proper user management for confidentiality, integrity, and accessibility of data. Where possible, partners will use certified solutions for storing data. Partners will transfer data using either secure encryption transportation protocols or by trusted encryption techniques, including proper transfer of encryption key(s) – as necessary. In case not-aggregated data (or not anonymised data) are located on non-national servers, a permission request from participants is needed. All data storage will be on the EU territory, except for the data generated by Interdigital, an associated partner residing in the United Kingdom.

6 ETHICAL ASPECTS

The project does not include any ethics deliverables. The usage of AI methods in the 6G-SHINE project does not bring any ethical issues in relation to objectives, methodology, and impact. The 6G-SHINE project will apply AI methods over data that is synthetically generated with simulations/emulation software platforms or measured in laboratory facilities. Such data refers to radio signal characteristics such as signal power, signal envelope, and data traffic measured in laboratory facilities related to control applications in machinery. For the latter, 6G-SHINE will only measure and use data traffic characteristics related to packet size and packet interarrival times and not the content of the actual communication.

Since this project deals with the basic research design of radio system components, 6G-SHINE will not use any data related to humans and animals. The AI solutions developed in the project will therefore not stigmatise or discriminate people. Also, the AI solutions are not expected to interact, replace or

influence any human decision-making process since the designed solutions will only be used for decision-making of radio electronics for the sake of improving wireless communication performance and reducing energy consumption. The AI solutions will not lead to any negative social and/or environmental impact nor raise any other ethical concerns.

7 OTHER ISSUES

Each partner is expected to follow their national and institution's procedures for data management.

8 CONCLUSIONS

This document provides information about the data types to be generated and used by the project consortium partners, information on formats, storage and making the data FAIR are also provided. The objective of this document is to report on data management issues handled during the implementation of the 6G-SHINE project. Its scope comprises the overall data management plan for the data handling and the reporting and publishing of project datasets.

Daily data management is an important task undertaken by all partners in the consortium according to the elements described in this document and through continuous collaboration between the coordinator, the administrative project manager, the partner responsible for its development and the rest of the consortium.

As this report is generated at the early stage of the project's implementation, it is considered a living document that will evolve along with the project.

Revised and extended versions of this DMP will be prepared according to the GA and internal reviews agreed upon by the partners.

While partners must make sure they respect the defined data management strategy and its procedures, the project coordinator will ensure that all data management issues in the project will be handled promptly and fairly.

REFERENCES

- [1] Data management Plan template for Horizon Europe, https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/temp-form/report/data-management-plan_he_en.docx (EU Grants: Data management plan (HE):V1.1 – 01.04.2022).
- [2] D6.1. Dissemination, Exploitation and Communication Plan for the 6G-SHINE project <https://6gshine.eu>
- [3] General Contract Clauses: Consortium Agreement
- [4] Date Cite <https://datacite.org/>
- [5] OpenAIRE, European project supporting Open Science <https://www.openaire.eu/>
- [6] VBN, Aalborg University's research portal <https://vbn.aau.dk/en/>
- [7] Figshare, online open access repository <https://figshare.com/>
- [8] D1.1 Quality Project Management Handbook <https://6gshine.eu>