



D1.4

# Data management plan

Project title: **Enhancing Integration and Interoperability of  
CCAM eco-system**

Project acronym: **IN2CCAM**

Horizon Research and Innovation Actions  
Project No. 101069573  
Call HORIZON-CL5-2021-D6-01



<b>Dissemination level</b>	Public (PU) - fully open
<b>Work package</b>	WP1 - Project Management and Quality Assurance
<b>Deliverable number</b>	D1.4 - Data Management Plan
<b>Status - Version</b>	Final – V1.0
<b>Submission date</b>	30/04/2023
<b>Keywords</b>	Data management

#### Quality Control

	Name	Organisation	Date
Peer review 1	Mónica Díaz de Mendivil	VICOM	24/04/2023
Peer review 2	Jason Sioutis	ICCS	26/04/2023
Quality reviewer	Marie-Laure Watrinet	LIST	14/04/2023

#### Version history

Version	Date	Author	Summary of changes
0.1	9/01/2023	Eva Rodrigues, Rita Santiago	First draft of the document
0.2	24/02/2023	Eva Rodrigues, Rita Santiago	Ethics, Privacy & Security Considerations, Data Archiving & Preservation, Data Management, Sharing & Open Access

0.3	26/04/2023	Eva Rodrigues, Rita Santiago	Update after peer-review
0.4	26/04/2023	Eva Rodrigues, Rita Santiago	Finalisation before generation of 1.0
1.0	26/04/2023	Eva Rodrigues, Rita Santiago	Final version 1.0

## **LEGAL DISCLAIMER**

The IN2CCAM project is funded by the European Union. However, the views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or CINEA. Neither the European Union nor the granting authority, CINEA, can be held responsible for them.

The IN2CCAM project Consortium members shall have no liability for damages of any kind, including without limitation direct, special, indirect, or consequential damages that may result from the use of these materials subject to any liability which is mandatory due to applicable law.

© 2022 by IN2CCAM Consortium.

## TABLE OF CONTENTS

<b>1</b>	<b>EXECUTIVE SUMMARY</b>	<b>8</b>
<b>2</b>	<b>INTRODUCTION</b>	<b>9</b>
2.1.	Project intro	9
2.2.	Purpose of the deliverable	9
2.3.	Intended audience	10
2.4.	Structure of the deliverable and its relation with other work packages/deliverables	10
<b>3</b>	<b>ETHICS, PRIVACY &amp; SECURITY CONSIDERATIONS</b>	<b>11</b>
3.1.	Ethics Requirements	11
3.2.	Data Privacy and Personal Data	13
3.3.	Data Security	13
<b>4</b>	<b>DATA ARCHIVING &amp; PRESERVATION</b>	<b>14</b>
<b>5</b>	<b>DATA MANAGEMENT, SHARING &amp; OPEN ACCESS</b>	<b>15</b>
5.1.	Data Handling and Management	16
5.2.	Sharing and Open Access	17
<b>6</b>	<b>FAIR DATA</b>	<b>17</b>
6.1.	Findability	17
6.2.	Accessibility	18
6.3.	Interoperability	19
6.4.	Re-usability	20
<b>7</b>	<b>CONCLUSION</b>	<b>21</b>
<b>8</b>	<b>REFERENCES</b>	<b>22</b>

## LIST OF TABLES

<i>Table 1: Tool to characterise data curation and preservation.....</i>	<i>15</i>
<i>Table 2: Template for data handling and management summary .....</i>	<i>16</i>

## ABBREVIATIONS AND ACRONYMS

Abbreviation	Meaning
CA	Consortium Agreement
DL	Deliverable Leader
DMP	Data Management Plan
EC	European Commission
FAIR	Findable, Accessible, Interoperable, and Reusable
GA	Grant Agreement
GDPR	General Data Protection Regulation
IPR	Intellectual Property Rights
ISO	International Organization for Standardization
KPI	Key Performance Indicator
QM	Quality Manager
WP	Work Package

## 1 EXECUTIVE SUMMARY

The Data Management Plan (DPM), deliverable D1.4 of IN2CCAM, relates to Task 1.4 – Data governance - of WP1 – Project Management and Quality Assurance. This document aims to provide information from all partners on the data used in the project, the data acquisition and transmission tools they will collect and archive from the local demonstrations in the database, and the datasets collected. The data/research outputs will be managed per the FAIR principles (Findable, Accessible, Interoperable, and Reusable). Their treatment will ensure data protection and compliance with the regulation on protecting private data (GDPR).

Intellectual Property Rights (IPR), licensing agreements, interoperability and re-use of data will be considered to comply with the principle of being “as open as possible, as closed as necessary”. Thus, the exploitation opportunities, the protection of generated IP, the confidentiality obligations, the security obligations, and the obligations to protect personal data will always be considered before making the data openly accessible.

The document is structured in four parts: firstly, a description of the ethics and privacy considerations, secondly, a description of how data will be archived and preserved, thirdly a description of the data handling and management; and finally, a description of the FAIR data.



---

## 2 INTRODUCTION

---

---

### 2.1. Project intro

---

IN2CCAM project, according to the vision of the Horizon Europe 2021-2027 framework program, aims to accelerate the implementation of innovative technologies to make it possible to include autonomous vehicles for transporting passengers and goods.

Cars, buses, and trucks without drivers but under the control of futuristic instruments based on artificial intelligence techniques, will be able to remedy human errors with significant impact on society in terms of safety (i.e., the reduction in the number of road accidents) caused by human error; the environment (i.e., reducing transport emissions and congestion by facilitating the flow of traffic and avoiding unnecessary travel); inclusiveness (i.e. ensuring inclusive mobility and good access for all as elderly or disabled people physical problems).

Twenty-one partners from 9 different European countries will collaborate under the guidance of Prof. Fanti and her team from the Automatic Controls laboratory to create a series of physical, digital, and operational solutions that will be implemented in 6 pilot cities: Tampere (Finland), Trikala (Greece), Turin (Italy), Vigo (Spain), Bari (Italy) and Guimarães (Portugal) via Associação Quadrilátero.

---

### 2.2. Purpose of the deliverable

---

The main purpose of the Data Management Plan is to describe the data management life cycle for the data to be collected, processed and/or generated by the project. The datasets identified should be according to the data management policy and the FAIR principles (Findable, Accessible, Interoperable, and Re-usable). For that reason, in IN2CCAM, a template for the dataset has been customised and shared within the consortium.

The Data Manager appointed is Sofia Almeida and will:

- Identify and describe the evaluation data that will be collected, processed or generated;
- Define what data will be made publicly available for research purposes and what sensitive data will not be shared because of IPR or confidentiality issues;
- Specify the different types of data and the methodologies and tools for collecting, storing, curating, preserving and sharing the data;
- Ensure data protection and compliance with regulation on the protection of private data (GDPR).

The current document corresponds to the first version of the IN2CCAM Data Management Plan (DMP), therefore composed of preliminary information and frameworks that will be followed soon to collect and characterise the project's data sets (which will be presented in the next iteration of the current deliverable, to be submitted by M36 – October 2025).

Moreover, the data and research outputs collected and managed by IN2CCAM are of three types:

Type 1: metadata and map data to describe the model of the living labs, including land-use data such as residential buildings, firm and school locations, and associated characteristics. Such data are existing numerical and text data from the LLs and are used to organise the test phase and LLs and used to manage the test phase and the simulations.

Type 2: data collected during the development and testing phases, i.e., data on vehicle locations, vehicle capabilities, public transportation schedules, traffic flow observations and other relevant information (buildings, points of interest, people counting, etc.) that must consider privacy, data security and cyber security requirements to enable real-time access to vehicle-generated and infrastructure-based sensing data.

Type 3: data collected during the analysis and evaluation phase both on the field and by simulation, but also during the research and development phases, regarding feedback from the authorities and end-users, via surveys and other feedback collection mechanisms.

---

## 2.3. Intended audience

---

The dissemination level of D1.4 is public (PU) but is meant primarily for (a) all members of the IN2CCAM project consortium and (b) the European Commission (EC) services.

---

## 2.4. Structure of the deliverable and its relation with other work packages/deliverables

---

The deliverable is divided into 5 sections:

- Section 1: short introduction regarding the document's purpose and the datasets identified at this stage of the project.
- Section 2: details an overview of the datasets that will be generated/collected in IN2CCAM and a presentation of the methodology that will be used to identify and characterise the different data sets that will be generated or acquired, their origin, purpose, and its relation to the objectives of the project, as well as the type and format of the data.
- Section 3: Ethics, Privacy & Security Considerations;

- Section 4: Data Archiving & Preservation;
- Section 5: Data Management, Sharing & Open Access;
- Section 6: FAIR Data
- Section 7: addressing the main highlights of the deliverable.

It is related to all work packages and deliverables of the project since it presents how all the datasets should be processed during the lifetime of the IN2CCAM project.

## 3 ETHICS, PRIVACY & SECURITY CONSIDERATIONS

All IN2CCAM consortium is fully aware of the ethical aspects of all research activities involved in the project. Ethics and privacy compliance task will ensure this. Any ethical issue that might arise during the project's lifetime will be carefully analysed and will follow the ethical rules and standards of Horizon Europe, as well as the Ethics and privacy compliance task will ensure this hose reflected in the Charter of Fundamental Rights of the European Union and the General Data Protection Regulation (GDPR) (Regulation (EU) 2016/679).

### 3.1. Ethics Requirements

The internal ethical agenda of IN2CCAM follows the guidelines of various expert communities in data ethics (e.g. the European group on ethics in science and new technologies to the European Commission). The ethical agenda includes a code of conduct to specify correct behaviour and corresponding principles to data collection and processing (see D1.6 Ethics and privacy compliance addresses). The following six principles are contemplated and addressed within IN2CCAM's ethical code of conduct:

1. Ownership - Who is the owner of the data?
2. Transaction transparency - What is the access given to the owner, and how transparent is the access? Transparent access for the individual's data must be established, and the individuals should have full and transparent access to the algorithms used to generate and aggregate the data sets.
3. Consent - Which individuals or other entities must consent to use the data? Individuals or other entities shall be explicitly informed of what personal data moves to whom, when, and for what purpose the owner of the data is.

4. Privacy - What are the efforts and measures to ensure data privacy? In terms of data processing, there must be an effort to preserve the individual's privacy.
5. Currency - If applicable, what is the financial value of the personal data, and how is that communicated to the data owner? Individuals shall be explicitly informed of any financial transactions resulting from their data.
6. Openness - How many of the aggregate data sets are freely available?

Following these principles, we ensure that the data generated and processed within the IN2CCAM is used responsibly and sustainably.

The ethical code of conduct shall further complement compliance with data protection laws and current regulations. The main aim of the code of conduct of the Riot-ES project is to reflect a principle that promotes honesty and genuine transparency in data management.

Besides the afore-mentioned six principles, any data generation/acquiring activity involving humans will be strictly held confidential during the research phase. This entails the following practical actions each project partner must follow when interacting with volunteers/individual stakeholders:

1. Explicitly inform all participants about the six principles of the code of conduct indicated above and allow them to provide their consent to the data management process. (Why is the data being collected? How is it going to be used? How long will it be stored? How can it be amended by the individual concerned?).
2. Ensure the volunteers that no personal or sensitive data will be centrally stored. Additionally, the volunteers shall be informed about the data security measures to avoid the potential identification of individuals.

The briefing and information should be conducted in the volunteers' native language. Additional information provided must include (i) a written description of the project and its goals, (ii) the project's progress and the related testing, evaluation, and validation procedures (protecting Riot-ES confidential information), and (iii) information about unrestricted disclaimer rights on their agreement.

---

## 3.2. Data Privacy and Personal Data

---

IN2CCAM project will engage demonstrations in cities from different countries, collecting multiple data sets and carrying out other data measurements in each city. Hence, the project stakeholders and the individuals affected by the data collection will vary from task to task. The collection, processing, and transmission of personal data will be analysed under the General Data Protection Regulation (EU) 2016/679 (GDPR)

Any additional regulations at the national level that do not fall under the GDPR and apply to data protection or any other sensitive information, will also be considered.

IN2CCAM project will respect and protect the privacy of all stakeholders and individuals. To be clear and transparent, the project will also follow a stakeholder management procedure, ensuring that stakeholders and individuals are fully informed about their rights, aim, and data handling. The stakeholder management procedure obliges the project partners to obtain consent whenever personal data is collected and processed, as described above, implementing relevant data handling practices and protocols to avoid the potential identification of individuals. This will include participants' data sets related to activities that use techniques such as interviews, questionnaires, and mobility data collection.

Data collected within the project will only be processed under the following preconditions:

- a) When the data subject has given her/his consent;
- b) When the processing is necessary for the performance of or the entering into a contract;
- c) When processing is necessary for compliance with a legal obligation;
- d) When processing is necessary to protect the vital interests of the data subject.

To this end, personal data collected and processed by the Riot-ES project will be anonymised and stored in such a way that it is not possible to identify the participants.

---

## 3.3. Data Security

---

For each data set (which will be classified in Table 1 and Table 2), the partners will state the provisions and measures to be implemented to ensure data security, privacy, and ethical requirements.

The secure management of information will adhere to the guidelines of relevant standards (e.g. ISO/IEC 27001 and 27002; Code of practice for information security management) to ensure the triad of cyber security:

- I. **Confidentiality** – Preventing unauthorised disclosure of information;
- II. **Integrity** – Assuring that data cannot be modified in an unauthorised manner;
- III. **Availability** – Making information available for authorised users.

## 4 DATA ARCHIVING & PRESERVATION

In this section, the description of the procedures for the long-term curation and preservation of the data are presented. The procedures will indicate how long the data shall be stored, where it will be stored, and the associated costs.

POLIBA, as project coordinator, is responsible for the general maintenance of the project repository, which is hosted in Sharepoint and includes deliverables, meetings documentation (agendas, minutes, and presentations), administrative documents (GA, CA, amendments), and any other document used for the development of the project. All partners are responsible for supporting the document management process. The server is only accessible to consortium team members.

Public deliverables will be published and curated on the project website, while internal data sets will be backed up to allow recovery for re-use and/or verification. Primary data will be archived for a minimum of years, to be determined by the data responsible/partner generating the data.

The classification system presented in Table 1 shall facilitate appropriate measures for the data curation and preservation procedures.

Table 1: Tool to characterise data curation and preservation

ID	Title of data set	Repository			Long-term preservation plan			
		Storage location	Type of repository	Reasons to use this repository	Time period	Approximated end volume	Storage repository	Associated costs for preservation

## 5 DATA MANAGEMENT, SHARING & OPEN ACCESS

IN2CCAM will implement a comprehensive data management system complying with ethics, privacy, and security considerations and facilitating swift processing between the project participants. After the data sets have been generated/acquired, the procedures of how the data is managed shall ensure trackability, transparency, and usability among the consortium partners.

In general, IN2CCAM DMP details 4 different categories of data generated or acquired in the project:

1. **Research data** – all the data necessary to evaluate the quantitative KPIs of the project and data necessary to validate the results presented in public deliverables or scientific publications. The consortium believes in and applies the concepts of open science and benefits arising from the European innovation ecosystem and economy by facilitating data reuse on a larger scale. The DMP covers the entire research data life cycle and must be consistent with exploitation and IPR requirements. Hence, research data linked to exploitable results will not be put into the open domain if they compromise its commercialisation prospects or have inadequate protection, which is also an H2020 obligation. The Project Coordinator will ensure that provisions on Scientific publications and guidelines on Data Management in H2020 are adhered to. As indicated, scientific research data should be findable, accessible, interoperable and reusable (FAIR) to ensure it is soundly managed beyond the original purpose for which it was collected.
2. **Operational and observational data** – all the data, raw data generated/acquired as well as curated data during the implementation, testing and operation of the demonstration activities), and data from qualitative activities including surveys, interviews, fieldwork data or engagement activities (observational data, such as the

one that will be collected under WP4 activities on citizen engagement). Particularly, sensitive data provided by consortium partners for the demonstration scenarios and personal data of individual stakeholders will be kept strictly confidential to protect their competitive advantage and in terms of personal data anonymised and secured to maintain compliance with GDPR.

3. **Monitoring and evaluation data** – all the data related to the monitoring of project specific KPIs. This data will be regularly reported and published in relevant repositories with respective access rights set there.
4. **Documentation, instruments and reusable knowledge** – all the data and documentation produced by IN2CCAM consortia, including specific documentation of the project and demonstration and implementation activities, such as tools, equipment, instruments, software, and underlying source code. The DMP covers the entire research data life cycle and must be consistent with exploitation and IPR requirements.

---

## 5.1. Data Handling and Management

---

IN2CCAM will ensure that all the research data generated is findable, accessible, interoperable, and reusable (FAIR), complying with the Horizon Europe Guidelines on FAIR Data Management. To implement FAIR Data Management, IN2CCAM will facilitate using a data handling system following the template presented in Table 2.

*Table 2: Template for data handling and management summary*

DATASET XX	Dataset name:
	Partner/Living Lab:
Description	
Purpose	
Existing data and/or external sources?	
Types and formats of data	
Re-use of existing data	
Origin	
Visibility	
Expected size	
Utility	
Storage	



Security & Privacy considerations

Exploitation/Dissemination

The template will serve as a guideline to detail the description, purpose, and relevance of the data sets and the methodology and collection procedure used to obtain them.

For data sets that are further used in publications and/or deliverables as a result of IN2CCAM, this will be detailed within the project's WP7 - Communication, Awareness, Dissemination and Exploitation.

---

## 5.2. Sharing and Open Access

---

IN2CCAM's results and achievements will be disseminated among the scientific community, industry, and key stakeholders. Open Access will be provided as much as possible for all peer-reviewed scientific publications. These publications will be stored in an Open Access repository during and after the project's life.

---

# 6 FAIR DATA

---

---

## 6.1. Findability

---

To make all the public information produced by the project findable for the researchers and stakeholders, the dataset used will include search keywords that will optimise possibilities for re-use. A good document nomenclature that facilitates data searching will also be provided. This naming and versioning protocol has already been defined in the D1.2 Quality management plan.

The deliverables should follow the nomenclature below:

"Project\_Name\_DN.N – Name\_vX.Y.docx".

Where:

- Project Name: Refers to the project's short name (IN2CCAM).
- DN.N: Is the deliverable number.

- Name: Refers to the name of the deliverable that should be matched exactly with the name of the deliverable as defined in the proposal.
  - The “y” in Vx.y may be used internally only to number draft versions.
  - A version should be Vx.0 only when it is submitted to the EC, before that the number should be V(x-1).y. It is changed to Vx.0 by the QM when the document is ready to be submitted and the pdf is generated.

In the context of the project, depending on the nature of the data, the following type of metadata could be used:

- Structural: Facilitates the navigation between data and resources.
- Descriptive: Identification of the data in terms of title, date, abstract, author or keywords.
- Administrative: Identifies who can access the data and the restrictions to be applied in the file, the location of the data, etc...

---

## 6.2. Accessibility

---

IN2CCAM data will be accessible through different channels that will ensure broad visibility and raise awareness of the project and its results.

For Scientific publications, IN2CCAM adopts a combined strategy for Open Access, allowing maximum impact for broad dissemination without increased cost, offering access to restricted materials. These publications will be carried out by publishing either in green or gold open access channels. Green open access, where the authors will deliver their own final published versions or will allow free and open access to them.

However, if this option is inappropriate, the project is considering reserving the budget to pay for published versions or will allow free and open access to them. However, if this option needs to be more appropriate, the project is considering reserving a budget to pay for providing public access to their papers through the gold open access.

The deliverables that have been defined in the proposal with nature “Public”, will be available through the project website after they are reviewed and approved by the EC. For the rest of the deliverables, the accessibility is limited for the consortium members in the project repository.

Deliverable D7.1 Communication & Dissemination Plan shows the different channels where the data will be accessible. This document presents the different channels used for the project, namely:

- Project website;
- Social media channels, such as Twitter and LinkedIn;
- Newsletters;
- External media;
- IN2CCAM concept image;
- Roll-up banner;
- Brochure;
- Leaflet;
- Video;
- Papers and reports.

---

### 6.3. Interoperability

---

Making data interoperable will facilitate the exchange and reuse between organisations and research institutions. To meet the required interoperability standards, IN2CCAM is analysing all datasets to have a common data and metadata capture and creation. The standards to follow are:

1. metadata will use keywords
2. metadata will contain vocabulary according to the FAIR Principles.

Therefore, the bibliographic metadata will be in standard format and include the following:

- a. the terms “European Union (EU)” and “Horizon Europe”.
- b. the name of the action, acronym and grant number.
- c. the publication date, and length of the embargo period if applicable, and a persistent identifier.

Once the project is advancing and all the information used is recognised and gathered, additional data on making information interoperable will be provided in an updated version of the Data Management Plan.

---

## 6.4. Re-usability

---

The reusability of the data within this project will be published after the submission and approval of the public deliverables and one year after the end of the Project. The open share dataset created along the project lifetime will be reusable in any project platforms: IN2CCAM website, project repository, Gitlab, among others.

Regarding the quality evaluation, we have established procedures (see D1.2 Quality management plan) to guarantee that the deliverables are submitted with a high level of quality. Each deliverable will have two organisations as peer reviewers that analyse the document in two ways: the format (keywords, table of figures, executive summary, etc.) and the most crucial art, the content (clarity, innovation, references, etc...) and, besides that also a quality review by the Quality manager.

## 7 CONCLUSION

The present deliverable D1.4 - Data Management Plan - constitutes a guideline complying with European and national legislation on the acquiring, handling, processing, and archiving data generated during the IN2CCAM project and beyond the project's end.

This deliverable is a preliminary version of the DMP for the IN2CCAM project at the time of delivery: April 2023 (M6). The IN2CCAM's project leader will regularly communicate with the other consortium members to refine and update the DMP.

The DMP will ensure that all the consortium partners are aware of the practices that shall be followed within the conduction of day activities.

## 8 REFERENCES

ISO/IEC 27002:2022, Information security, cybersecurity and privacy protection — Information security controls.

ISO/IEC 27005:2011, Information technology — Security techniques — Information security management systems.

Regulation, P. (2016). Regulation (EU) 2016/679 of the European Parliament and of the Council. Regulation (eu), 679, 2016.

Von Solms, R. (1998). Information security management (3): the code of practice for information security management (BS 7799). Information Management & Computer Security, 6(5), 224-225.