

6G-MIRAI –

Data Management Plan

Project information

Project name	Machine Intelligence based Radio Access Infrastructure
Project acronym	6G-MIRAI
Grant agreement	101192369
Call	HORIZON-JU-SNS-2024
Topic	HORIZON-JU-SNS-2024-STREAM-B-01-05
Type of action	HORIZON JU Research and Innovation Actions
Start date	1 April 2025
Duration	36 month

Document information

Associated WP	WP5
Associated Task	All
Associated Deliverable	D5.2
Main authors	Tobias Ley
Contributors	Henning Sanneck
Reviewers	Henning Sanneck, Paul Almasan Puscas
Type	R
Dissemination level	PU
Comment	

Document revision history

Version	Date	Changes	Author
v0.1	2025-07-31	1 st version of deliverable	Tobias Ley
v1.0	2025-09-30	Updates after review and set to full revision	Tobias Ley

Disclaimer

The content of this document reflects only the author's view. The European Commission is not responsible for any use that may be made of the information it contains. While the information contained in the documents is believed to be accurate, the authors(s) or any other participant in the 6G-MIRAI consortium make no warranty of any kind with regard to this material including but not limited to the implied warranties of merchantability and fitness for a particular purpose. Neither the 6G-MIRAI Consortium nor any of its members, their officers, employees, or agents shall be responsible or liable in negligence or otherwise howsoever in respect of any inaccuracy or omission herein. Without derogating from the generality of the foregoing neither the 6G-MIRAI Consortium nor any of its members, their officers, employees, or agents shall be liable for any direct or indirect or consequential loss or damage caused by or arising from any information advice or inaccuracy or omission herein.

Co-funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the SNS JU (granting authority). Neither the European Union nor the granting authority can be held responsible for them. 6G-MIRAI project has received funding from the Smart Networks and Services Joint Undertaking (SNS JU) under the European Union's Horizon Europe research and innovation program under Grant Agreement No 101192369.

© 6G-MIRAI Consortium. This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation, or both. Reproduction is authorised provided the source is acknowledged.

Contents

Acronyms and definitions	3
Introduction	4
1 Data Summary	4
2 FAIR data	7
2.1 Making data findable, including provisions for metadata	7
2.2 Making data accessible	7
2.3 Making data interoperable	9
2.4 Increase data re-use	10
3 Other research outputs	11
4 Allocation of resources	11
5 Data security, integrity and privacy	11
6 Ethics	12
7 Other issues	12
8 References	12

Acronyms and definitions

3GPP	3 rd Generation Partnership Project
DMP	Data Management Plan
EC	European Commission
EU	European Union
FAIR	Findable, Accessible, Interoperable, and Reusable

GA	Grant Agreement
GDPR	General Data Protection Regulation
IP	Intellectual Property
IPR	Intellectual Property Right
WP	Work Package

Introduction

This document D5.2 Data Management Plan (DMP) describes the principles for data management, life cycle handling and governance. It covers data artifacts like documents, software and numerical data. It defines the methods to secure the generation and handling of “FAIR” data making it Findable, Accessible, Interoperable, and Reusable.

The requirements of the Horizon Europe Model Grant Agreement require that a data management plan (DMP) is established and regularly updated. The requirements for research data management of Horizon Europe are described and analyzed in the Annotated Grant Agreement, article 17.

All project participants shall adhere to the principles and guidelines described in the DMP.

1 Data Summary

Types and formats of data:

- Text-based documents (e.g., reports, research results, white papers, workflows, processes, research papers, analysis, specifications, other documents containing text and graphics).
- Software.

- Numeric data (e.g., real time data, data from simulators, generative data, data from lab and test environments) used to train AI/ML models, used for benchmarking, data as specific results.

Origin of data: 6G-MIRAI is focusing on the lower layers of the RAN stack, mainly on layer 1 and 2 (i.e., Physical and MAC layers). The data collected is data from base stations and devices, numeric data of the characteristics of the system (e.g., physical channel characteristics, transmission characteristics, node/device performance, etc.). No end user or data from humans are expected to be collected.

Re-use of data: Data and information from previous EU projects will be re-used case by case where it is suited. Data from partners' internal projects will be used as input and potential as benchmark for the 6G-MIRAI project. Several partners have existing lab set-ups, simulators and test environments, which produced output, data and research results (both from previous internal projects as well as from ongoing parallel projects).

Numeric data generation: Data collection and data generation for training AI/ML models will be done in several different ways. Data will be mainly collected from lab test environments and from simulators, since currently there are no live 6G networks existing. Further data generation using generative models is an option to enlarge the data sets. Collecting and creating authentic data is important for the development of the AI/ML models. Further, data from existing network set-ups can be used for benchmarking.

Expected size of numeric data: The expected size of the data will vary depending on the use case and the lab set-up.

Use of data outside the project: Main users outside the project is academia, for further research. For this purpose, 6G-MIRAI plans to publish research outcome as documents, open-source software and open-source suitable numerical data sets, which are post-processed (cleaned, secure, consistent) and documented.

Classification of data:

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 they can be published as open research data, while the complete, unprocessed data are expected not to be made open.

C – Closed; no part of this data is expected to be made accessible as open research data: Data is classified as sensitive and restricted, due to intellectual property rights, pending IPR and other reasons.

Datasets listed as ‘Limited’ or ‘Closed’ are restricted because of beneficiaries’ legitimate interests, as industrial partners expect commercial exploitation from these data.

Open Data Repositories (Public) as defined in the Dissemination Plan:

- 6G-MIRAI-HARMONY Webpage (<https://6g-mirai-harmony.eu/>) for general information and for linking research outputs in an accessible form
- Zenodo (<https://zenodo.org/>) for all research outputs (as of September 2025, the 6G-MIRAI-HARMONY Zenodo community [ZENODO] is part of the EU Open Research Repository <https://zenodo.org/communities/eu/>)
- ArXiv (<https://arxiv.org/>) for selected technical paper publications
- EU Portal for Project reporting according to EU guidelines.
- Documents classified as restricted/internal are documents that are needed for project planning and execution as well as documents which contain private/ sensitive data and are thus not exposed to the public. These documents are handled internally in a secure, access-controlled repository.
- Zenodo and, where applicable, European Open Science Cloud - EU Node (<https://open-science-cloud.ec.europa.eu/>) for reusable data, publications and software.
- Zenodo and, where applicable, GitHub for software and data sets as well as SW documentation.

Restricted Data Repositories (Internal): Documents classified as limited or closed are documents that are needed for project planning and execution as well as documents which contain private/ sensitive data and are thus not exposed to the public. These documents are handled internally in a secure, access-controlled repository. 6G-MIRAI is using state of the art

data repositories, which are secure, and access controlled. For security reasons not listed here.

Repository for EU-Japan collaboration: A dedicated document and data sharing repository is maintained for facilitate the collaboration between the two partner projects in EU 6G-MIRAI and in Japan HARMONY. This repository is as well internal and restricted. Documents and Data shared can be made public and for that reason transferred to a public repository, if decided by the project. Data and background rights need to be observed.

Responsibilities: It is the responsibility of each partner to provide FAIR data and to follow the principles of the data management plan. The owner of the data is responsible for the content, security, ethics compliance and the legal aspects.

2 FAIR data

2.1 Making data findable, including provisions for metadata

All data shall follow the FAIR principles and provide metadata information about the dataset (description, dates of deposit, authors, venue and embargo), HORIZON Europe funding grant project name, acronym and number, licensing terms, persistent identifiers for the dataset, authors involved in the action, organization, and search keywords. Further, descriptive and structural information shall be provided. Specific metadata format (e.g., Zenodo) needs to be observed (All data deposited in Zenodo or other relevant repositories will be identified with a Digital Object Identifier (DOI) whenever possible to increase findability and citability).

Metadata harvest and indexation will be available depending on the repositories' protocols.

A specific template has been created for 6G-MIRAI to provide metadata.

2.2 Making data accessible

Repository:

Open Data Repositories (Public) as defined in the Dissemination Plan:

- 6G-MIRAI-HARMONY Webpage (<https://6g-mirai-harmony.eu/>) for general information and for linking research outputs in an accessible form
- Zenodo (<https://zenodo.org/>) for all research outputs (as of September 2025, the 6G-MIRAI-HARMONY Zenodo community [ZENODO] is part of the EU Open Research Repository <https://zenodo.org/communities/eu/>)
- ArXiv (<https://arxiv.org/>) for selected technical paper publications
- EU Portal for Project reporting according to EU guidelines.
- Documents classified as restricted/internal are documents that are needed for project planning and execution as well as documents which contain private/ sensitive data and are thus not exposed to the public. These documents are handled internally in a secure, access-controlled repository.
- Zenodo and, where applicable, European Open Science Cloud - EU Node (<https://open-science-cloud.ec.europa.eu/>) for reusable data, publications and software.
- Zenodo and, where applicable, GitHub for software and data sets as well as SW documentation.

Data:

Open data: Data classified as Open will be made public using the 6G-MIRAI selected Open Data Repositories. 6G-MIRAI will open-source specific datasets to support repeatability of research outcome as well as to support the community in further research activities in the domain. For this purpose, the data will be post-processed and cleaned to make sure that the data is usable by third parties.

Restricted data: Data classified as limited or closed will not be made available. Main reason for introducing restricted data is the protection of background and intellectual property rights, competitive considerations as well as legal aspects. Restricted data can play an important role for creating 6G-MIRAI research findings, results and outcome, which in turn can be made public, without making the underlying data public itself. Having this option will allow the partners to use internal, confidential data in the research process, and even though the data is not made generally available, the community will benefit from the research findings. The owner of the data can decide to open-source parts of the restricted

data, for example as limited sample or as anonymized data to support the possibility to reproduce research outcome and for further research activities in the domain. Further, there is the possibility to share restricted data between consortium partners, which are bound by Non-Disclosure Agreement (NDA).

Sensitive personal data: 6G-MIRAI is not handling any human or personal data.

Metadata:

Whenever possible, metadata will be made openly available and licensed, as per the 6G-MIRAI Grant Agreement. 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.

2.3 Making data interoperable

Version control: As far as possible, the existing format and structure of version control of the selected repositories shall be used.

Naming convention: Files shall use the 6G-MIRAI acronym in the title when applicable.

Formats: The format requirements of the selected repositories shall be respected and followed. General and public used formats should be selected to secure interoperability. Proprietary formats shall be avoided.

Common vocabulary for data types: 6G-MIRAI is handling specific data, especially data from lower layers in the RAN stack. This means that data and features can be very cryptic and are not self-explainable. Data sets which are shared and/or open-sourced shall have a supporting description and explanation attached.

References to other data sets: Where possible, the supporting documentation shall include references to other data sets as well as to further information and public available research results and/or software.

2.4 Increase data re-use

Publications in form of conference papers, journals, poster sessions, are an important tool to increase re-use of data. Whenever possible, the relevant datasets and software shall be made open-source and referenced in the publications. This will support the re-use of data, reproducibility of research outcomes and to further support the broader community.

Important aspects are correct handling of metadata, proper documentation, information on usage, data quality assurance as well as information about licensing and terms. Data provenance and historical records shall be made available as part of the metadata where applicable.

Quality Assurance: Dedicated reviews of data including all forms of data (documents, software, numerical data) shall be held before data is made public available to secure high quality for use and re-use. For this purpose, dedicated review meetings will be held with the involved authors as well as selected experts as reviewers.

Process for 6G-MIRAI:

1. Email for GA Approval (If Necessary): For certain outputs, particularly publications, prior approval from the General Assembly (GA) might be required as per our consortium agreement. Email to the internal WP4 mailing list includes a link to the current version of the work in the 6G-MIRAI Sharepoint publications folder.
2. Upload to Zenodo: Once output / publication has been cleared internally and accepted / published externally, the authors are responsible for uploading the appropriate version (e.g., author version of an IEEE publication) to Zenodo 6G-MIRAI-HARMONY Community <https://zenodo.org/communities/6g-mirai-harmony/>.
3. Register the publication internally via the 6G-MIRAI-HARMONY in Sharepoint (include metadata and additional information, in particular DOI and the Zenodo URL).

3 Other research outputs

The 6G-MIRAI Data Management Plan includes types and formats of data:

- Text-based documents (e.g., reports, research results, white papers, workflows, processes, research papers, analysis, specifications, other documents containing text and graphics).
- Software.
- Numeric data (e.g., real time data, data from simulators, generative data, data from lab and test environments) used to train AI/ML models, used for benchmarking, data as specific results.

The Data Management Plan and the FAIR principles applies for all three types mentioned.

4 Allocation of resources

Responsibilities: It is the responsibility of each partner to provide FAIR data and to follow the principles of the data management plan. The owning partner of the data is responsible for the content, security, ethics compliance and the legal aspects.

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, integrity and privacy

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 non-aggregated data (or non-anonymised data) are located on non-national servers, a permission request from participants is needed. All data storage will be on the EU territory.

Data that will be shared with the partner project HARMONY in Japan. The owning partner is responsible for the clearance as well as for the secure handling of the transfer.

Privacy considerations will be paramount. Sensitive or personally identifiable information (PII) will be handled with utmost care, and appropriate anonymization techniques will be applied where necessary. Personal data collected as part of the project will be strictly limited to fulfilling project objectives. Participants' informed consent for data use will be mandatory, ensuring transparency and respect for privacy. The processing of personal data within the project will strictly adhere to the General Data Protection Regulation (GDPR), as outlined in REGULATION (EU) 2016/679. This

regulation ensures the protection of natural persons regarding the processing of personal data and the free movement of such data.

6 Ethics

The partners of 6G-MIRAI are committed to upholding the ethical principles delineated in Article 14 of the Grant Agreement, which espouse fundamental values. Each beneficiary pledges to uphold and safeguard the core values of the EU, including respect for human dignity, freedom, democracy, equality, the rule of law, and human rights, including those of minorities. Furthermore, beneficiaries will execute project activities with the utmost integrity, adhering to pertinent EU, international, and national laws, encompassing the EU Charter of Fundamental Rights and the European Convention for the Protection of Human Rights and Fundamental Freedoms and its Supplementary Protocols. They will also operate by the principles of proportionality, privacy rights, personal data protection, physical and mental integrity, non-discrimination, environmental protection, and the preservation of human health.

6G-MIRAI does not foresee to handle human/personal data, or data that can be classified as Ethical. Main source of data used in 6G-MIRAI is data from the lower layers of the RAN stack, usually very cryptic, numerical data from the base station of device function themselves, not related to end user. The functions of the mobile network, where 6G-MIRAI is using AI to optimize and automate are in real-time, sub-second level. These functions are not operated by humans and thus do not raise concerns of interaction, replacement or influence of human decision-making process.

7 Other issues

Each partner is expected to follow their national and institutional procedures for data management.

8 References

6G-MIRAI_Deliverable_D4.1_Coop-Diss-Impact_v1.0

6G-MIRAI_Deliverable_D5.1_Project-Management-Plan_v1.0