



Project: **SEAWave**

## **Data Management Plan, DMP**

Work Package: WP11

Deliverable: D11.2

Deliverable No.: D40

## Abstract

The pervasiveness and social-economic dependence on wireless technology has steadily increased over the last three decades. Currently, the 5<sup>th</sup> generation (5G) New Radio (NR) cellular system is being deployed to unlock the potential of new applications that require the connection of many more devices (Internet of Things), higher data rates and low latency (autonomous driving, 'Factory of the Future'). 5G operates in two frequency bands, 5G NR FR1 and 5G NR FR2. Many exposure parameters of 5G are similar to those of 2G-4G. However, there are also many differences that lead to major knowledge gaps, all of which will be addressed by the SEAWave project. SEAWave will (i) quantify the differences in exposure patterns between 2G-4G and 5G for the entire population including children; (ii) provide new tools and instruments for reliable exposure evaluation of base stations, local networks in factories, and enduser devices; (iii) provide the means to minimise exposure; (iv) generate important new scientific data for assessing the health risk from exposure to the new frequency bands (FR2), especially with regard to the potential (co-)carcinogenicity of skin exposure and other hazardous effects; and (v) provide knowledge for effective health risk communication and dissemination to various stakeholders. The purpose of data collection within this project is therefore i) the quantification of exposure, ii) the creation of new tools and instruments for this purpose, and iii) the generation of new scientific data for assessing, in a solid fashion, possible health risk for 5G-FR2 exposures.

## Project Details

Project name	SEAWave
Grant number	101057622
Start Date	01 Jun 2022
Duration	36 months
Scientific coordinator	Prof. T. Samaras, Aristotle University of Thessaloniki (AUTH)

## Deliverable Details

Deliverable related number	D11.2
Deliverable No.	D40
Deliverable name	Data Management Plan, DMP
Work Package number	WP11
Work Package name	Project and Consortium Management
Editors	Caterina Merla, ENEA - Italy

---

Distribution	Public
Version	1
Draft/final	Final
Keywords	Data Management, Data security, Open Research Data, FAIR Data

## Contents

1	Introduction .....	6
2	Data Summary.....	6
2.1	Purpose of the data collection/generation and relation to the objectives of the project 6	
2.2	Types and formats of the data .....	7
2.3	Size of the data.....	7
2.4	Targeted users of the collected data .....	7
3	FAIR Data.....	7
3.1	Making data findable, including provisions for metadata .....	7
3.1.1	Discoverability of data (metadata provision) .....	7
3.1.2	Identifiability of data.....	7
3.1.3	Naming and conventions used .....	7
3.1.4	Approach towards keywords .....	8
3.2	Making data openly accessible .....	9
3.2.1	Data to be made publicly available and rationale for keeping some data closed....	9
3.2.2	Methods to access the data.....	10
3.2.3	Restricted area access.....	10
3.3	Making data interoperable .....	10
3.3.1	Data interoperability and used vocabulary .....	10
3.4	Increase data re-use (through clarifying licences).....	11
3.4.1	Data licensing .....	11
3.4.2	Reusability at the end of the project .....	11
3.4.3	Data quality assurance process .....	11
3.4.4	Re-usability duration.....	11
4	Allocation of resources .....	11
5	Data security .....	12
5.1	Data recovery .....	12
5.2	Secure storage.....	12

---

5.3	Transfer of sensitive data.....	12
6	Ethical aspects.....	12
7	Use of the DMP within the project.....	13
8	List of acronyms .....	13

## 1 Introduction

The Data Management Plan (DMP) within the project is a part of the action called Open Research Data (ORD) in Horizon Europe in which SEAWave participates. This document describes the initial Data Management Plan (DMP) for SEAWave and is based on the Guidelines Exploitation & Open science in Horizon Europe.

A DMP describes the data management life cycle for the data to be collected, processed and/or generated by a Horizon Europe project. As part of making research data findable, accessible, interoperable and re-usable (FAIR), a DMP should include information on:

- 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).

## 2 Data Summary

### 2.1 Purpose of the data collection/generation and relation to the objectives of the project

The purpose of data collection/generation is to gather evidence of the developed quantification of the 5G exposure, the creation of new tools and instruments for this purpose, and the generation of new scientific data for assessing, in a solid fashion, the possible health risk for 5G-FR2 exposures. Additionally it will be necessary to collect data from different environments (ex. urban areas and smart industries) in order to permit the assessment of exposure in real 5G scenarios to establish appropriate standards and validate the results coming out from the biological investigations *in vitro*, *in vivo* and on human volunteers.

Therefore, the collected data will include the experimental and theoretical procedures for characterization of the exposure along with the protocols for biological testing. Finally, simulation data and codes, for example for the automatic control of measurement instruments, or the dosimetric assessment will be also collected.

---

## 2.2 Types and formats of the data

The data will be collected in text, numerical and image formats and gathered in files whose extension are defined by the equipment/software used for generation and collection. Some examples include but are not limited to .xlsx (Excel spreadsheet), .ppt (PowerPoint), .mat (Matlab), .txt (text), .s2p (touchstone), .avi (video). Data will be generated by individuals or groups of researchers in all involved institutions.

## 2.3 Size of the data

The expected size cannot be predicted at this stage but it is reasonable to assume that it will hit the tens of Gigabyte range.

## 2.4 Targeted users of the collected data

The data will be useful to members of the scientific community who are willing to reproduce and build on the described experiments, develop similar technologies or test new algorithms.

# 3 FAIR Data

## 3.1 Making data findable, including provisions for metadata

### 3.1.1 Discoverability of data (metadata provision)

Considering the strongly interdisciplinary nature of the SEAWave project, the consortium favors the adoption of a broad and domain agnostic metadata standard that the EU recommends to its member states for recording information about research activity: the Common European Research Information Format (CERIF) standard is described at <http://www.eurocris.org/cerif/main-features-cerif>.

An additional advantage of a CERIF inspired standard is that SEAWave's DMP managing institution (ENEA) currently uses a research information system developed by Elsevier that implements the CERIF standard (PURE).

### 3.1.2 Identifiability of data

For publication data unique identifiers such as Digital Object Identifiers will be used. For other data the identification mechanism described in "Naming and convention used" will be adopted.

### 3.1.3 Naming and conventions used

The following structure is proposed for a SEAWave data set identifier:

"Project"\_"Date"\_"Time"\_"Name"\_"Type"\_"Extension"\_"Place"\_"Creators"\_"Target user"\_"Other"

---

Where:

- “Project” is the project name (SEAWAVE by default).
- “Date” is the date in format “YYMMDD” which is chosen to allow data that was taken at similar dates to be stored in close locations. For the same reason the date and time fields are set to precede the name field.
- “Time”: is the time in format “HHMMSS” if relevant, or NA by default.
- “Name” is a short name for the data.
- “Type” describes the type of data (e.g. publication, measured data, simulation data, protocol description ...).
- “Extension” describes the data file extension.
- “Place” describes the location where the data were produced.
- “Creators” defines the individual(s) who generated the data.
- “Target user” defines the target audience of the data, if known.
- “Other” is an optional field for additional details (whose default value is NA).

For example:

“SEAWAVE”\_“141122”\_“093030”\_“Skinprofile”\_“Measured”\_“png”\_“ENEA”\_“Cr C. Merla”\_“Partners and public”\_“NA”

is a file named “Skinprofile” that was taken on November 14<sup>th</sup> 2022 at 9:30 AM and contains image data with png extension. Such data was generated in ENEA by C. Merla and its storage target SEAWave partners, as well as general public. A simple excel spreadsheet has been created and will be distributed to all partners for a highly automatized generation of file names using the described format. An example of the file name generated using such a tool is visible in Figure 1.

#### 3.1.4 Approach towards keywords

For publication data the official keywords list provided by the publisher will be used. For other data keywords will be selected by the data owner.



	A	B	C	D	E	F	G	H	I	J	K	L
1	Field	Entry			Notes							
2	Project	SEAWAVE										
3	Date	141122			Year Month Day format! Ex. 170519 for the 19th of May 2017							
4	Time	121834			24h format. Type NA if not relevant							
5	Name	Skinprofile										
6	Type	Measured data			Ex. publication, measured data, code, protocol, ...							
7	Extension	png			File extension if known							
8	Place	ENEA			You can specify additional details such as particular lab							
9	Creator(s)	C. Merla										
10	Target User	Partners and public										
11	Other	NA			If you want to add anything else							
12												
13	<b>File Name:</b>	SEAWAVE_141122_121834_Skinprofile_Measured data_png_ENEA_C. Merla_Partners and public_NA										
14		<i>(Cell D13 is the recommended file name for uploading to Zenodo repository.</i>										
15		<i>You can copy and paste it to any other document,</i>										
16		<i>but please do not mess with its definition in this file!)</i>										
17												
18												

Figure 1: simple Excel utility to be distributed to all partners for the generation of data name according to the DMP convention.

## 3.2 Making data openly accessible

### 3.2.1 Data to be made publicly available and rationale for keeping some data closed

**Publications:** Partners will be free to publish and disseminate their own results according to the procedure defined and agreed in the Consortium Agreement. The consortium will comply with the Grant Agreement open access clause for the publications generated from the project, but will deposit them into institutional (closed) repositories (like the one of the Aristotle University of Thessaloniki, <https://ikee.lib.auth.gr>) before moving them to public data repositories like Zenodo (<https://zenodo.org>). The timing and approach in moving publications to the public repository is similar to those for the other data and is discussed in next session.

**Other data:** SEAWave's partners strive for maximum openness of data collected and generated during the project but reserve the right to evaluate which data will be made publicly available along with the time for publication on a case by case basis. The "Guidelines for FAIR Data Management in Horizon Europe" recognize the need to balance openness and protection of scientific information, commercialization and Intellectual Property Rights (IPR), privacy concerns, security as well as data management and preservation questions. It is expected that the dominant causes for enforcing data access restriction during SEAWave will be protection of IPR and commercialization strategies. It is also expected that the openness stance regarding individual items can be reviewed and updated periodically. For example, test results or experimental protocols can be made publicly available after the consortium has filed for the corresponding patents. Other measurement data may become publicly available after reasonable time that will allow project partners to publish them.

---

The decision as to data openness and availability/embargo time will be made through a vote held by the Project Management Committee (PMC). If the amount and quality of data is deemed to require an extraordinary PMC consultation, a meeting will be scheduled at the earliest convenience. Otherwise, the PMC will hold a vote in the frame of the scheduled consortium meetings.

### 3.2.2 Methods to access the data

SEAWave has chosen the Zenodo (<https://zenodo.org>) repository for storing the project data and a SEAWave project account has been thereby created. Zenodo is a repository supported by CERN and the EU OpenAire project, is open, free, and searchable and structured with flexible licensing allowing for storing all types of data: datasets, images, presentations, publications and software. Additionally:

- The repository has backup and archiving capabilities.
- The repository allows for integration with github.com<sup>3</sup> (a platform providing a free and flexible tool for code developing and storage) which could be used for storing of code generated during the project (ex. code for data analysis and automated measurement setup drivers).
- The repository can be set to restrict access to the data under embargo status until a chosen date; then the content becomes publically available automatically.
- Zenodo assigns all publicly available uploads a Digital Object Identifier (DOI) to make the upload easily and uniquely citable.

Finally, the documentation about the software needed to access the data will be included by means of a text file that will be periodically updated.

### 3.2.3 Restricted area access

If an embargo is sought to give the consortium time to publish or seek IPR protection, data will be accessible through Zenodo.org to consortium members only until the agreed embargo expiration date.

## 3.3 Making data interoperable

### 3.3.1 Data interoperability and used vocabulary

The depositors will strive for adhering to standards for formats, as much as possible compliant with available (open) software applications as from the CERIF guidelines. They will also strive for using a standard vocabulary for all data types present to allow inter-disciplinary interoperability.

---

### 3.4 Increase data re-use (through clarifying licences)

#### 3.4.1 Data licensing

The flexible licensing capability embedded in Zenodo will be leveraged to partition the repository space in an open area and a restricted access area with the aim to transfer as much data as possible to the open area at the earliest convenience. Sharing of data with restricted access will be possible only by the depositor's approval.

#### 3.4.2 Reusability at the end of the project

The data produced and/or used in the project will be useable by third parties, both during and after the end of the project as far as it is placed in the open area of the Zenodo repository. Access by third parties will be encouraged through dissemination initiatives for example by sharing the repository address and basic access instructions during conference presentations.

#### 3.4.3 Data quality assurance process

The DMP manager will periodically assess compliance of the repository entries to the preset format and content standards. The Plan is a living document whose content concerning the data management will be updated from its creation (month 6 of the project) to month 18, to the end of the project (month 36).

#### 3.4.4 Re-usability duration

The length of time for which the data will remain re-usable will not be enforced by SEAWave partners after the end of the project (unless it is deemed that further IPR protection steps need to be taken). However it is foreseeable that re-usability will depend on the demonstrated technology obsolescence.

## 4 Allocation of resources

The chosen repository (Zenodo) is free of charge for educational and informational use. While no resources were specifically devoted to making SEAWave's data FAIR, all partner institutions have budgeted dissemination costs supporting Open Access publication. Therefore, they will make sure that peer-reviewed journal article they publish is openly accessible, free of charge, according to the Grant Agreement.

For some publishers supporting a green route to Open Access of journals, special issues and conference proceedings a post-print version of the publication will be made available in the Zenodo repository. This version is after the peer-review changes have been made, but it does not typically include the publication-specific formatting. This version may also be referred to as the

---

author's final draft, accepted author manuscript (AAM) or the author's final peer-reviewed manuscript.

For example, the IEEE supports this green route to Open Access for the *IEEE Transactions on Microwave Theory and Techniques*. Green Open Access is also supported by Wiley's *Bioelectromagnetics*.

## 5 Data security

### 5.1 Data recovery

By relying on the Zenodo repository, SEAWave's research output will be stored safely in the same cloud infrastructure as research data from CERN's Large Hadron Collider and using CERN's battle-tested repository software INVENIO (a fully customized digital library framework).

All files uploaded to Zenodo are stored in CERN's EOS service in an 18 petabytes disk cluster. Each file copy has two replicas located on different disk servers.

### 5.2 Secure storage

Metadata and persistent identifiers in Zenodo are stored in a PostgreSQL instance operated on CERN's Database on Demand infrastructure with 12-hourly backup cycle with one backup sent to tape storage once a week.

### 5.3 Transfer of sensitive data

Transfer of sensitive data will occur uniquely from/to the Institutional Clouds or Repositories of the partners that the consortium has chosen for internal data storage and transfer.

## 6 Ethical aspects

The ethics aspects have been covered in the proposal and by obtaining (a) any ethics committee opinion required under national law and (b) any notification or authorization for activities raising ethical issues required under national and/ or European law.

The documents submitted upon request by the coordinator to the Commission Services may be added to the Zenodo repository.

---

## 7 Use of the DMP within the project

The plan is used by the SEAWave partners as a reference for data management (naming, providing metadata, storing and archiving) within the project each time new project data are produced.

The project partners are introduced to the DMP and its use as part of WP11 activities. Relevant questions from partners will be specifically addressed within WP11. The Work Package will also provide support to the project partners on using Zenodo as the data management tool.

## 8 List of acronyms

<b>Acronym</b>	<b>Meaning</b>
CERIF	Common European Research Information Format
DOI	Digital Object Identifier
DMP	Data Management Plan
FAIR	Findable, accessible, interoperable and re-usable
ORD	Open Research Data