



Document Number: H2020-ICT-52/RISE-6G/D8.6

Project Name:

**Reconfigurable Intelligent Sustainable Environments for 6G Wireless Networks
(RISE-6G)**

Deliverable 8.6

Data Management Plan

Date of delivery: 22/03/2023
Start date of Project: 01/01/2021

Version: 2.0
Duration: 36 months



Deliverable D8.6

Data Management Plan

Project Number:	H2020-ICT-52 / 101017011
Project Name:	Reconfigurable Intelligent Sustainable Environments for 6G Wireless Networks

Document Number:	H2020-ICT-52/RISE-6G/D8.6
Document Title:	Data Management Plan
Editor(s):	NKUA
Authors:	George C. Alexandropoulos, Kyriakos Stylianopoulos
Dissemination Level:	PU
Contractual Date of Delivery:	31/06/2021 (with optional updates throughout the Project)
Security:	Public
Status:	Final
Version:	2.0
File Name:	RISE-6G_D8.6_Rev_2.docx



Abstract

The RISE-6G vision capitalises on the latest advances on Reconfigurable Intelligent Surfaces (RIS) technology for radio wave propagation control, to conceive and implement intelligent, sustainable, and dynamically programmable wireless environments that go well beyond the 5G capabilities and pave the way for the 6G vision. To fulfil its objectives, the project will work on a number of activities including manufacturing and validation of meta-surfaces, modelling and characterization of the wireless medium, and incorporation and validation of signal processing techniques. The design and development of the above items is foreseen to generate a certain amount of data sets, as well as potentially incorporate, process, and extend existing open or proprietary collections of data.

Therefore, a series of procedures should be established, followed by the consortium, and made known to the interested public bodies (legislators, standardisation fora, other ICT-52 projects, independent researchers, and industrial entities). These procedures on the one hand are designed to ensure that all collected data conform to the EU and national legislations (e.g. the General Protection Data Regulation (GDPR)). Additionally, the developed policies should be designed to adhere to the EU's Open Data policies and, specifically, to the Findable, Accessible, Interoperable, Re-usable (FAIR) principles to enhance the impact of the project's outcomes.

This document constitutes the Data Management Plan (DMP) of RISE-6G and provides a detailed information on the procedures that are implemented and followed regarding data-collection, storage, sharing and dissemination, preservation, and long-term access. It provides a comprehensive and detailed plan for the responsible and ethical management of the data generated and includes policies on data formats, quality, access and use, ownership, and responsibility, as well as general management of the Intellectual Property Rights (IPR). The present handbook is addressed both to external parties but, more importantly, to the project Partners and is aimed to be consulted when new project-level data collections are generated and catalogued. This is intended as a living document, and it will be regularly reviewed and updated to ensure compliance with best practices and standards.

Keywords

Data Management Plan, FAIR principles, IPR, GDPR, Open Access, Open Data



Contents

1	Introduction	8
2	Data Management Plan under the Open Data Pilot initiative	9
2.1	European Community strategic framework for DMP	9
2.2	Principles and guidelines for developing a DMP	10
3	RISE-6G FAIR DMP	12
3.1	Types of data under RISE-6G	12
3.2	Practices and Procedures	13
3.2.1	Storage and repositories	13
3.2.2	Naming conventions and unique identifiers	15
3.2.3	Formats	15
3.2.4	Standards	15
3.2.5	Metadata	16
3.2.6	Access, sharing, and licensing	16
3.2.7	Security	16
3.2.8	Responsibility	16
3.2.9	Target audience	17
3.2.10	Archiving and preservation	17
3.2.11	Management of intellectual property	17
3.2.12	Selection and retention periods	18
3.2.13	Ethics	18
3.2.14	Privacy and data protection	18
3.2.15	Budget	18
3.2.16	Legal requirements	18
3.3	Evaluation on the FAIR principles	18
3.3.1	Findable Data	19
3.3.2	Openly Accessible Data	19
3.3.3	Interoperable Data	19
3.3.4	Re-usable Data	19
3.4	Data not relevant to the DMP	19
4	Generating a new data item	19
5	References	21



List of Tables

Table 1-1 - List of Acronyms.....	vii
Table 3-1 Types of data expected to be produced by RISE-6G.....	12
Table 3-2 Data storage and open access options used by RISE-6G	13
Table 4-1 Form to be completed and circulated prior to the publication of a data item.....	20



Table 1-1 - List of Acronyms

Acronym	Full text
AI	Artificial Intelligence
ASCII	American Standard Code for Information Interchange
CA	Consortium Agreement
CSV	Comma Separated Values
DOI	Digital Object Identifier
DMP	Data Management Plan
EC	European Commission
EU	European Union
FAIR	Findable, Accessible, Interoperable, Re-usable
GB	Giga-Byte(s)
GDPR	General Protection Data Regulation
H2020	Horizon 2020
HTTPS	HyperText Transfer Protocol Secure
IP	Intellectual Property
IPR	Intellectual Property Rights
JSON	JavaScript Object Notation
MB	Mega-Byte(s)
PDF	Portable Document Format
RIS	Reconfigurable Intelligent Surface
SSL	Secure Sockets Layer
URL	Uniform Resource Locator
XML	extended Markup Language



1 Introduction

The present data management handbook (Deliverable D8.6) defines the data management procedures and policies established by the RISE-6G project. The main content in D8.6 is the Data Management Plan (DMP) which has been compiled to be harmonised with the H2020 FAIR (Findable, Accessible, Interoperable, Re-usable) principles and the corresponding template made public by the European Commission¹. For the rest of the document, the term *data item* will be used to refer to a single, complete, standalone, piece of data generated/modified/published or otherwise associated with the project. Note that a data item may refer to a collection of individual files. The main key aspects considered in this deliverable are the following:

1. **Data type and characterization:** RISE-6G entails an array of activities ranging from meta-surface prototyping to incorporation of Artificial Intelligent (AI) algorithms. Therefore, a plethora of data collections is expected to accumulate, which each type requiring different practices and policies.
2. **Naming conventions and identification:** The naming rules, patterns, and conventions, based on each data type, as well as unique/univocal identifiers of the data is a crucial aspect of making the data items findable.
3. **Format and metadata:** To make data interoperable and reusable, it is important to consider the format in which the data will be stored. This includes both the file format, such as CSV or JSON, and the structure of the data, such as a relational database or flat file. In this project, standard file formats will be used (refer to Table 3-1). Additionally, metadata are a very powerful tool that will enhance the understanding of the published data. Metadata entries may be stored at collection-level (e.g. at top level directory, repository, data-base), item-level, or even within the data files themselves (e.g. in image and multimedia files). Metadata shall include descriptive information when applicable such as titles, keywords, and author information.
1. **Quality and utility:** To ensure that data is accurate, complete, and reliable, it is important to have rules in place for data quality control, including validation and cleaning of data, and to document the data-quality process. The utility of the data refers to the designated purpose for the generation of a specific data item (e.g., who is intended to use it, for which purpose, via which software).
4. **Storage and access:** For each data item, a decision needs to be made on where it will be stored. Different types may be accommodated better by different storage options. Regardless, the storage medium selected for each data item shall confront to the general FAIR principles, Open Access mandates, and regulations. Some storage options may be controlled by the consortium, while others may rely on ubiquitous public repositories (GitHub², Zenodo³, ArXiv⁴, etc.) Multiple access levels may be granted to different entities to ensure data integrity.
5. **Sharing and use:** The storage option selected shall include a straightforward and established sharing mechanism. Additionally, each published data item will be accompanied by one or multiple (for specific sub-items) licenses (e.g. GPLv3⁵, MIT⁶, Creative Commons⁷). In general, the licenses will be selected to be as open as possible, however

¹ https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm

² <https://github.com/>

³ <https://zenodo.org/>

⁴ <https://arxiv.org/>

⁵ <https://www.gnu.org/licenses/gpl-3.0.en.html>

⁶ <https://opensource.org/licenses/MIT>

⁷ <https://creativecommons.org/licenses/>



restrictive licensing may be occasionally applied to conform to IPR directives or for ethical reasons.

6. **Ownership and Responsibility:** It's important to establish clear rules and agreements related to who is responsible for the data and its management, including data curation, preservation, and disposal. In general, the creator(s) or the storage administrator(s) shall be responsible for adherence to the current DMP.
7. **Archiving and preservation:** Procedures for storage, backup, recovery, and transfer will be incorporated. In general, those procedures are already implemented by the majority of storage options. Ensuring that the data is preserved for long-term access and availability is an important aspect of FAIR. To ensure long-term preservation, the storage options to be selected must guarantee public access for a number of years in the future.
8. **Compliance:** To ensure that the data is being handled in a responsible and ethical manner, it is important to consider the legal and regulatory requirements that apply to the data, such as data protection and intellectual property laws. This is a responsibility of the owner of each data item.
9. **Monitoring and evaluation:** It is important to establish a process for monitoring the implementation of FAIR rules, and for evaluating their effectiveness over time to ensure that they are being followed and to improve them if necessary.

The above key aspects are to be considered per data type or per individual data item, although exceptions and special cases may arise due to confidentiality concerns and IPR compliance. The rest of this document is organized as follows: Section 2 provides a general introduction of the purpose, structure, and guidelines of the Data Management Plan in the context of the European Commission plans for the Open Data Pilot initiative and the mandates for the H2020 projects. Section 3 includes the description of the data expected under RISE-6G and the processes, conventions, rules, protocols, etc. implemented under the Open Data initiative considering the key aspects listed above to ensure that FAIR principles are satisfied. Section 4 provides a guideline to the internal partners for publishing a data item that is to be followed when RISE-6G publishes its results.

2 Data Management Plan under the Open Data Pilot initiative

2.1 European Community strategic framework for DMP

The European Commission (EC) has long established the idea of Open Data as a proponent in capitalising on the public benefits, and therefore, the return of investments of EU-funded research. In the context of Horizon 2020, both research publications and data collections are to be made publicly available by mandate [1] in the spirit of “accelerating scientific discovery, enabling new forms of data-intensive research, and allowing research finding to be systematically taken up by European business and industry” [2].

The RISE-6G Consortium, established under the umbrella of H2020, has a contractual agreement⁸ to participate to the EC's Open Research Data Pilot, which well-complements the Partners' views on open research initiatives. Projects participating in H2020 are required to publish the following types of data items:

- Research publications under “Gold” or “Green” open access [2].
- Data items required for validation of the Project's results and publications.
- Data items constituting or complementing miscellaneous project's results, as specified by their respective DMPs.

⁸ Grant agreement ID: 101017011. DOI: [10.3030/101017011](https://doi.org/10.3030/101017011)



Additionally, projects should deposit their public data to selected open access repositories ensuring that third parties are able to access, understand, exploit, and disseminate their results free of charge. Specifically, two main pillars to the Open Research Data Pilot are⁹: developing a DMP and providing open access to research data, if possible.

The conditions you must adhere to, are:

- Develop (and keep up-to-date) a Data Management Plan.
- Deposit your data in a research data repository.
- Ensure third parties can freely access, mine, exploit, reproduce and disseminate your data.
- Provide related information and identify (or provide) the tools needed to use the raw data to validate your research.

The Pilot applies to:

- The data (and metadata) needed to validate results in scientific publications.
- Other curated and/or raw data (and metadata) that you specify in the DMP.

To achieve the aforementioned objectives, the DMP is one of the key ingredients of successful data management¹⁰. A DMP describes the data management life cycle for the data to be collected, processed and/or generated by a Horizon 2020 project. As part of making research FAIR, a DMP should include information on [3]:

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

It is important to highlight the fact that the DMP, and therefore the present deliverable is not a static set of rules. Instead the DMP needs to be updated over the course of the project whenever significant changes arise, such as (but not limited to) [3]:

- New data,
- changes in consortium policies (e.g. new innovation potential, decision to file for a patent), or
- changes in consortium composition and external factors (e.g. new consortium members joining or old members leaving).

2.2 Principles and guidelines for developing a DMP

A DMP is a document that outlines how data will be collected, processed, stored, and shared during a research project. It is an essential component of any research project, as it helps researchers to ensure that their data is managed in a responsible way. Importantly, DMPs should include clear descriptions and rationale for the procedures regarding data collection, storage, and access. Furthermore, DMPs should follow relevant national and international recommendations and comply to national or EU regulations regarding data protection laws (e.g. GDPR) and IPR. The preparation of a DMP should be made in consultation of the consortium Partners and disciplinary stakeholders. When compiling the handbook, research activity requirements

⁹ <https://www.openaire.eu/what-is-the-open-research-data-pilot>

¹⁰ https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm



and outcomes should be anticipated, to the best of the authors' ability, and the contents should be subject to regular reviews and amendments as the research activity of the Consortium progresses.

Concretely, the key aspects listed in Section 1 are to be considered for each data item published by a Horizon project. In the following, a list of disambiguated terminology is provided that is designed to explain and give context to the terms that will be used in the latter sections:

- **Name and identifier:** The text used as identifier of a data item. This must be unique and refers to the top-level structure of the item (i.e., folder name, database name, repository name, file name). For text documents this is different from the title appearing in the document. For some items, external identifiers may be used (e.g., DOI¹¹ / arXiv identifier¹²).
- **Description:** Accompanying (usually short) text that presents a human-readable explanation of the contents of the relevant data item. The description may contain: (i) link to in-depth explanatory material, (ii) existing or anticipated data volume, (iii) existing or anticipated data formats, (iv) data quality indicators, (v) standards the item conforms to, (vi) relationship with other RISE-6G or external data items, (vii) intended use (including potential software needed).
- **Formats:** File formats used as part of this item. Note that some data items may be available as multiple formats (e.g. DOCX/PDF), while in others, a file format tag is not applicable (e.g. database / code repository).
- **Standards:** Reference to existing suitable standards of the discipline, if applicable.
- **Metadata:** Accompanying information for a data item. The format and location of the metadata is type- and storage- dependent. Different storage options and file types may allow for metadata to be included during the creation of the item. In other cases, the metadata may simply appear as a readable (text-like) file stored along the data item. In case of existing standard metadata practices for specific items/collections/files, it is advisable to be followed.
- **Data access and sharing:** How each data item is to be shared and under which conditions. It outlines technical mechanisms, software, authentication, etc. for the dissemination of the data. Restrictions such as embargo periods should be clearly indicated and each item must be accompanied by a respective license.
- **Security:** Concerns the integrity of the data against malicious actions, protection from unauthorized access, etc. It is especially important for cases of confidential information, IP managements, content sensitive to public rights. Restrictions, multi-level permissions, and embargoes may need to be introduced for security-critical data items.
- **Responsibility:** Refers to names of individuals or Institutions, part of the consortium or external co-operators, that are responsible for data management per data item.
- **Target audience:** Refers to groups of potentially interested parties for a data item.
- **Intellectual Property Rights:** Entities or persons may hold the intellectual property rights to published data items, and they should be clearly stated. IP protection procedures may need to be included if necessary. Any copyright constraints (e.g. copyrighted data) should be noted.
- **Archiving and preservation:** Concerns the long-term storage, presence, and availability of the data and the relevant metadata. This concern must be satisfied to harmonize

¹¹ <https://www.doi.org/>

¹² https://arxiv.org/help/arxiv_identifier



the DMP with the EC Open Data Pilot which states that the research outcomes must be publicly available even after the end of the project.

- **Selection and retention periods:** Refers to how and what data will be selected for archiving, how long the data will be held, and plans for eventual termination of the data availability.
- **Ethics and privacy:** How informed consent is to be handled, if applicable, and how private data will be protected including aspects of confidentiality and other ethical issues that may arise throughout the project.
- **Budget:** Costs associated with the preparation, storage, publication, and archiving of the data. The budget should also specify the source of the funding for each data item (i.e. from the Consortium budget, Institutional funds, third-party costs, etc.).
- **Data organization:** Refers to naming conventions, versioning, and other procedures for publishing data in an organised manner.
- **Quality assurance:** Procedures to ensure the quality of the data published by the project.
- **Legal requirements:** All relevant federal or funder requirements for data sharing. The procedures set by the DMP should adhere to those requirements.

3 RISE-6G FAIR DMP

This section includes the concrete practices and procedures that comprise the data Management Plan of RISE-6G. The first step concerns a detailed taxonomy of types of data expected to be produced by the project, their purpose, file formats, extensions, etc. This is important for establishing procedures that best take into consideration the different requirements in managing a wide collection of data.

3.1 Types of data under RISE-6G

Table 3-1 below presents a non-exhaustive list of data types expected to be produced by RISE-6G. The data are categorised according to their nature and possible formats are listed alongside. In the following section, specific details of the DMP will be presented by taking into consideration different data categories.

Table 3-1 Types of data expected to be produced by RISE-6G

Type of data	Category	Possible file formats/extensions	Description
Publication	Text/Document, Web document	PDF, HTML	Papers published in conferences, journals, etc.
Deliverable (Report)	Text/Document	DOCX, PDF	Technical documents reporting topic-related results as specified in RISE-6G Technical Annex.



Code	Source code file (MATLAB®, Python, etc.)	.py, .m	Public code repositories, implementations of algorithms, APIs, numerical simulations, etc.
Compiled code	binary	.exe, .pyc, .o, .asv	Pre-build binary (closed source) code developed as part of the project.
Code results	Binary / text / spreadsheet / markup	.mat, .npy, .h5, .csv, .tsv, Flat file (Fixed-width), xlsx, XML	Numerical results produced by the execution of code files. They can be either in human readable format or in programming-readable files.
Technical Datasheet	Document	PDF	Technical specifications of technological results (e.g., meta-surfaces).
Visual & media results	Image / Video	PNG, SVG, JPEG, PDF, EPS, MP3, MP4, WAV, AVI, HTML	Results of the project that are best conveyed in a media form (e.g., demonstrations, simulations).
Measurements	Binary / text / spreadsheet / markup	.mat, .npy, .h5, .csv, .tsv, Flat file (Fixed-width), xlsx	Numerical values arising from measurements of the proof-of-concept / prototyping aspects of the project. Those are treated similarly to code results.
Experimental results	Any	Any	Similar to code results, but concern only results attained via proof-of-concept
Configuration / setup	Text / markup / source code	JSON, YML, XML, TXT, .md	Configuration files that are needed to specify the exact parameters of other results (simulations, proofs-of-concept, etc.
Metadata / README	Text / Any	TEXT, .md	Accompanying files that provide additional information or instructions on other results.
Presentations / Tutorials	Document / Media	.pptx, PDF, MP4, WAV, AVI	Results and advancements that are best conveyed via a presentation or a tutorial. Those may include either media presentations with voice-over/video or written documents (e.g., slides)
Miscellaneous	Any	Any	Unclassified data.

3.2 Practices and Procedures

3.2.1 Storage and repositories

The selection of a storage medium is probably one of the most important factors in distributing Open Data. Crucially, given the wide variety of data categories related to this project, there is no single solution that satisfies all the requirements equally. Therefore, it has been decided that RISE-6G will utilize a number of storage and publishing options. Nevertheless, an effort has been made to provide as few different storage options as possible to increase the findability of the data. In general, the RISE-6G website has been selected as the main point of access from the outside world, so it contains both individual data items/files as well as references to data hosted elsewhere. All different storage options are listed in Table 3-2, and their use with respect to their associated data types is discussed below.

Table 3-2 Data storage and open access options used by RISE-6G

Storage	Use (types of data)
GitHub ¹³	Source code, binary code, small result/data files

¹³ <https://github.com/RISE-6G>



RISE-6G website ¹⁴	Publications (references only), Deliverables,
Zenodo ¹⁵	Text/technical documents, presentations, large result/data files
Publisher websites	Gold access publications (full text)
ArXiv ¹⁶	Green access publications (full text)
ResearchGate ¹⁷	Green access publications (full text)

- **Source code:** The source code made public by the project will be hosted on **GitHub**, since its use is ubiquitous within the programming community. The project maintains an active organization account and the standalone code bases are being posted as repositories¹⁸. The code-related functionalities and integration GitHub provides make it the perfect tool for this type of data. Note that the purpose of those repositories is strictly to provide open access rather than contribute to active code developing. To that end, repositories under the RISE-6G GitHub account may be “forked”¹⁹ from their original repository.
- **Publications:** By definition, published papers appear in the **websites of the respective publishers**. In case of a Gold Access, those websites constitute the main storage option for the specific papers. For publications falling under Green Access, an open access copy of the text/document will be hosted in some **open access repository** (namely *ArXiv* and *ResearchGate*). To provide a single collection point, the list with all RISE-6G’s publications is included on the **project’s website** and it is periodically updated. That list contains each paper in its proper reference form as well as a link to the respective Open Access file from the external hosting website. This approach has been adopted for the publications since (a) papers will be uploaded in the aforementioned locations in the first place, (b) those websites provide high-quality meta-data and indexing options.
- **Deliverables:** The deliverables constitute the main collections of advancements and results of the project. Therefore, they require the highest level of findability. To that end, the deliverables (of “public” dissemination level) are to be published on the project’s website.
- **Text Documents, presentations:** Various standalone documents of text-like format (including technical datasheets, presentations, etc.) will be uploaded to Zenodo, under the RISE-6G community. Zenodo will also host important collaborative publications (e.g., ones prepared alongside other ICT-52 papers [4], or the RISE-6G inaugural publications [5],[6]). The decision to resort to a third storage option for multiple documents has been made because Zenodo provides automatic metadata, DOI, reference extraction, etc., therefore making it a better option than the website. To increase the documents’ findability, all material uploaded to Zenodo will be accompanied by a related announcement on RISE-6G website which will include the relevant links.
- **Numerical and experimental measurements and results:** Such documents may be either in binary or in structured ASCII format (e.g., CSV). Their size is also difficult to determine since they may take up several GBs of data. Data items that are associated with published code repositories will be hosted in GitHub, alongside their respective

¹⁴ <https://rise-6g.eu/>

¹⁵ <https://zenodo.org/communities/rise-6g>

¹⁶ <https://arxiv.org/>

¹⁷ <https://www.researchgate.net/>

¹⁸ <https://github.com/orgs/RISE-6G/repositories>

¹⁹ <https://docs.github.com/en/get-started/quickstart/fork-a-repo>



codebases, assuming they fit the platform's size limits²⁰. For larger data files, Zenodo will be used as a repository²¹ and a link will be placed in the item that references the data (i.e., publication, codebase, deliverable). The storage of even larger files will be evaluated per case.

- **Metadata, configuration, and setup files:** Those are not treated as standalone files; therefore, they are intended to be stored alongside their corresponding data items. It is important to note that metadata may not be stored as separate files, but they may be embedded on the storage platform that displays the data (e.g., Zenodo, ArXiv).
- **Visual and Media files:** In cases of standalone files, they will be stored either on Zenodo, or on the website depending on their primary use. In cases of accompanying files to other data (e.g., plots of results, videos of presentations), they are to be stored alongside the primary data file.

3.2.2 Naming conventions and unique identifiers

- **Deliverables:** Deliverables are to be named under the following pattern: *RISE-6G_D[X].[Y]_V[Z].[W]_[NAME].[EXT]* where [X], [Y], [Z], [W] are numerical values that correspond to the deliverable's number, and version [NAME] refers to the name of the current deliverable (optional) and the [EXT] is either .docx or .pdf.
- **Publications:** Publications are distributed by the respective hosting platforms. Therefore, it is not possible to specify a naming convention for the actual files. Nevertheless, a DOI or a similar unique identifier, generated by the publishing sites, is included in the metadata on every document.
- **Codebases:** It is inconvenient to specify naming conventions and patterns for code files and codebases since programming languages apply their individual rules. Therefore, no naming patterns are specified for code files with respect to their interoperability. It is important to note that GitHub repositories already include metadata regarding authors, versions, etc., as well as a unique identifier which corresponds to the URL of the repository.
- **Rest of the files:** The rest of the files (to be published on Zenodo) will follow the following convention: *RISE-6G_[TYPE]_[NAME]_V[X].[Y].[EXT]* where [TYPE] corresponds to one of the data types specified in the previous section, [NAME] refers to the name of the document/collection, [X] and [Y] specify the version of the document and [EXT] corresponds to the extension.

3.2.3 Formats

For the file formats that are used, please refer to Table 3-1.

3.2.4 Standards

Apart from the standards implied by the file formats, the following standards are used:

- **Publications:** IEEE formats²² and templates or formats specified by the journal/conference or other publishing agents will be used.

²⁰ GitHub current file size limit is 25MB.

²¹ Zenodo's file size limit is 50GB.

²² <https://www.ieee.org/conferences/publishing/templates.html>



Document: H2020-ICT-52/RISE-6G/D8.6

Date: 22/03/2023

Status: Final

Security: Public

Version: 2.0

- **Code:** Best practices corresponding to each language/environment will be used (e.g., PEP styles²³).
- **Deliverables:** The deliverable reports will be harmonised with the EC directives regarding H2020 projects²⁴.

For the rest of the data items, relevant standards and guidelines will be evaluated on a per-case basis.

3.2.5 Metadata

Metadata may be categorized on three broad categories: Metadata that are embedded within the primary document, metadata that are created during the upload process of the data items on the respective storage options and are thus visible through the browser, and metadata that are created as standalone files and are placed together with the primary data item(s). For the first two cases, the metadata practices to be followed are provided by the storage options themselves (e.g., publication metadata on IEEEXplore²⁵) and are thus standardised. For the latter case, the metadata files are suggested to be written in a widely used format (e.g., XML, JSON) if possible.

3.2.6 Access, sharing, and licensing

All storage options listed above, provide unrestricted and free access to all entities (people or automated tools). Furthermore, sharing/export links are provided by the hosting websites, namely with relevant metadata:

- PDF for publications
- Text references for publications
- Bibliography formats for publications (BibTeX, RIS²⁶, etc.)
- Git repositories or ZIP archives for code repositories
- ZIP archives for collections of data

Each published data item will be accompanied by its individual license, either explicitly (as a separate file or meta-data), or being implied by the terms & conditions agreed upon when uploading the data to the respective storage options. Concretely, RISE-6G utilises the GPLv3, MIT, and creative-commons licenses for its free and open source code and open data.

3.2.7 Security

The security of the published data relies on the security provided by the hosting platforms. The external platforms are provided by trusted institutions and are widely used, so the employed security protocols and procedures are expected to be of the highest level. In an equivalent manner, the website of the project has been developed under the highest security practices (i.e., HTTPS/SSL encryption, Up-to-date software, active maintenance, and support throughout the lifetime of the project).

3.2.8 Responsibility

Under this project, the entities responsible for the upload, management, content, quality, and accessibility of each data item are set to be the original author(s)/creator(s). Under the set of

²³ <https://peps.python.org/pep-0008/>

²⁴ https://ec.europa.eu/research/participants/docs/h2020-funding-guide/grants/grant-management/deliverables_en.htm

²⁵ <https://ieeexplore.ieee.org/Xplore/home.jsp>

²⁶ In this context, RIS refers to the Research Information System file format used for bibliographic indexing [https://en.wikipedia.org/wiki/RIS_\(file_format\)](https://en.wikipedia.org/wiki/RIS_(file_format))



responsibilities also falls the responsibility of ensuring that each published item conforms with the procedures specified in the current DMP.

3.2.9 Target audience

In general, the data made available to the public by RISE-6G are primarily targeted at the following audiences: (European) legislative bodies, the academic community (especially those involved with the areas of metamaterials, material sciences, wireless communications, communications engineering, system engineering, electrical and electronic engineering, computer science and computer engineering, informatics and telecommunications, mathematics and physics), industrial entities (especially those associated with the manufacturing, infrastructure, telecommunications, or computer software/engineering industries) and amateurs/enthusiasts/students of the relevant fields described above. In practice, it is necessary to assume some level of familiarity with the field of research for the understanding of the contents, utilization, and inference of the data.

3.2.10 Archiving and preservation

The storage options selected are expected to provide a high level of archiving benefits. Since the external platforms are widely-used tools with firm policies, the longevity of the data is ensured. Similarly, the website, by mandate, will be operational for a number of years after the conclusion of the project. Note that data stored onto the website are intended to be duplicated in other platforms. Finally, the results of RISE-6G are also mandated to be uploaded to the EC portal, which provides public access and guarantees their longevity.

3.2.11 Management of intellectual property

IPR is an important aspect that RISE-6G consortium will carefully consider. Scientific results and measurements defined during the project are public domain research, intended to be used in international fora to advance technological development and scientific knowledge. Therefore, methods, techniques, obtained architectures and functionalities should be available for peer-review and self-adaptation. Valuable IPR that might be generated during the project from the work in the areas of new technological innovations with direct product potential will be protected by the consortium and/or single partner entity within the project. The IPR shall be shared with reasonable rules strictly adhering to the H2020 contract rules.

For handling patents, the consortium will also apply proven methods used in previous accepted projects. The partners will inform the consortium about technologies, algorithms, etc., offered for use in the Work Packages that they have patented, are in the process of patenting, or consider patenting. Similarly, if patentable methods and techniques are generated within project activities, the patenting activities will aim to protect the rights of all partners participating in these activities. Lists of patents related to the project, whether adopted, applied or generated will be maintained for reference, and included in reports made to the EC. The Consortium Agreement (CA) will provide rules for handling confidentiality and IPR to the benefit for the Consortium and its partners. All the project documentation will be stored electronically and as paper copies. Classified Documents will be handled according to proper rules regarding classification (as described above), numbering, locked storing, and distribution limitations. The policy that will govern the IPR management is driven by the following principles, to be detailed in the Consortium Agreement: (1) Policy for Ownership and Protection of knowledge, (2) Dissemination and Use policy, (3) Access rights for use of knowledge, (4) Confidentiality, (5) Ownership of results / joint ownership of results / difficult cases (i.e., pre-existing know-how so closely linked with result difficult to distinguish pre-existing know-how and result), (6) Legal protection of results (patent rights), (7) Commercial exploitation of results and any necessary access rights, (8) Commercial obligations, (9) Relevant Patents, know-how, and information Sublicense and, (10) Pre-existing



know-how excluded from contract. Nevertheless, many specific IPR cases that will need a concrete solution from the bases previously fixed may exist. In these conflict situations, the General Assembly will be the responsible to arbitrate a solution.

Please also refer to Section 3.4 for data expected not to be made available under the current DMP specifications potentially due to intellectual property rights issues.

3.2.12 Selection and retention periods

Due to intellectual property rights (see also Section 3.2.11), as well as reasons of confidentiality, academic integrity, ethics and privacy (see also Sections 3.2.13 and 3.2.14), certain data items may be made public at a later stage (after an embargo period), be made public in an altered or partial form, or not be made public at all. Since such cases are difficult to be anticipated beforehand, evaluation and decision-making will be carried out on a per-item case by the project partners involved and potentially third parties with intellectual property interests or other conflicts.

3.2.13 Ethics

RISE-6G generates, collects, utilises, and processes data that are not associated with ethical concerns. Nevertheless, at cases where ethical concerns may potentially be relevant, the project will ensure that the complete pipeline (from data generation to Open Access publishing) conforms to the EC ethics guidelines²⁷.

3.2.14 Privacy and data protection

RISE-6G, in general, generates, collects, utilises, and processes non-personal and non-sensitive data. Therefore, it is largely disassociated with privacy concerns. Nevertheless, RISE-6G will also carefully collect, process and treat confidential data in full respect of GDPR policies. In particular, when personal data are collected and used, RISE-6G will elaborate a data protection impact assessment. In addition, technical measures might be applied including encryption, limitation of the amount of collected personal data and deletion of data no longer needed. Finally, involved people in public field-trials will have the right to see what personal data have been collected and stored and how RISE-6G will make use of it, despite the fact that the public field trials will be designed to minimise the need for personal data collection.

3.2.15 Budget

Part of the budget allocated to RISE-6G is directed to Open Access and other data-related operations. The specific details are given in the Grant Agreement. In other cases, costs related to data management will be directly allocated by the respective institutions. Nevertheless, RISE-6G has made efforts to limit the budget allocated to data publishing by relying on free/already existing solutions.

3.2.16 Legal requirements

All stages of the pipelines associated with the management of the data will be ensured that they comply with the current EU regulations (e.g., GDPR) and national legislation. The primary responsibility of this task falls under the responsible institutions for each data item as specified above.

3.3 Evaluation on the FAIR principles

The current section contains a self-evaluation of the processes described above with respect to their adherence to the requirements of findability, accessibility, interoperability, and re-usability with the aim to illustrate how those practices achieve FAIR data.

²⁷ In this context, RIS refers to the Research Information System file format used for bibliographic indexing [https://en.wikipedia.org/wiki/RIS_\(file_format\)](https://en.wikipedia.org/wiki/RIS_(file_format))



3.3.1 Finable Data

To ensure findability, RISE-6G promotes a “single-point-of-contact” policy via its website. The rest of the data can be retrieved from there by following hyperlinks to a limited number of well-known repositories. The naming conventions, unique identifiers, and additional metadata are designed to enhance the discovery process of the intended data items.

3.3.2 Openly Accessible Data

All storage options provide free-of-charge, registration-free, and unconstrained access to the stored data. Their longevity and archiving are also ensured due to the policies of the storage options.

3.3.3 Interoperable Data

Data published under RISE-6G are designed to adhere to current standards, use most convenient file formats, provide ample metadata, and be stored at wide-access platforms with established procedures for harmonization. Therefore, a high degree of interoperability is expected being produced by this project.

3.3.4 Re-usable Data

The metadata specified to describe the rest of the data to the highest degree possible, while the sharing licenses, restrictions, and IPR rules are designed to allow as few restrictions as possible.

3.4 Data not relevant to the DMP

While RISE-6G makes efforts to make as much data as possible openly accessible, there are specific categories of data that do not fall under the Open Access mandate for varied reasons, such as:

- Internal project files relating to its operation, management, and research procedures (e.g., internal presentations, email or other correspondence, initial/raw versions of later published data etc.).
- Initial or internal data relating to research and development activities (e.g., editable versions of the published documents/publications, research plans, future ideas, backups, etc.
- Code, software, or other material used as part of project results but developed out of the scope RISE-6G (by third parties or even by individuals) and made available to the project as part of a collaborative initiative.

4 Generating a new data item

As far as items such as deliverables and publications are concerned, it is relatively straightforward to ensure that the processes set by the DMP in section 3 are followed due to the detailed and standardised procedures that need to be followed when uploading those specific items. For the rest of the data, however, there is a potential danger of not adhering to all DMP practices when making them public. Additionally, such items are difficult to catalogue. To remedy for those facts, this section provides a concrete procedure to be followed by all partners (all collaborations of partners) when producing a new data item. Concretely, the following steps should be carried out by the authoring/creator partner(s):

- **Step 1:** Complete a copy of the form appearing in Table 4-1 while ensuring adherence to the established practices. Refer to Section 3.2 for disambiguation.



- **Step 2:** Share the intention of publishing a new data item under Open Access alongside the specified form and the contents of the data with the Consortium via the general mailing list using a relevant subject line.
- **Step 3:** Allow a period of 30 days for any partner to raise potential objections that will be solved alongside the Project Coordinator.
- **Step 4:** Proceed with the upload of the public data ensuring that the information is as specified in the circulated form.

Table 4-1 Form to be completed and circulated prior to the publication of a data item

Authoring/Creator Institution(s)	Short name. Multiple Institutions are separated by commas.
Name of the top-level file	Respecting the naming conventions of Section 3.2.2 if applicable.
Storage platform	Must correspond to the ones listed in Table 3-2.
Object identifier or URL (if known)	May be possible to retrieve during the upload process.
Type of data	Refer to Table 3-1.
Upload date	e.g. 2023-01-20
Brief description	Some text to serve as a guideline.
Version	Of the form [X].[Y].
License	Name of the re-use license or “protected” for non-open data.
Embargo or retention period	Specify the duration of the retention period or leave blank.
Other restrictions	List any that apply or leave blank.
Ethics/Privacy/Legal concerns apply?	Yes/No. (If Yes, the concerns must be resolved prior to the publication of the data).



5 References

- [1] European Commission, "Commission launches pilot to open up publicly funded research data", Press release IP-13-1257, Brussels, December 2013.
- [2] European Commission, "Towards better access to scientific information: Boosting the benefits of public investments in research", as part of: *Communication from the commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions*, July 2012, document number: 52012DC0401, procedure number: 201832.
- [3] European Commission, Directorate-General for Research & Innovation (2016) H2020 Programme Guidelines on FAIR Data Management in Horizon 2020, Version 3.0. Luxembourg, European Commission, Directorate-General for Research & Innovation 12pp. DOI: <http://dx.doi.org/10.25607/OBP-774>
- [4] Nielsen, Lars, Gavras, Anastasius, Dieudonne, Michael, Mesogiti, Ioanna, Roosipuu, Priit, Houatra, Drissa, & Kosmatos, Evangelos. (2022). Beyond 5G/6G KPIs and Target Values. Zenodo. <https://doi.org/10.5281/zenodo.6577506>
- [5] E. C. Strinati *et al.*, "Wireless Environment as a Service Enabled by Reconfigurable Intelligent Surfaces: The RISE-6G Perspective," *2021 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*, Porto, Portugal, 2021, pp. 562-567, doi: 10.1109/EuCNC/6GSummit51104.2021.9482474.
- [6] E. C. Strinati *et al.*, "Reconfigurable, Intelligent, and Sustainable Wireless Environments for 6G Smart Connectivity," in *IEEE Communications Magazine*, vol. 59, no. 10, pp. 99-105, October 2021, doi: 10.1109/MCOM.001.2100070.