

# SINNOGENES

Storage INNOvations for Green ENERgy Systems

DELIVERABLE D1.2

Data Management Plan v1

Call: **HORIZON-CL5-2022-D3-01**

Type of Action: **IA**

Project Acronym: **SINNOGENES**

Project ID: **101096992**

Duration: **48 months**

Start Date: **01/01/2023**



Co-funded by  
the European Union

## Copyright

Copyright © 2023 the SINNOGENES Consortium. All rights reserved.

The SINNOGENES Consortium consists of:

Number	Short name	Legal name
1	UNISY	UNISYSTEMS LUXEMBOURG SARL
1.1	UNIS GR	UNI SYSTEMS SYSTIMATA PLIROFORIKIS MONOPROSOPI ANONYMI EMPORIKI ETAIRIA
2	UBE	UBITECH ENERGY
3	ART	ARTELYS
4	RINA-C	RINA CONSULTING SPA
5	CIRCE	FUNDACION CIRCE CENTRO DE INVESTIGACION DE RECURSOS Y CONSUMOS ENERGETICOS
6	FBK	FONDAZIONE BRUNO KESSLER
7	MINDS	METAMIND INNOVATIONS IKE
8	CINT	CINTECH SOLUTIONS LTD
9	UGE	UNIVERSITA DEGLI STUDI DI GENOVA
10	CW	CAPWATT, S.A.
11	INESCTEC	INESC TEC - INSTITUTO DE ENGENHARIADESISTEMAS E COMPUTADORES, TECNOLOGIA E CIENCIA
12	FEUP	UNIVERSIDADE DO PORTO
13	CARTIF	FUNDACION CARTIF
14	CIEMAT	CENTRO DE INVESTIGACIONES ENERGETICAS, MEDIOAMBIENTALES Y TECNOLOGICAS-CIEMAT
15	INYCOM	INSTRUMENTACION Y COMPONENTES SA
16	FHA	FUNDACION PARA EL DESARROLLO DE LAS NUEVAS TECNOLOGIAS DEL HIDROGENO EN ARAGON
17	SCHN	SCHNEIDER ELECTRIC ESPANA SA
18	DLR	DEUTSCHES ZENTRUM FUR LUFT - UND RAUMFAHRT EV
19	SAND	SANDDORN GMBH HERZBERG
20	HEDNO	DIACHEIRISTIS ELLINIKOU DIKTYOU DIANOMIS ELEKTRIKIS ENERGEIAS AE
21	IPTO	INDEPENDENT POWER TRANSMISSION OPERATOR SA
22	UoA	ETHNIKO KAI KAPODISTRIAKO PANEPISTIMIO ATHINON
23	CERTH	ETHNIKO KENTRO EREVNAS KAI TECHNOLOGIKIS ANAPTYXIS
24	EWf	ENERGY WEB STIFTUNG (ENERGY WEB FOUNDATION)
25	TPG	TRANSPORTS PUBLICS GENEVOIS
26	UNIGE	UNIVERSITE DE GENEVE
27	Hitachi	HITACHI ENERGY SWITZERLAND LTD.



## Executive Summary

SINNOGENES project primary aim is to develop the Storage INNOvations (SINNO) energy toolkit, which is a complete framework of methodologies, tools and technologies that will enable the grid integration of innovative storage solutions, while demonstrating sustainability, technical performance, independence from cost and geographical characteristics. It will utilize successful energy storage business cases and systems at different scales and timeframes.

The present document is the first version of the Data Management Plan (DMP). Data management within a research and innovation project is essential and the DMP is the document that regulates the standards on how the data is obtained, used, and stored. Additionally, describes how the data will be available for other researchers across the world. The Data Management Plan (DMP) is the key document presenting the data management processes to be carried out in the WP1.

The SINNOGENES Data Management Plan outlines how the data collected during the SINNOGENES project will be handled, not only while the project is running but after its completion too. This deliverable also describes which standards and methodologies will be followed for the efficient collection of data and the appropriate dissemination of them.

The F.A.I.R. (Findable, Accessible, Interoperable, Re-usable) Data Management guidelines, outlined by EC for Horizon Europe projects, along with EC's DMP template, were followed for the formation of this deliverable. A first version of the project's DMP (*Data Management Plan* as a deliverable) will be submitted covering the first 6 months of the project. The DMP should be updated over the course of the project, whenever significant changes arise, such as new data and/or changes in consortium policies.

The Data of the SINNOGENES project will be generated via accessible information, project partners' contributions, external evaluators' activities, third party beneficiaries and the demonstration cases.

All the sets of data produced will be anonymized for impact assessment and research purposes. All the personal data will be treated in line with the REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016 and it will be accessible for authorized users only (validated via authentication processes).

This report illustrates the data handling framework within the SINNOGENES project.



## Table of Contents

<b>Disclaimer</b> .....	<b>2</b>
<b>Copyright</b> .....	<b>3</b>
<b>Executive Summary</b> .....	<b>4</b>
<b>Table of Figures</b> .....	<b>6</b>
<b>Table of Tables</b> .....	<b>6</b>
<b>Table of Acronyms</b> .....	<b>6</b>
<b>1 Introduction</b> .....	<b>7</b>
1.1 Purpose of the document.....	7
1.2 Target Audience .....	7
<b>2 Data Management in Horizon Europe</b> .....	<b>8</b>
<b>3 SINNOGENES Data Management Summary</b> .....	<b>10</b>
3.1 Purpose of the data collection/generation.....	10
3.2 Types and formats of data generated/collected. ....	10
3.3 Data Protection .....	11
3.4 Artefacts and Access Rights. ....	11
3.5 Existing data re-used. ....	13
3.6 Expected size of the data .....	13
<b>4 FAIR data</b> .....	<b>14</b>
<b>4.1 Making data findable, including provisions for metadata</b> .....	<b>14</b>
4.1.1 Discoverability of the data .....	14
4.1.2 Data identification mechanisms .....	15
4.1.3 Naming conventions used.....	15
4.1.4 Approach towards search keywords.....	15
4.1.5 Clear versioning of the documents.....	16
4.1.6 Standards for metadata creation (if any) .....	16
<b>4.2 Making data openly accessible</b> .....	<b>18</b>
<b>4.3 Making data interoperable.</b> .....	<b>20</b>
4.3.1 Interoperability of data assessment .....	20
4.3.2 Vocabulary use .....	21
<b>4.4 Making data re-usable.</b> .....	<b>21</b>
4.4.1 Increase data re-use through clarifying licenses. ....	21
4.4.2 Data quality assurance process .....	21
4.4.3 Length of time for which the data will remain re-usable.....	21
<b>5 Allocation of resources</b> .....	<b>22</b>
5.1 Data management responsibilities.....	22
5.2 Cost of potential value of long-term preservation.....	22
<b>6 Data Security</b> .....	<b>23</b>
<b>7 Ethical Aspects</b> .....	<b>24</b>
7.1 Confidentiality.....	25



<b>7.2</b>	<b>Personal Data</b> .....	<b>25</b>
<b>7.3</b>	<b>Intellectual Property Rights (IPR)</b> .....	<b>25</b>
<b>8</b>	<b>Conclusions</b> .....	<b>26</b>

## Table of Figures

<i>Figure 1: Template to be used for project documentation metadata overview.</i> .....	17
<i>Figure 2: Open access to scientific publication and research data in the wider context of dissemination and exploitation</i> .....	18

## Table of Tables

<i>Table 1: Clarifications of terms</i> .....	9
<i>Table 2: Types of Data Generated/Collected</i> .....	10
<i>Table 3: Artefact types</i> .....	11
<i>Table 4: Proposed Document History Table</i> .....	16
<i>Table 5: Document History table example</i> .....	16
<i>Table 6: Datasets accessibility</i> .....	19
<i>Table 7: Accessibility details</i> .....	20

## Table of Acronyms

Acronym	Definition
<b>DMP</b>	Data Management Plan
<b>EC</b>	European Commission
<b>ERC</b>	European Research Council.
<b>EU</b>	European Union
<b>FAIR</b>	Findability, Accessibility, Interoperability, Re-usability
<b>HE</b>	HORIZON Europe
<b>RI</b>	Research Item
<b>WP</b>	Working package



# 1 Introduction

Research and innovation projects like SINNOGENES produce extensive sets of data. The origins of these data could extend from social science research and laboratory testing to field studies and observations. However, following the data analysis and the project completion, this volume of data usually remains unused. So, it is important that the researchers working on the SINNOGENES project should think about the data that they will produce and ways to be exploitable by other researchers across the globe. This is the purpose of the Data Management Plan (DMP).

## 1.1 Purpose of the document

The Data Management Plan (DMP) aims to provide a strategy for managing data generated and collected during the project while optimizing access to these research data. The DMP is a document that evolves through the progress of the project and will continuously outline how the SINNOGENES research data will be handled during and after the project, and it will be reviewed and updated at regular intervals.

The DMP describes the types of research data that will be generated/collected during the course of the project, the standards that will be used, methods that the research data will be preserved and ways that the datasets will be shared for verification or re-use.

SINNOGENES DMP covers:

- Data handling during and after the project.
- Types and formats of data that will be generated/collected.
- Methodologies and standards that will be applied.
- How the data will be openly accessible.
- How the data will be safe and secure.

It also includes an overview of the datasets that will be produced by the project. The next versions of the DMP will get into more detail and describe the practical data management procedures implemented by the SINNOGENES project.

## 1.2 Target Audience

The target audience for this deliverable is:

- European Commission.
- EU Parliament.
- Partners and Advisory Group in the SINNOGENES project.
- Horizon Europe projects and other energy related projects (clustering activities).
- Organizations and specialists involved in the SINNOGENES demonstration cases.
- Relevant public and private organizations.



## 2 Data Management in Horizon Europe

According to the European Research Council (ERC) established by the European Commission (EC):

*“Under the Horizon Europe all projects that generate research data have to submit a DMP (at the latest, six months after the start of the project), deposit such data in a ‘trusted’ repository and provide access to them, under the principle ‘as open as possible, as closed as necessary’. There are also a number of requirements concerning the bibliographic and administrative metadata of deposited data, which also have to be made openly accessible to enhance findability and facilitate reuse.”*

The DMP offers an analysis of the main features of the data management policy that will be used by the grantees regarding all the datasets that will be generated by the project.

The scientific research data should be accessible, intelligible, and assessable, interoperable, and useable beyond the original purpose of collection.

Participation in the Open Research Data Pilot means that the research data which generated for the project have to be saved in a research data repository where any interested third party -without cost- could access, exploit, reproduce and disseminate:

- the data required to validate the results presented in scientific publications (including metadata).
- other data as specified and within the deadlines set in the data management plan.

The SINNOGENES DMP also follows the Guidelines on FAIR Data Management in Horizon Europe, released by the European Commission Directorate - General for Research & Innovation. This Horizon Europe FAIR DMP template has been designed to be applicable to any Horizon Europe project that produces, collects, or processes research data. According to these guidelines the management of data should be based on basic principles, which determine how research outputs should be processed so that they can be more easily accessed, understood, exchanged, and re-used. This means that data must be findable, accessible, interoperable, and re-usable.



Table 1: Clarifications of terms

<b>Research data</b>	Research data is the evidence that support research conclusions and includes data that have been collected, observed, produced, or obtained from commercial, government or other sources, for subsequent analysis and synthesis to produce original research results. These results are used to produce research papers and submitted for publication.
<b>Open research data</b>	Openly accessible research data can typically be accessed, mined, exploited, reproduced, and disseminated, free of charge for the user.
<b>Secondary data</b>	Secondary data are data that already exist, regardless of the research to be conducted.
<b>Open access</b>	Online access to research outputs provided free of charge to the end-user.
<b>Metadata</b>	Metadata is data that provide information on other data which can make finding and/or working with these data easier.
<b>Research data repositories</b>	Database infrastructures that are set up in order to manage, share, access, and archive research datasets.

A template with the FAIR principles is provided by EC as a guideline. The template represents the set of questions that someone should answer with a level of detail appropriate to the project.

The template proposes the following issues to be addressed:

- Data Summary
- FAIR data
- Allocation of resources
- Data security
- Ethical aspects
- Other issues

Each of these issues has its own sections with matters that has to be addressed. The first version of DMP is not required to provide answers to all the issues, however as DMP is intended to be a living document, more detailed information can be made available in later versions.





### 3 SINNOGENES Data Management Summary

A Data Management Plan is a key element to ensure data is well managed. Thus, below we will firstly identify the type of data that will be produced in the context of the project:

- **Data generated from accessible information** such as publications on relevant sectors of the power industry area, developments in the storage technologies or other areas related to the objectives of the project.
- **Data produced from project partners' activities**, such as expert meetings, evaluation reports, reviews, processing of technical data, development of the business process, demonstration sites and other related activities in order to achieve project goals.
- **Data obtained from third party beneficiaries and use cases**, such as details of projects that are submitted under any conditions, interviews, and presentations.

#### 3.1 Purpose of the data collection/generation

During the SINNOGENES project, data will be generated, collected, stored, and processed by project partners, stakeholders, and third-party participants. The purpose is the exploitation of these data for research, impact assessment and evaluation of deliverables/proposals.

#### 3.2 Types and formats of data generated/collected.

An overview of the different data sets that will be produced during the SINNOGENES project is shown in the table below. It includes the data type, the data origin, the related WP number, and the format in which the data will be stored.

Table 2: Types of Data Generated/Collected

Data type	Origin	WP	Format
Associates contacts collection	Openly available	1,6	.xls, .pdf, .txt
Quantitative data	Raw data	1,2,3,4,5,6	.xls, .csv, .txt, .pdf
Expert interview data	Primary data	1,2,6	.doc, .xls, .csv, .txt, .pdf
Demonstration generated data	Raw data	3,4,5	.xls, .doc, .csv, .txt, .pdf, .xml
Workshops data	Primary data	2,3,4,5	.doc, .xls, .pdf
Validation data	Primary data	3,4,5,6	.xls, .csv



### 3.3 Data Protection

All data generated during the SINNOGENES project will be anonymised for impact assessment and research purposes. Personal data of the participants will be gathered only for the project's submission process, and their informed consent will be required. All personal identifiers (known only by project partners who safekeep these records) will be encrypted and stored securely.

The personal data protection processes will follow the REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and until valid, the repealing Directive 95/46/EC (General Data Protection Regulation). The relation of applications will be coded and will be available for external evaluators with such a coding. In case data needs to be transferred to non-EU partners, we will obtain approvals from the Data Protection Office.

### 3.4 Artefacts and Access Rights.

In the first months of the project, partners asked for their input of artefacts that they will contribute to the project. Table 3 outlines the artefact contribution of the partners and whether the artefact is publishable. It should be noted that all future additions and updates will be documented under the relevant deliverables.

Table 3: Artefact types

Partner	Artefact Type	Artefact	Publishable/ Non-publishable
FBK	Research Item	Use cases IEC-62559 descriptions	Publishable
SCHN	Software	New functionalities related with Remote Terminal Units. Possibility to include the RTU utilization in a joint paper with Demo3 partners	Publishable
CERTH	Software	Digital twin of Ikaria demo, T5.2, in association with IPTO, HEDNO and CIRCE.	The infrastructure of the digital twin is publishable. However, this is part of a non-public deliverable, and the data should be handled carefully.
CERTH	Software	Life Cycle Assessment (LCA) and Life Cycle Cost (LCC) analysis of the SINNOGENES solution with the use of CERTH's VERIFY platform, T6.2.	Publishable.
CERTH	Software	Tools for management of transport fleets and synergy with the energy networks, T5.1.	The tools' methodology is publishable. However, this is part of a non-public deliverable, any data, should be handled carefully.
CARTIF	Research Item	Deliverable D3.3	Publishable
CARTIF	Research Item	Research papers	Publishable
CARTIF	Research Item	Evaluation methodology	Non-publishable
CIRCE	Research Item	Deliverable D2.1	Publishable
CIRCE	Research Item	Flexibility questionnaire	Non-publishable
CIRCE	Software	WP5 Optimisation and Flexibility algorithm	Non-publishable



CIRCE	Dataset	Operation costs and emissions for generation units	Non-publishable
HEDNO	Research Item	Papers regarding the Greek demo (Ikaria)	Publishable
HEDNO	Research Item	Internal report for load flow and protection studies from past years	Non-Publishable
HEDNO	Dataset	Time-series measurements (15 minutes resolution) - for local generation / production in the Greek demo	Non-Publishable
HEDNO	Dataset	Photos, maps of the greek demo topology / screenshots of legacy SCADA system	Non-Publishable
IPTO	Dataset	Excel file with static elements of greek demo to serve as input for the INTEMA platform (Task 5.2)	Non-Publishable
IPTO	Research Item	First draft of deliverable D5.2 with thorough description of Greek demo and its electrical topology	Publishable
UBE	Software	SINNOGENES IT architecture and relevant components	Publishable
UBE	Research Item	Deliverables D1.2, D1.4, D2.2, D2.4, D5.2, D6.1	Publishable
UNISY	Research Item	Deliverables D1.1, D1.4	Publishable
INYCOM	Research Item	Deliverable D3.1, D3.4	Publishable
CIEMAT	Research Item	Scientific papers in international Journals and Conferences	Publishable
CIEMAT	Research Item	Deliverable D3.2	Publishable
INESCTEC	Research Item	Deliverable D4.1, D4.4	Publishable
DLR	Research Item	Deliverable D4.2	Publishable
SAND	Research Item	Deliverable D4.3	Publishable
UoA	Research Item	Deliverable D5.3	Publishable
ART	Research Item	Deliverable D6.7	Publishable
RINA-C	Research Item	Deliverable D6.8	Publishable
UGE	Software	Contribution to the Optimal Steady State Energy Dispatch Tool (T4.1)	Non-publishable
UGE	Dataset	Reconstructed profile of Italian Hydro power plant (T5.2)	Publishable
UGE	Dataset	Results of the Replicability Studies (T6.3)	Non-publishable
UGE	Research Item	Dissemination papers (T6.3)	Publishable

It is clear from the table above that the partners' contribution of artefacts consists of:

- publishable research items (reports, deliverables, papers).
- non-publishable RIs (evaluation methods, questionnaires).
- non-publishable Datasets (mainly from Demo sites).
- publishable Software with limitations sharing specific data of the project.



### 3.5 Existing data re-used.

Data will be collected throughout the project. No data that has been collected in previous research will be used for the SINNOGENES project. Also, sharing data with non-EU members who are not part of the consortium is not anticipated.

### 3.6 Expected size of the data

The SINNOGENES project will generate research datasets (i.e., results of the technologies, services of the demos, etc.), publications, new services proposal through the Open Call, dissemination material, etc. Due to the complexity of the project and the scope of the work, the expected size cannot be estimated at this point.



## 4 FAIR data

SINNOGENES project follows the FAIR principles and supports the re-use of the research data generated throughout the lifespan of the project.

FAIR, through the set of guidelines that offer, can help to make the data:

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

### 4.1 Making data findable, including provisions for metadata.

This document initiates the data management plan to support the effective collection and integration of the SINNOGENES data. Storage, processing and sharing (among project participants) will occur via data exchange platforms (i.e., Microsoft365), while interaction with the public will be achieved through the official project website.

Furthermore, data will be stored at the coordinator's Dropbox repository and will be kept for a minimum of 5 years after the end of the project.

A naming agreement will include a brief description of contents, the institution collecting the data and when was published.

Version numbering will only be an issue if a participant requests withdrawal of their data in which case a version number will be added to the filename.

Data will be anonymized meaning that data will not identify any individuals and so real names of participants will NOT be distributed. Data will be shared only in relation to publications (deliverables and papers). As a result, the publication will serve as the main piece of metadata for the shared data. When this is not seen as being sufficient for the comprehension of the raw data, a report will be shared along with the data, explaining their meaning and methods of acquisition.

#### 4.1.1 Discoverability of the data

It is essential to integrate data from the participants' activities in the open calls so that the datasets to be reused. Considering the FAIR data principles (meta)data should:

- have a unique and persistent identifier.
- contain enough metadata for the interpretation of the data.
- be filed in a searchable source.

Following these principles, not only the data but also their authentication and authorization details will be retrievable.



#### 4.1.2 Data identification mechanisms

All the documents associated with the project will be identified with the project's name and a unique, continuing document type. Versioning of the document should be part of the document name and title.

Documents related to project activities (i.e., WP activities, deliverables), will use the corresponding number (WP or deliverable) followed by a short title.

For example:

- SINNOGENES\_D10.1-Project Web Portal-v0.1.pdf
- SINNOGENES\_D1.2-Data Management Plan-v0.1.pdf

#### 4.1.3 Naming conventions used.

Each set of data produced (dataset, deliverables, etc.) will be named in a standardized way and will include a table with the record of the versions of the specific document.

Suggestions for the project's documents naming are below:

- Choose easily readable identifier names (short and meaningful).
- Do not use acronyms that are not widely accepted.
- Do not use abbreviations or contractions.
- Avoid Language-specific or non-alphanumeric characters.
- Add a two-digit numeric suffix to identify new versions of one document.
- Dates should be included back to front and include the four-digit years: YYYYMMDD.

Deliverables: **SINNOGENES\_[Deliverable Code]-[Deliverable Title]\_[Partner]-vA.BB** i.e.: SINNOGENES\_D1.1-Project Management Handbook-v1.00 (*for submission to the Commission*)

For datasets: **WP [Work Package number] P [Pilot number; pilot activity number] - [description of the activity]** i.e.: WP4 P1.4 Results of demonstration performance.

#### 4.1.4 Approach towards search keywords

Documents related to the project's activities will be produced by using templates agreed by the consortium. These templates include a keywords section to make documents findable.

The information submitted by the applicants to the open calls will use keywords related to the topics covered by the SINNOGENES such as:

- Combination of Energy Storage technologies
- Advantages of grid-connected Energy Storage
- Flexibility Markets
- Innovative Energy Storage solutions
- Market Design
- SINNO energy toolkit architecture
- Etc.



The keywords are used in order to easily identify documents related to the project and will be used throughout the submission process.

#### 4.1.5 Clear versioning of the documents

Only documents created by the consortium will be versioned, for this purpose templates include 3 descriptors to identify the versions and status of the documents: Moreover, partners, following the recommendations included in section “Naming conventions” will identify the different versions by using a two-digit number following the descriptor Draft. A document reviewed by another partner should be returned to the principal author by including rev + acronym of the organization.

Only the principal author will change the draft number and will add the word FINAL to documents ready to be sent to the EC.

Table 4: Proposed Document History Table

Version	Date	Change History	Author(s)	Organisation
0.1	Xx/yy/zzzz	Initial version	Partners Name	Organisation

The document history included in the document template should be filled in as follows:

Table 5: Document History table example

Version	Date	Change History	Author(s)	Organisation
0.1	Xx/yy/zzzz	Initial version	Mr A	XYZ
0.2	Xx/yy/zzzz	Under review	Mrs B	ABC
0.3	Xx/yy/zzzz	Draft	Mr A	XYZ
0.4	Xx/yy/zzzz	Under review	Mrs B	ABC
	Xx/yy/zzzz	Issued	Mr A	XYZ

#### 4.1.6 Standards for metadata creation (if any)

Basic metadata will be used to facilitate the efficient retrieval of information by project partners and external evaluators. As a result, all documents related to the project have to include in the front-page information about the author(s), the editor(s), the WPs, the dissemination level, and the version.



**SINNOGENES**  
Storage **INNO**vations for Green **ENERgy** Systems

DELIVERABLE Dx.x  
DELIVERABLE TITLE

Call: **HORIZON-CL5-2022-D3-01**

Type of Action: **IA**

Project Acronym: **SINNOGENES**

Project ID: **101096992**

Duration: **48 months**

Start Date: **01/01/2023**

 Co-funded by the European Union

Figure 1: Template to be used for project documentation metadata overview.





## 4.2 Making data openly accessible.

Data will be made available subject to ethics and participant agreement. However, the personally identifiable nature of the data collected within SINNOGENES means that in most cases it would be difficult to release collected data. All the available data will be stored in the coordinator's Dropbox repository. Whoever requests access should contact the Project Coordinator describing their intended use of a dataset.

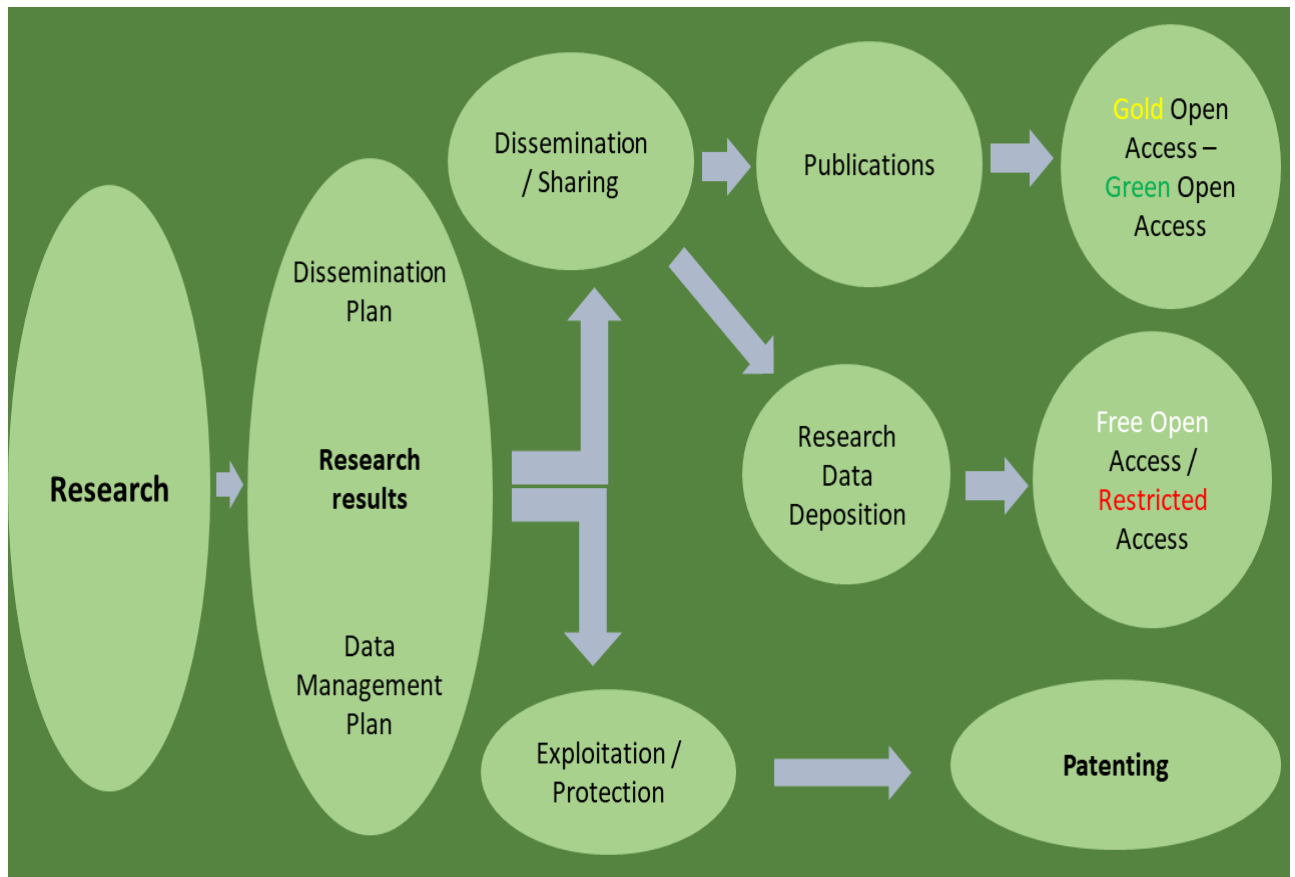


Figure 2: Open access to scientific publication and research data in the wider context of dissemination and exploitation

The Project Coordinator will send a terms and conditions document for them to sign and return. Upon return, the dataset will be released. Documentation will be included with the release of the data.

The following table is highlighting:

- which data are produced and used in the project.
- which data will be made openly available.



Table 6: Datasets accessibility

	Data Type	Data openly available	Justification
1	Stakeholder contacts collection	No	The data related to the stakeholder's contacts will not be published as primary data due to privacy and security concerns.
2	Public security services collection	Yes	
3	Digital security solutions collection	Yes	
4	Quantitative survey data	Yes	
5	Expert interview data	No	The data from the expert interviews (recordings, protocols, and transcriptions) will not be published as primary data due to privacy and security concerns. Anonymization is not considered as an alternative, because the sample size allows drawing conclusions on the respondents.
6	Focus groups data	No	The data from the expert interviews (recordings, protocols, and transcriptions) will not be published as primary data due to privacy and security concerns. Anonymization is not considered as an alternative, because the sample size allows drawing conclusions on the respondents.
7	Workshops data	No	The data from the expert interviews (recordings, protocols, and transcriptions) will not be published as primary data due to privacy and security concerns. Anonymization is not considered as an alternative, because the sample size allows drawing conclusions on the respondents.
8	Validation cycles data	No	The data from evaluation survey will not be published due to privacy and security concerns. Anonymization is not considered as an alternative, because the sample size allows drawing conclusions on the respondents.

As indicated above, the following data sets will be made openly accessible: Data type 2 (Public security services collection) 3 (Digital security solutions collection) and 4 (Quantitative survey data).

The following table describes the accessibility details of these datasets.



Table 7: Accessibility details

	Data Type	Accessibility	Availability and required software tools	Information on metadata and additional data information
2	Public security services collection	Public	Filterable and searchable database; can be accessed with a state-of-the-art web browser	No metadata needed; additional information will be provided on the platform
3	Digital security solutions collection	Validated professionals	Filterable and searchable database; can be accessed with a state-of-the-art web browser	No metadata needed; additional information will be provided on the platform
4	Quantitative survey data	Public	Cleaned primary data; can be accessed with SPSS or Excel	No metadata needed; additional information will be provided on the platform

No specific software tools will be needed to access any of the data. Anonymized data sets will be saved and stored in word, pdf or excel to facilitate its exploitation and guarantee their long-term accessibility.

All Data will be stored and secured on Microsoft365 platform and on coordinator's Dropbox account.

### 4.3 Making data interoperable.

Making data interoperable means that both data and metadata must be read by automation and have to use a consistent terminology.

#### 4.3.1 Interoperability of data assessment

All partners will be responsible for storing the data in a specific user-friendly format, taking into consideration the real and current needs of the possible practitioners interested in using, integrating, or exploiting the data generated during the project. The assessment of data interoperability will be updated in future reviews in order



to guarantee the SINNOGENES data are suitable for the needs of a broad spectrum of researchers.

#### 4.3.2 Vocabulary use

The vocabulary used in the project is a very standard and common language within the business creation culture and logistics. Vocabulary won't represent any barrier to data interoperability and/or re-use.

### 4.4 Making data re-usable.

#### 4.4.1 Increase data re-use through clarifying licenses.

Due to the sensitive nature of the data, they will be available on the application/Microsoft365 platform/shareportal and their use will be restricted to the research use of the licensee and colleagues on a need-to-know basis. This non-commercial license is renewable after 2 years, data may not be copied or distributed and must be referenced if used in publications. These arrangements will be formalized in a User Access Management license which describes in detail the permitted use of the data.

#### 4.4.2 Data quality assurance process

The project coordinator will be responsible for assuring the quality of the data by making sure the datasets follow the FAIR principles included in this DMP, and that the data is updated.

Personal data will be processed following the EU, national and international laws and considering the "data quality" principles listed below:

- Accurate and up to date.
- Processed securely, fairly, and legitimately.
- Processed in line with data subjects' rights.
- Kept for no longer that required and only for the purpose of the project.

Data quality assurance process will be led in accordance with the REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data.

#### 4.4.3 Length of time for which the data will remain re-usable.

The Consortium's goal is to maintain data re-usable for as long as it is possible after the end of the project. A first period of 5 years has been established. However, this time can be extended, subject to the partners' agreement. This period can vary depending on the value of the data after the end of the project.



## 5 Allocation of resources

### 5.1 Data management responsibilities

Data will be stored at the Collaboration portal (Microsoft365), set by the Coordinator as the project's repository, and will be kept for 5 years after the end of the project. The Coordinator is responsible for managing the repository for SINNOGENES including all data management issues related to the project.

The analyses of the observed research data will be presented in the form of publications, which the consortium will publish in open access scientific journals. The costs related to open access will be claimed as part of the Horizon Europe grant.

Each WP leader will be responsible for the storage of the data resulting from the project's activities and the uploading of them in the SINNOGENES Microsoft365 web portal, or other storage systems for sharing the information of the project.

The SINNOGENES coordinator supported by the WP leaders will be responsible for the constant update of this document and for developing a strategy to encourage:

- identification of the most-suitable data-sharing and protection methods.
- efficient use of data, assuring clear rules on its accessibility.
- quality of the data.
- safe storage in a user-friendly interface.

### 5.2 Cost of potential value of long-term preservation

The cost of data storage and maintenance are not going to require extra funding once the project ends. As per the value of the data, it is important to consider that the issues covered by the project act in response to a current need. Thus, the data from this project will have a direct impact in the future but might not be of relevance as the challenges are being tackled or replaced by other priorities.



## 6 Data Security

In order to secure the processing of personal data against publishing to unauthorized people, processing in violation of the law, change, loss, damage, or destruction of the SINNOGENES data exchange platform (SharePoint-Microsoft365) practices measures such as:

**Physical protection:** Only a limited number of essential personnel can gain access to datacenters. Their identities are verified with multiple factors of authentication, including smart cards and biometrics. There are on-premises security officers, motion sensors, and video surveillance. Intrusion detection alerts monitor anomalous activity.

**Network protection:** The networks and identities are isolated from the Microsoft corporate network. The service is administrated by a dedicated Active Directory domain, separate domains for test and production, and the production domain is divided into multiple isolated domains for reliability and security.

**Application security:** Engineers who build features follow the security development lifecycle. Automated and manual analyses help identify possible vulnerabilities. The Microsoft security response center helps triage incoming vulnerability reports and evaluate mitigations.

**Content protection:** Data is encrypted at the disk level using BitLocker encryption and at the file level using keys.

**Data Transit:** When data transits into the service from clients, and between datacenters, it's protected using best-in-class encryption.

**Recoverable:** Datacenters are geo-distributed within the region and fault tolerant. Data is mirrored in at least two datacenters to mitigate the impact of a natural disaster or service-impacting outage.

**Audited and Compliant:** Regulatory compliance is fundamental to Microsoft 365. The service complies with regulatory and compliance standards.



## 7 Ethical Aspects

The SINNOGENES consortium is aware of ethical, privacy, copyright and data protection issues that might be triggered by the activities of the project. Consequently, this section revolves around ethical and legal compliance issues such as:

- the consent for data protection and sharing.
- the protection of the individuals' identity and the ways that the personal data will be handled.
- the safe storage and secure transfer of data.
- copyright and intellectual property rights (IPR) issues.

To deal appropriately with the aforementioned issues the SINNOGENES consortium scrutinized the legal framework that governs the activities of the project i.e. a) the European Union ePrivacy Directive 2002/58/EC, b) the Directive 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data thus repealing Council Framework Decision 2008/977/JHA and c) the (EU) Regulation 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data thus repealing Directive 95/46/EC.

Additionally, the consortium studied the ethical guidelines for research projects in the EU under Horizon Europe which must comply with ethical principles and relevant national and EU legislation i.e., a) HE Framework Programme Regulation 2021/695: Eligible actions and ethical principles (Article 18) and Ethics (Article 19), b) HE Model Grant Agreement: Ethics (Article 14 and Annex 5) c) EU Grants: HE Programme Guide: Guidelines on serious and complex ethics issues (July 2021). d) EU Grants: HE Programme Guide: How to complete your ethics self-assessment (July 2021).

The SINNOGENES consortium will act in accordance with the above-mentioned legislation and Horizon Europe ethical principles as it relates to any individual that might be involved in the project either as a participant or not.

Work package 1 (Deliverable 1.4) of the SINNOGENES project deals with the ethics requirements in relation to the project's objectives, methods, processes, tasks, and results. These ethics requirements mainly relate to the processing of personal data and with data protection. Therefore, in this deliverable will not be covered requirements that are dealing specifically with the procedures in respect to the data protection.

In order to ensure that all ethical aspects are considered and that the SINNOGENES project is compliant with all legal requirements and ethical issues, a general strategy has been designed by the Ethics Requirements that will be defined by deliverables from WP1.

Ethics Requirements that were set by the EC, during the ethics check process before the signature of the GA, have been considered by the consortium. General guidelines



have been provided to the project and the consortium on all aspects such as data protection, privacy issues, research participants' safety, etc.

### 7.1 Confidentiality

SINNOGENES partners must retain any data, documents, or other material as confidential during the implementation of the project. Further details on confidentiality can be found in Article 13 of the Grant Agreement.

### 7.2 Personal Data

Part of this project will be the collection of personal data. This data will be stored, analysed, and used anonymously. The individuals will be informed about the use of their information, and they will have to agree to the data collection while providing their approval in the form of written consent. The identity of any individual interviewed or in any other way engaged in the project (e.g., by email, correspondence, newsletter) will be protected by anonymization of the data.

### 7.3 Intellectual Property Rights (IPR)

SINNOGENES embarks to ensure that data access and sharing activities will be meticulously implemented in compliance with the privacy and data collection rules and regulations, as they are applied nationally, internationally and in the HORIZON Europe rules. Primary data collected through the interviews from sources outside of the consortium may be available either to all or specific partners upon authorization of the interviewees. This kind of data will not be available to the public. Concerning the results of the project, these will become publicly available based on the Access Rights as described in the Consortium Agreement (Article 16).





## 8 Conclusions

The first data management plan describes the collection and handling of the project's data along with information on the importance of FAIR data and its application. Security measures and ethical aspects of the project are addressed at an initial level. It should be noted that the DMP is a living document that will be continuously updated by all relevant partners throughout the project's duration.

